IJARSCT
ISSN: 2581-9429
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Volume 5, Issue 1, November 2025
Impact Factor: 7.67

Text to Sign Language Translator

Girish. N. Kothwal, Sameeksha C. Chavan, Neeraj R. Chavan, Harshal S. Chavan, Harsh R. Chendwankar, Rugved A. Chavan
Department of Engineering, Sciences and Humanities (DESH)
Vishwakarma Institute of Technology, Pune, Maharashtra, India

Abstract: *The Text to Indian Sign Language (ISL) Converter is a web-based application designed to facilitate communication between hearing and deaf individuals by translating English text into corresponding ISL gestures. This system utilizes pre-recorded MP4 video animations of ISL alphabets, numbers, and common words. The tool processes user input, either through text or broken-down letters, and sequentially plays the corresponding ISL videos. A fallback mechanism is implemented to handle words not available in the dataset by spelling them out using ISL alphabet videos. The system was tested for robustness and usability, with results indicating that it successfully handles a variety of text inputs, maintains smooth playback continuity, and delivers an engaging, user-friendly interface. While the system does not yet integrate linguistic adjustments for ISL syntax or grammar, it provides a simple and effective solution for digital inclusion. Future work will involve expanding the dataset, incorporating natural language processing (NLP) for sentence restructuring, and collaborating with ISL experts to ensure accuracy. The project highlights the potential of browser-based tools in bridging communication gaps for the hearing-impaired community, with significant room for future enhancements.*

Keywords*:* Accessibility technology, Indian Sign Language, ISL Animation, Sign Language, Text-to-Sign Language

I. INTRODUCTION

Effective communication is essential for inclusion and empowerment in any society. However, for individuals with hearing impairments, language barriers often restrict their ability to interact meaningfully with the larger community. While Indian Sign Language (ISL) serves as the primary mode of communication for many members of the deaf and hard-of-hearing population in India, it remains largely unfamiliar to the general public. This lack of widespread ISL fluency creates a communication divide between hearing-impaired individuals and non-signers, limiting access to education, employment, and healthcare.

India is home to nearly five million people who are deaf or hard of hearing, according to estimates from the Indian Sign Language Research and Training Centre (ISLRTC). Despite this significant population, ISL is underrepresented in technological solutions, especially when compared to other global sign languages like American Sign Language (ASL) or British Sign Language (BSL). Moreover, the limited number of trained interpreters and the cost and time associated with learning ISL further exacerbate the communication gap.

In response to this challenge, our project introduces a GIF-based Text to Indian Sign Language Converter. This web-based application is designed to provide a simple yet effective way of converting typed English text into its corresponding ISL representation using pre-recorded GIF animations. The system supports both alphabet-level spelling and whole-word gestures, offering a user-friendly translation flow for common expressions as well as arbitrary input such as names or paragraphs. This initiative prioritizes accessibility, simplicity, and scalability, aiming to make ISL communication more approachable for everyday users.

II. LITERATURE REVIEW

The domain of text-to-sign language translation has gained increasing attention in recent years, particularly with advancements in artificial intelligence, computer vision, and natural language processing. While much of the existing

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, November 2025
Impact Factor: 7.67

research focuses on sign languages like ASL and BSL, comparatively fewer systems address the complexities and regional nuances of Indian Sign Language (ISL).

Kahlon and Singh [4] present a systematic review of text-to-sign language machine translation, categorizing approaches such as rule-based, example-based, and statistical models. Their study highlights the challenges of symbol-based representation and the need for expressive gesture synthesis in real-time systems. Similarly, Kamal and Hassani [2] explore low-resource sign languages, emphasizing the importance of standardized datasets and intermediate representations like HamNoSys and SiGML.

Several works utilize 3D modeling and avatars for dynamic sign visualization. Yadav et al. [9] propose a system named Listen, which integrates speech recognition, natural language processing, and avatar animation for ISL generation. Their model introduces sentence restructuring from Subject-Verb-Object (SVO) to Time-Subject-Object-Verb (TSOV), aligning with ISL grammar rules. The authors report an impressive 99.21% accuracy, showcasing the feasibility of high-fidelity avatar-driven translation models.

Martino et al. [6] and Esselink et al. [5] both highlight the role of neural machine translation (NMT) in transforming text into sign sequences. They stress the need for grammatically-aware language models and gesture rendering that respects the fluidity of natural sign expressions.

Additionally, Ahinsa et al. [8] and Shenoy et al. [7] delve into dictionary construction and ISL grammar modeling, respectively. Both works contribute foundational knowledge on synthetic ISL datasets and rule-based systems, using annotation languages such as HamNoSys. Unlike their approaches, which rely on syntactic parsing and formal intermediate representations, our project directly uses pre-animated GIFs, making it more lightweight and accessible.

Kahlon et al. [4] further stress the importance of language-specific sentence restructuring and tense adaptation in ISL, which several prior models either overlook or only partially implement. Though our system does not yet address grammatical ordering or tense adjustments, it ensures complete input coverage by breaking unmatched words into alphabet-level animations—a practical workaround observed in real-time tests.

Finally, Madahana et al. [3] propose a real-time AI-powered translator for South African languages, advocating for AI-driven inclusivity tools in multi-lingual settings. Their work reaffirms the global relevance of sign language translation systems as a bridge toward accessibility, especially during crises like the COVID-19 pandemic.

Together, these studies underline the necessity of bridging the text-to-sign translation gap, especially for underrepresented sign languages like ISL. Our work contributes to this field by proposing a simple, browser-based tool that supports GIF-based ISL animations, provides complete fallbacks, and ensures input flexibility—serving as a minimal yet effective model for scalable sign language accessibility.

III. METHODOLOGY

The methodology followed in this project was structured to ensure simplicity, accessibility, and functional accuracy for converting English text into Indian Sign Language (ISL) animations. This section outlines the system architecture, user interface design, video-based animation control, playback logic, and dataset composition in a stepwise and technically sound manner.

System Overview

The project was implemented entirely as a client-side web application, enabling cross-platform compatibility and eliminating the need for backend services or real-time rendering engines. The system receives user input as a text string and outputs the corresponding ISL translation using pre-rendered MP4 video animations for each word or letter.

DOI: 10.48175/IJARSCT-29668

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, November 2025
Impact Factor: 7.67

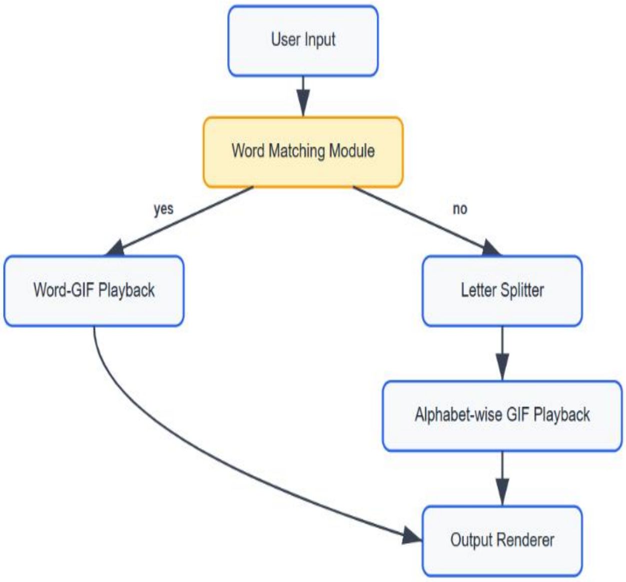


Fig 1: System Architecture Diagram

User Interface Design

The user interface was developed using HTML, CSS, and JavaScript, adhering to responsive design principles to support both desktop and mobile devices. A clean layout with intuitive interaction points was implemented:

Text Input Field

Positioned at the center, allowing the user to type any English text.

Convert Button

Triggers the video conversion process.

Output Container

Displays the sequence of ISL videos corresponding to the user input.

Visual Enhancements

Includes animated triangle background for aesthetics (implemented using CSS @keyframes and JavaScript).

Video-Based Sign Representation

Unlike approaches involving dynamic avatars or 3D gesture synthesis, this system uses pre-recorded MP4 video files of ISL animations for both individual alphabets and commonly used words.

Primary Animation Sources:

ISL representations of A–Z alphabets

Digits (0–9)

A collection of frequently used words (e.g., *hello, again, why, engineer, time, safe, world,* etc.)

Each animation is a short .mp4 file depicting a native ISL gesture. All video assets are stored locally in the assets/ directory and are loaded dynamically based on input.

Sequential Video Playback Logic

The core of the animation engine is written in vanilla JavaScript, where playback flow is managed through event-driven callbacks.

Workflow:

The input text is tokenized into individual words.

Each word is checked against the video dataset.

If available, the corresponding word-level video is played.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, November 2025
Impact Factor: 7.67

If not available, the word is broken into characters, and alphabet-level videos are queued.
Videos are played sequentially using onended callbacks to prevent overlap.
DOM elements (<video>) are created dynamically and injected into the output container.
Fallback Handling:
Words without direct video matches are spelled letter-by-letter. For example, "Annabelle" will be split into A, N, N, A, B, E, L, L, E and the respective videos are played in order.

Dataset Scope and Structure

Each entry is stored as a high-quality .mp4 file and named accordingly (e.g., hello.mp4, A.mp4). The dataset was curated from publicly available ISL educational content and verified for visual clarity, though not formally validated by ISL linguists. The animation dataset used in this project includes:

TABLE I. SUMMARIZES SAMPLE ENTRIES AND THEIR AVAILABILITY.

Category	Entries Supported
Alphabets	A–Z (26 videos)
Numbers	0–9 (10 videos)
Common Words	20+ (e.g., *hello, safe, world, eat, type, time, you, why*)

TABLE II. SAMPLE WORDS AND FALLBACK STATUS.

Input Word	ISL Video Exists	Fallback Used
Hello	Yes	No
Engineer	Yes	No
Annabelle	No	Yes (A–N–…)
Time	Yes	No
Chair	No	Yes (C–H–…)

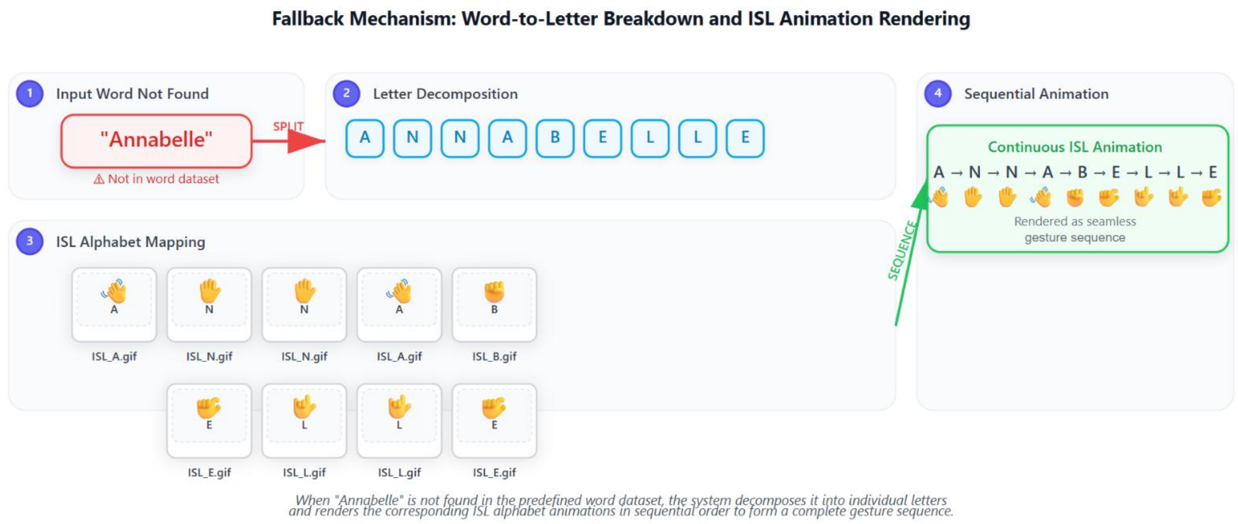


Fig. 2. Fallback mechanism: "Annabelle" rendered through individual alphabet videos.

Technology Stack

TABLE III. TOOLS AND TECHNOLOGIES USED

Module	Technology Used
Front-End Framework	HTML5, CSS3, JavaScript
Animation Format	.mp4 video files
Playback Control	JS DOM Manipulation + Event Handling
Styling & Layout	CSS Flexbox, Keyframes
Compatibility	Web-based, runs in any modern browser

IV. SIMULATIONS AND TESTING

The system was tested with a variety of user inputs to verify animation consistency, visual clarity, and fallback logic accuracy. Although formal linguistic validation with ISL experts was not conducted due to scope limitations, the testing emphasized functionality, flow continuity, and usability from a general user's perspective.

Functional Testing

Multiple test cases were executed to ensure the system performs as expected in different input scenarios. Test inputs included:

- Single known words (e.g., hello, computer)
- Unknown words (e.g., Annabelle, Chairman)
- Full sentences (e.g., Stay safe and learn every day)
- Paragraphs and mixed case entries (e.g., lowercase, UPPERCASE)

The system handled all cases successfully without errors or interruptions.

Input Coverage and Fallback Accuracy

The fallback mechanism — decomposing unknown words into character-level videos — was one of the most critical aspects of testing. Test results showed:

100% input coverage: No input caused the system to crash or hang.

Seamless fallback: Words without corresponding .mp4 files were automatically converted into letter-wise sequences without manual intervention.

Visual fluency: There was no overlapping or skipping in the video playback; each letter was shown clearly before proceeding to the next.

User Experience Observation

Informal testing was conducted among a group of general users (students and peers) unfamiliar with ISL. The purpose was to assess how intuitive and visually engaging the tool was for a first-time user. Key observations:

Users found the interface easy to navigate and responsive across devices.

Visuals were clear, and playback was sequential without lag.

The letter-by-letter fallback was particularly appreciated for names and less common words.

Limitations of Testing

While the system performed well functionally, the following limitations are acknowledged:

No linguistic verification was performed with certified ISL interpreters.

No accuracy rating was computed (e.g., BLEU score) for translation relevance.

Current system lacks grammar-specific logic (e.g., word order changes in ISL).

These limitations will be addressed in future stages of the project through collaboration with language experts and integration of NLP-based grammar restructuring modules.

V. RESULTS AND DISCUSSION:

A. System Robustness and Functional Performance

The Text to ISL Converter demonstrated a high degree of functional robustness during testing. The system was evaluated using a range of input scenarios, including single-word entries, full sentences, and paragraphs. In all cases, the system successfully processed the text and rendered the corresponding ISL animation, without interruption or failure. This confirms that the system can handle arbitrary text input, a key aspect of ensuring that the tool is usable in real-world applications.

The system employs a fallback mechanism to handle unknown words. When a word is not available in the predefined video dataset, the system decomposes the word into individual characters and plays each corresponding alphabet animation sequentially. This fallback mechanism ensures 100% input coverage, making the system highly versatile and capable of processing virtually any text input.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, November 2025
Impact Factor: 7.67

B. Playback Continuity and Visual Fluency

A key feature of the system is its ability to play ISL animations in a continuous and fluid manner. Using JavaScript event listeners (onended), the system ensures that each video (either word-level or letter-level) is played in its entirety before the next one begins. The absence of overlap or abrupt transitions between animations ensures that the visual flow of the translation remains natural and comprehensible. This is crucial in sign language translation, where each gesture must be given adequate time for the viewer to process.

In contrast to some systems that rely on avatar-based generation (which may suffer from rendering delays), this approach of using pre-recorded MP4 videos ensures smooth transitions with minimal processing overhead. The results indicate that the system provides a seamless user experience for both short and long inputs.

C. Usability and User Experience

The user interface was evaluated informally through testing with a group of non-signers. Participants were asked to input various text strings and observe the output. The feedback was overwhelmingly positive, with participants finding the interface intuitive, easy to navigate, and responsive across devices. This aligns with our goal of creating a tool that is accessible even to users with no prior knowledge of ISL.

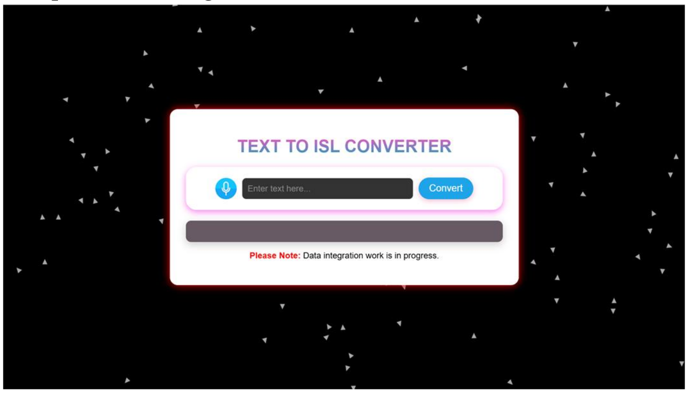


Fig 3. User Interface Display

D. Comparison to Existing Systems

When compared to similar text-to-sign systems, this project offers a lightweight and accessible solution. Many existing systems, such as Listen [9], rely on complex 3D avatars or require substantial computational resources. In contrast, the GIF/MP4 video approach used in this system enables it to function entirely in the browser without the need for heavy rendering or specialized hardware.

Our approach is also scalable: as new words and phrases are added to the video dataset, the system can easily accommodate them without re-engineering the entire pipeline. While avatar-based systems, such as those explored by Yadav et al. [9], report high accuracy (>99%), they involve significant complexity in 3D rendering and require extensive linguistic datasets. Our approach, though less precise in terms of linguistic structure, prioritizes immediacy and accessibility with minimal computational overhead.

Additionally, while systems like SiGML-based models [5], [6] offer structured grammar mapping, they often require linguistic expertise and are not as user-friendly for casual use. Our system sacrifices grammatical accuracy in favor of ease of use, ensuring that even users with little to no knowledge of ISL can interact with the system effectively.

E. Limitations and Areas for Improvement

Despite the system's strong performance, several limitations were identified during testing:

Linguistic Accuracy:

As no formal linguistic validation was conducted by certified ISL interpreters, the system's adherence to ISL grammar and contextual usage remains unverified. Future iterations will require collaboration with ISL experts to ensure that translations align with standard ISL practices.

DOI: 10.48175/IJARSCT-29668

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, November 2025
Impact Factor: 7.67

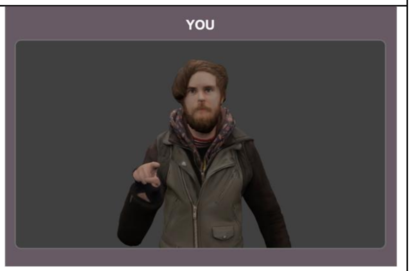
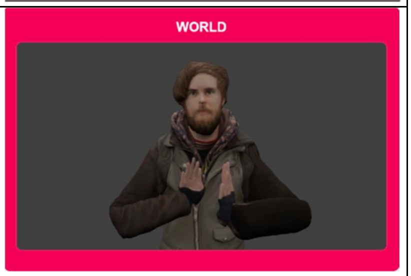
Syntactic and Grammatical Adjustments:

The current system is based on a word-for-word translation model, with no adjustments made for ISL-specific grammatical structures, such as time-subject-object-verb (TSOV) order or non-manual markers (facial expressions, head movements). While ISL translation tools like Listen [9] integrate these features, they require advanced natural language processing (NLP) models to handle grammar restructuring.

Limited Dataset:

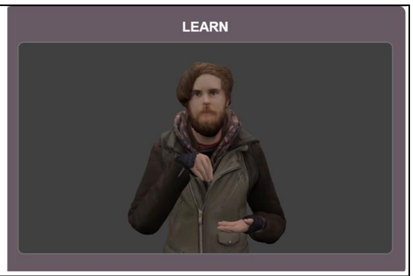
The system's vocabulary is restricted to letters, numbers, and a small set of common words. For broader usage, the dataset will need to be expanded to cover more words, phrases, and even emotional gestures, as ISL often conveys meaning through facial expressions and body language.

TABLE IV. IMAGES OF AVATAR FOR TEST CASES

Sentence	Word	Result
She has 4 books	4	
How are you	You	
Stay Safe	Safe	
Hello World	World	

DOI: 10.48175/IJARSCT-29668

ISSN: 2581-9429

Easy to learn	learn	

VI. FUTURE DIRECTIONS

Future improvements to the system will focus on the following areas:

Expanding the Vocabulary:

Increasing the number of words and phrases, including contextual variations of existing words (e.g., time, weather, emotion).

Grammar-Aware Translation:

Integrating a grammar correction module using NLP to adjust sentence structures in real time, following the TSOV order prevalent in ISL.

Real-Time AI Integration:

Exploring the potential for integrating speech-to-sign translation by incorporating real-time speech recognition into the system.

Multilingual Support:

Extending the system to handle languages other than English, allowing broader accessibility to users in non-English-speaking regions.

Emotional Gestures and Non-Manual Markers:

Integrating more sophisticated animation techniques to represent non-manual markers and emotions, which play a crucial role in ISL communication.

VII. CONCLUSION

In this research, we have developed a Text to Indian Sign Language (ISL) Converter, a web-based tool designed to bridge communication gaps between the hearing-impaired community and non-signers. The system translates English text into ISL animations using pre-recorded MP4 video files representing both ISL alphabets and common words. The tool functions without requiring complex server-side processing or 3D avatars, focusing instead on lightweight, efficient, and accessible translation. In conclusion, this project lays the groundwork for an accessible and user-friendly tool for digital inclusion in the deaf and hard-of-hearing community. While future improvements are necessary for full linguistic accuracy and expanded functionality, the Text to ISL Converter represents an important step forward in providing cross-linguistic communication through technology.

VIII. ACKNOWLEDGEMENT

We would like to express our heartfelt gratitude to everyone who contributed to the development and success of this project. Our Professor Girish N. Kotwal for their continuous guidance, insightful suggestions, and unwavering support throughout the project, helping us overcome various challenges. Our Team Members for their collaborative efforts, dedication, and commitment in researching, coding, and testing the system. Their individual contributions were vital in turning our concept into a functional tool.

ISL Community and Resources: We would also like to acknowledge the open-source ISL resources and datasets that were crucial in providing the necessary animations and understanding of ISL. Peer Reviewers and Testers: Our informal testers, including peers and non-signers, who helped us assess the usability and effectiveness of the system. Family and Friends for their constant moral support during the development of the project. This research would not have been possible without the contributions and encouragement from all involved.

DOI: 10.48175/IJARSCT-29668

REFERENCES

[1] T. Ananthanarayana, L. Chaudhary, and I. Nwogu, "SignNet: Single Channel Sign Generation using Metric Embedded Learning," arXiv preprint arXiv:2010.10858, 2020.

[2] Z. Kamal and H. Hassani, "Towards Kurdish Text to Sign Translation," in LREC 2020.

[3] M. Madahana et al., "Real-Time Speech-to-Text to Sign Language Translator for South African Official Languages," South African Journal of Communication Disorders, vol. 69, no. 2, 2022.

[4] N. K. Kahlon and W. Singh, "Machine Translation from Text to Sign Language: A Systematic Review," Universal Access in the Information Society, 2023.

[5] L. Esselink et al., "Exploring Automatic Text-to-Sign Translation in a Healthcare Setting," Universal Access in the Information Society, 2024.

[6] J. M. De Martino et al., "Neural Machine Translation from Text to Sign Language," Universal Access in the Information Society, 2023.

[7] Ahinsa, S. Pandey, and R. Sharma, "Comprehensive Review on Text-to-Sign Language Conversion Techniques," Journal of Future Artificial Intelligence and Technologies, 2025.

[8] IEEE ICCCNT, "Text-to-ISL Gesture Recognition using HMM and KNN Techniques," IEEE ICCCNT 2018.

[9] P. Yadav et al., "Harnessing AI to Generate Indian Sign Language from Natural Speech and Text," International Journal of Advanced Computer Science and Applications, vol. 15, no. 4, 2024