e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
( Peer-Reviewed, Open Access, Fully Refereed International Journal )
Volume:07/Issue:04/April-2025 Impact Factor- 8.187 www.irjmets.com

# AI ENABLED SPEECH TO SIGN LANGUAGE CONVERSION

**Ramya D[*1], Keerthana D[*2], Madhav Priyan S[*3], Nanthene R E[*4], Rajalingam P[*5]**

[*1]Assistant Professor, Department Of Computer Science And Engineering, Knowledge Institute Of Technology, Salem, Tamil Nadu, India.

[*2,3,4,5]Students, Department Of Computer Science And Engineering, Knowledge Institute Of Technology, Salem, Tamil Nadu, India.

DOI : https://www.doi.org/10.56726/IRJMETS72763

## ABSTRACT

Artificial intelligence (AI) has significantly advanced assistive technologies, particularly by enhancing communication solutions for individuals with hearing impairments. This work introduces a novel AI-powered web application designed to convert spoken language into sign language in real time, thereby improving accessibility for the Deaf and hard-of-hearing community. Leveraging cutting-edge speech recognition and natural language processing (NLP) techniques, the system accurately transcribes spoken input into text and maps it to corresponding sign language gestures. These gestures are then displayed as fluid animations or avatar-based representations, enabling seamless interactions between hearing and non-hearing individuals. The application utilizes deep learning methods to enhance the accuracy of speech-to-text conversion and relies on a comprehensive database of sign language movements to ensure precise translations. A key feature of this tool is its adaptability to various sign languages, such as American Sign Language (ASL) and British Sign Language (BSL), making it a versatile solution for global adoption. The platform is designed with a user-friendly interface, ensuring accessibility across multiple devices, including smartphones and desktop computers. Through the integration of machine learning, the system continuously improves its translation capabilities by learning from user interactions, resulting in increased efficiency and accuracy over time. This AI-driven solution aims to bridge the communication gap between hearing and non-hearing individuals by providing a simple yet effective method for converting speech into sign language. Its web-based architecture ensures wide availability without the need for specialized hardware. Representing a significant step toward inclusivity, this technology has the potential to transform educational, professional, and social opportunities for the Deaf community. Looking ahead, future enhancements could involve support for regional linguistic variations, improved gesture visualization, and the incorporation of augmented reality (AR) to create a more immersive user experience.

**Keywords:** Artificial Intelligence (Ai), Assistive Technology, Sign Language Translation, Speech Recognition, Natural Language Processing (Nlp), Machine Learning.

## I. INTRODUCTION

Communication is a cornerstone of human connection, yet individuals with hearing impairments often encounter barriers when engaging with those unfamiliar with sign language. While conventional speech-to-text tools partially address this divide, they fall short of translating spoken language into sign language—the preferred communication method for many in the Deaf community. Recent progress in artificial intelligence (AI) has paved the way for innovative approaches to enhance inclusivity and accessibility in communication. This paper presents an AI-powered web application designed to convert speech into sign language, significantly improving access for people with hearing challenges. By employing advanced speech recognition and natural language processing (NLP) technologies, the system transcribes spoken input and transforms it into corresponding sign language gestures. These gestures are rendered in real time via an animated avatar, facilitating smooth and natural exchanges between hearing and non-hearing individuals. Beyond improving communication, this technology supports greater social inclusion and expands opportunities in areas such as education, employment, and everyday interactions. The web-based nature of this AI-driven solution ensures broad accessibility across various devices, eliminating the need for specialized equipment or software. With only an internet connection, users can utilize the platform from virtually any location, making it a practical and scalable resource for diverse contexts, including educational settings, customer support, and public services. By integrating AI and web technologies, this initiative seeks to narrow the communication gap and foster a more inclusive environment for the Deaf and hard-of-hearing community. The structure of this paper is organized as follows: Section 2 provides a Literature Survey, Section 3 reviews the Existing System, Section 4 outlines the Proposed System, Section 5 details the Module Description, Section 6 presents the Results, Section 7 offers the Conclusion and Future Enhancements, followed by References.

## II. LITERATURE SURVEY

Sign Language Recognition (SLR) based on hand gestures is vital for enabling communication between individuals with hearing impairments and those without. However, the lack of a universal sign language standard leads to significant cultural variations such as Korean, American, and Japanese sign languages, each influenced by regional context. While existing SLR systems perform well within their specific domains, they often struggle to generalize across multiple sign languages or multicultural sign language (McSL) settings. To address this issue, a study introduced GmTC, an end-to-end SLR system that effectively translates McSL into text [1]. GmTC utilizes a dual-stream architecture: the first stream extracts graph-based features using superpixel values and a Graph Convolutional Network (GCN), while the second uses an attention-based network with Multi-Head Self-Attention (MHSA) and Convolutional Neural Networks (CNNs) to learn both short- and long-range dependencies. By fusing these feature streams, the model generates a rich representation for sign classification. Evaluated across five culturally diverse datasets, GmTC outperformed existing models in both accuracy and adaptability [1].

Another approach combined real-time sign language and audio conversion using AI technologies to enhance accessibility for individuals with hearing or visual impairments [2]. These systems utilize deep learning and computer vision to translate between speech and gestures, promoting inclusivity. Audio input is converted into visible gestural signals through avatars, using advanced Natural Language Processing (NLP) and AI algorithms. Key considerations in such systems include ensuring accuracy, maintaining data privacy, and minimizing ergonomic issues during prolonged use. Further research focused on creating inclusive SLR and Sign Language Translation (SLT) systems that support bidirectional translation between sign language and speech/text [4]. This study emphasized the need to handle the unique grammar and semantics of sign languages, as well as the scarcity of diverse datasets that include regional dialects like Indian Sign Language (ISL). The challenges in translating glosses into spoken text highlight the complexity of building robust SLT systems.

Other works proposed dual-language sign language platforms that support both ISL and American Sign Language (ASL), and further translate them into multiple Indian regional languages [3], [5]. These applications utilize CNNs and deep learning techniques to improve user interaction and accessibility. By accommodating regional language outputs and diverse sign systems, they foster effective communication and inclusivity for the Deaf and hard-of-hearing communities. A comprehensive sign language application was also developed to support sign learning and real-time communication [5]. This system recognizes alphabets and words, translates text to gestures, and provides multi-language voice output, prioritizing usability, efficiency, and accuracy. It addresses the gap in communication tools for users and learners of sign language by enhancing educational and accessibility experiences.

Several other studies have advanced this field further. For instance, grammar-based inductive learning (GBIL) models were used to improve sign spotting in continuous sign language videos [8], while novel models focused on Arabic Sign Language Recognition (ARSLR) using deep learning and gesture tracking [6]. More recent innovations integrated speech-to-sign conversion systems [7], [9], [10], offering real-time and multilingual capabilities to support seamless communication. Collectively, these studies demonstrate the growing effectiveness and versatility of AI-based SLR and SLT technologies. Despite challenges like language-specific grammar and dataset limitations, the field continues to evolve with applications tailored for cross-cultural accessibility, language diversity, and real-time interaction.

## III. EXISTING SYSTEM

Current methods of communication between speech users and the Deaf or hard-of-hearing community often rely on human interpreters, pre-recorded sign language videos, or simple text-based solutions. While traditional speech-to-text applications offer some level of accessibility by converting spoken words into text, they fall short in effectively supporting individuals who primarily use sign language as their mode of

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
( Peer-Reviewed, Open Access, Fully Refereed International Journal )
Volume:07/Issue:04/April-2025 Impact Factor- 8.187 www.irjmets.com

communication. Many existing systems that attempt to convert speech or text into sign language suffer from several critical limitations. A major drawback is the lack of real-time processing, which results in communication delays and reduces conversational flow. Additionally, these systems often depend on limited sign language databases, covering only a narrow range of vocabulary, which hinders their usability in everyday conversations. Sign languages possess unique grammatical structures that differ significantly from spoken languages, but most basic tools fail to accurately interpret and reproduce this complexity.

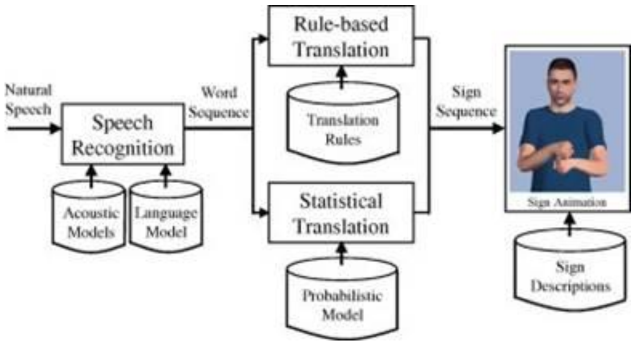
Another significant shortcoming is the absence of facial expressions and body movements, which are integral components of sign language, conveying tone, emotion, and context. Some solutions still rely on static or pre-recorded videos, making them unsuitable for dynamic or spontaneous conversations. Furthermore, the diversity of sign languages across different regions—such as American Sign Language (ASL), British Sign Language (BSL), and Indian Sign Language (ISL)—is often not accommodated, limiting the system's adaptability to global users. Personalization features are also scarce, leaving users unable to integrate regional signs, slang, or preferences into the system. High computational requirements pose another barrier, as many advanced solutions necessitate powerful hardware, making them less accessible to users with limited resources. Additionally, speech recognition often performs poorly in noisy environments, leading to inaccurate sign translations. Lastly, the high cost and limited availability of advanced systems make them inaccessible to a large portion of the intended user base, thereby hindering effective communication and inclusivity.

## IV. PROPOSED METHODOLOGY

The proposed system is an AI-enabled speech-to-sign language conversion web application designed to bridge the communication gap between hearing individuals and the deaf or hard-of-hearing community.This system utilizes advanced speech recognition technology to convert spoken language into text, which is then processed by an AI-powered module that translates the text into corresponding sign language gestures.The translated signs are displayed as animated sign language representations or virtual avatars, ensuring effective and real-time communication. The web-based nature of the application makes it accessible across various devices without requiring additional software installation. This system aims to enhance inclusivity by facilitating seamless interactions in educational, professional, and social environments.

**Proposed Model Diagram:**

System Architecture




**Workflow of the System**

1. User speaks into the microphone
2. Speech recognition converts audio into text
3. NLP module processes text & maps to sign language structure
4. Sign language converter generates animations/videos
5. Web app displays the sign output to the user

**Tools and Technologies:**

**ASP.NET - FRONT END:**

ASP.NET serves as a powerful framework for front-end development, distinct from being merely an updated version of ASP. It offers a robust combination of enhanced developer efficiency, strong performance, dependability, and seamless deployment capabilities. ASP.NET reimagines the development process from the ground up, maintaining accessibility for beginners while introducing a variety of innovative management techniques. This redesign ensures that it remains approachable for newcomers yet equips developers with advanced tools and methods to streamline their work effectively.

**VB.NET - MIDDLE END:**

VB.NET represents an advanced evolution beyond Microsoft's earlier ASP (Active Server Pages), marking a significant step forward in server-side scripting technology. It empowers developers to create robust, dependable, and scalable distributed applications, building on the strengths of its predecessors. Rooted in the Microsoft .NET framework, VB.NET leverages the framework's extensive features and tools to facilitate the development of web applications and services. While it shares some similarities with ASP, including compatible syntax, VB.NET goes far beyond, offering a rich set of capabilities that enhance the efficiency of building high-quality, scalable web solutions with fewer resources and less time. Operating within a compiled, .NET-based environment, it supports a range of .NET-compatible languages such as C#.NET, C#, JScript.NET, and VBScript.NET, providing developers with flexible options to craft C#.NET applications tailored to their needs.

**SQL SERVER 2005 - BACK END:**

Microsoft SQL Server is a relational database management system crafted by Microsoft, designed to function as a database server. Its core purpose is to efficiently store and retrieve data as needed by various software applications, whether they operate on the same machine or across a network, including the Internet. Microsoft offers a diverse range of SQL Server editions—at least twelve—tailored to meet the needs of different users, from small-scale, single-device applications to expansive, internet-facing systems handling numerous simultaneous users. The system employs T-SQL and ANSI SQL as its primary query languages, enabling robust data management and interaction.

## V. MODULE DESCRIPTION

**Speech Recognition Module:**

The Speech Recognition Module is pivotal in transforming spoken language into text through the application of Automatic Speech Recognition (ASR) technology. This component captures audio input from the user and converts the spoken content into textual form in real time. To achieve superior accuracy, the system employs noise-filtering techniques to minimize background interference and leverages Natural Language Processing (NLP) methods to refine the transcription process. The module can be developed using robust tools such as the Google Speech-to-Text API, DeepSpeech, or Whisper AI, each recognized for their effectiveness and dependability in handling real-time speech recognition tasks.

**Natural Language Processing (NLP) Module:**

Once the speech is transcribed into text, the Natural Language Processing (NLP) Module processes it for context and structure. It tokenizes and analyzes the text, identifies keywords and sentence patterns, and simplifies complex phrases into easily interpretable sentences for sign language translation. Technologies such as NLTK, SpaCy, or custom NLP models can be used for this purpose.

**Sign Language Translation Module:**

The Sign Language Translation Module maps the processed and simplified text to appropriate sign language gestures. It uses machine learning models trained on datasets specific to various sign languages such as ASL (American Sign Language), BSL (British Sign Language), and ISL (Indian Sign Language). This module ensures accurate word-to-sign matching and supports regional sign language variations

**Sign Language Animation Module:**

This module converts the translated text into visual sign language using animations. It employs technologies like WebGL, Three.js, or Blender to generate 3D avatar-based animations or use pre-recorded sign language videos. The animation is synchronized with natural expressions and body gestures to ensure smooth and realistic playback, enhancing comprehension for the end user.

**Web Application Interface Module:**

The Web Application Interface Module provides an interactive platform for users. Built using HTML, CSS, and JavaScript frameworks such as React.js, Vue.js, or Angular, it captures speech input, displays transcribed text in real time, and showcases corresponding sign language animations. It also offers customization options like language selection and speed control, ensuring a user-friendly experience across devices.

**Database & API Integration Module:**

This module manages data storage and retrieval functions for the application. It stores sign language word mappings, gesture animations, and logs user interactions for analysis and improvement. Technologies such as Firebase, MongoDB, or MySQL are used to ensure efficient data handling and API- based access to translation resources.

**AI Learning & Improvement Module:**

The AI Learning & Improvement Module enhances the system's performance over time using machine learning. By incorporating user feedback and real-world usage data, it refines speech recognition and sign translation accuracy. It also adapts to different accents, dialects, and sign language variations. Frameworks like TensorFlow and PyTorch are used to implement these AI-driven improvements.

## VI. SYSTEM TESTING:

The process of testing an integrated hardware and software system is an implementation process that helps ensure that the system works correctly and consistently before the start of a live operation. Testing is crucial to the success of the system. System Testing is a logical assumption that the goal would be accomplished if all parts of the device are correct.

**Unit Testing:**

Unit testing involves testing individual components of the AI- enabled speech-to-sign language conversion system to ensure they function as expected. Each module, such as speech recognition, text processing, and sign language animation rendering, is tested independently. Mock data is used to simulate real inputs, and test cases are designed to verify the correctness of each function, edge cases, and error handling. Automated unit tests help in early bug detection and improve the reliability of the application before integration.

**Integration Testing:**

Integration testing verifies that different modules of the web application work together seamlessly. It ensures proper communication between the speech-to text engine, natural language processing (NLP) module, and sign language animation generator. Data flow between these components is tested to identify issues such as data loss, incorrect format conversions, or API failures. Both top 47 down and bottom- up integration testing strategies can be used to validate smooth interaction between modules.

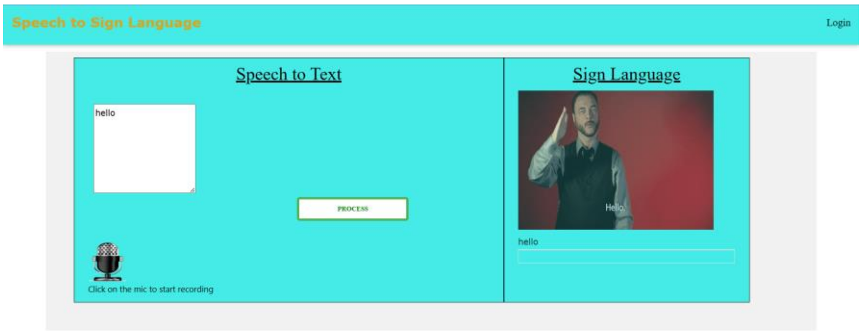
**Validation Testing:**

Validation testing ensures that the entire AI-enabled speech- to-sign language web application meets the specified requirements and user expectations. This phase involves testing the application in a real-world scenarios with diverse speech inputs and evaluating sign language output accuracy. It also includes usability testing to verify the responsiveness and accessibility of the web interface. The goal is to confirm that the system performs accurately, efficiently, and in a user- friendly manner before deployment

## VII. RESULT

The experimental framework for the AI-driven speech-to-sign language conversion web application integrates speech recognition, natural language processing (NLP), and sign language animation technologies. The system is structured around three core components: a speech-to-text module, a text-to-sign language translation module, and a visual sign language avatar. The speech-to-text module employs Automatic Speech Recognition (ASR) technology to transcribe spoken input into text in real time. Subsequently, this text undergoes processing through NLP techniques to enhance sentence coherence and contextual accuracy prior to its conversion into sign language gestures. The text-to-sign language translation module relies on a pre-trained AI model, developed using an extensive dataset encompassing sign language vocabularies and syntactic rules. This model converts the transcribed text into a series of sign language gestures, accommodating linguistic differences across systems such as American Sign Language (ASL) and British Sign Language (BSL).

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
( Peer-Reviewed, Open Access, Fully Refereed International Journal )
Volume:07/Issue:04/April-2025 Impact Factor- 8.187 www.irjmets.com




**Data flow diagram:**


The system also incorporates a database of predefined gestures and fingerspelling to enhance accuracy in cases where direct translations are unavailable. Additionally, machine learning algorithms refine the translations by adapting to user feedback and contextual improvements over time. Finally, the translated sign language output is visualized using an animated sign language avatar embedded within the web application. The avatar dynamically performs the generated sign sequences, ensuring a fluid and natural representation of sign language communication. The web application interface allows users to adjust settings such as sign speed and hand orientation for a personalized experience. This experimental setup is designed to improve accessibility for the Deaf and Hard of Hearing (DHH) community by bridging the communication gap between spoken language users and sign language users in real-time interactions.

## VIII. CONCLUSION

The development of an AI-enabled speech-to-sign language conversion web application represents a significant step toward bridging communication barriers between the hearing and deaf communities. By leveraging advanced speech recognition, natural language processing, and sign language generation technologies, the system provides an accessible and real-time translation solution. This innovation not only enhances inclusivity but also facilitates smoother interactions in educational, professional, and social settings. While the current implementation demonstrates promising results, ongoing improvements in accuracy, speed, and sign language variations will further enhance its effectiveness and usability.Looking ahead, this project aims to broaden its reach by incorporating support for a variety of sign languages, addressing the needs of diverse linguistic groups across the globe. The addition of AI-powered avatar technology to include facial expressions and body movements will make the sign language translations feel more natural and emotionally expressive. Progress in

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
( Peer-Reviewed, Open Access, Fully Refereed International Journal )
Volume:07/Issue:04/April-2025 Impact Factor- 8.187 www.irjmets.com

machine learning techniques is expected to enhance the system's ability to recognize speech, allowing it to better adapt to a range of accents and speaking styles with greater precision. Furthermore, adding mobile compatibility and the ability to function offline will improve the tool's accessibility, enabling smooth communication regardless of location or internet availability. Working closely with linguists, members of the Deaf community, and accessibility specialists will play a critical role in fine-tuning and improving the application to meet practical, real-world demands effectively.

## IX. REFERENCES

[1] S. M. Miah, M. A. M. Hasan, Y. Tomioka, and J. Shin, "Hand gesture recognition for multi-culture sign language using graph and general deep learning network," IEEE Open J. Comput. Soc. , vol. 5, 2024.

[2] A. Boobal, J. L. Jasmine, C. C. K. Reddy, C. A. Reddy, and C. B. V. S. Rohith, "Real-time sign language and audio conversion using AI," 2024.

[3] K. S. Sindhu, M. Mehnaaz, N. Biradar, P. L. Varma, and C. Uddagiri, "Sign language recognition and translation systems for enhanced communication for the hearing impaired," in Proc. 1st Int. Conf. Cognit., Green and Ubiquitous Comput. (IC-CGU), 2024.

[4] S. M. Antad, S. Chakrabarty, S. Bhat, S. Bisen, and S. Jain, "Sign language translation across multiple languages," in Proc. Int. Conf. Emerg. Syst. Intell. Comput. (ESIC), 2024.

[5] D. S. Priyadharshini, R. Anandraj, K. R. G. Prasath, and S. A. F. Manogar, "A comprehensive application for sign language alphabet and word recognition, text-to-action conversion for learners, multi-language support and integrated voice output functionality," in Proc. Int. Conf. Sci. Technol. Eng. Manage. (ICSTEM), 2024.

[6] A. M. Farouk, A. M. Zenhom, E. M. Abdelaleem, S. A. Fadel, K. K. A. Elsayed, and M. S. Mohammed, "A new approach for Arabic sign language recognition (ARSLR)," in Proc. 6th Novel Intell. Leading Emerg. Sci. Conf. (NILES), 2024.

[7] A. Deshmukh, A. Machindar, S. Lale, and P. Kasambe, "Enhancing communication for the hearing impaired: a real-time speech to sign language converter," in Proc. 27th Int. Symp. Wireless Pers. Multimedia Commun. (WPMC) , 2024.

[8] V. N. S. A. Amperayani, A. Banerjee, and S. K. S. Gupta, "Grammar-based inductive learning (GBIL) for sign-spotting in continuous sign language videos," in Proc. IEEE 7th Int. Conf. Ind. Cyber-Phys. Syst. (ICPS) , 2024.

[9] S. A. Shaba, S. I. Shroddha, M. J. Hossain, and M. F. Monir, "Translating one sign language to another sign language using deep learning model and GTTS," in Proc. IEEE 99th Veh. Technol. Conf. (VTC2024-Spring), 2024.

[10] S. Daga, A. Dusane, and D. Bobby, "With you – Indian sign language detection and alert system," in Proc. Int. Conf. Emerg. Smart Comput. Informat. (ESCI), 2024.