AI-GENERATED REAL-TIME SPEECH-TO-SIGN LANGUAGE TRANSLATION USING ANIMATED AVATARS

[1]K.Murugan, [2]Abdhul Azees S, [3]Chinnachamy P, [4]Dinesh M, [5]Gokulakannan C
[1]Assistant Professor, Department of Computer Science and Engineering,
Hindusthan Institute of Technology, Coimbatore
[2,3,4] UG student, Department of Computer Science and Engineering,
Hindusthan Institute of Technology, Coimbatore
[1] murugan.k@hit.edu.in,[2]azees.novo@gmail.com, [3]chinnachamy004@gmail.com, [4]md383755@gmail.com , [5]gokul9597c@gmail.com

ABSTRACT: AI-Generated Real-Time Speech-to-Sign Language Translation Using Animated Avatars is an intelligent assistive communication system designed to bridge the interaction gap between speech-impaired and hearing individuals. The system captures live speech input through a microphone and processes it using advanced speech recognition and natural language processing techniques to convert spoken language into meaningful textual representations in real time. The translated text is then mapped to corresponding sign language gestures, which are dynamically rendered through a 3D animated avatar for clear visual communication. The platform integrates artificial intelligence, computer vision concepts, and avatar-based animation technologies to ensure accurate gesture representation and smooth real-time performance. The system also supports multilingual speech input and provides an interactive user interface for enhanced usability and learning support. By combining real-time speech analytics, intelligent sign generation, and expressive avatar visualization, the proposed solution offers an efficient, scalable, and inclusive communication tool. This platform contributes to social accessibility, promotes inclusive digital interaction, and supports assistive technology development by enabling seamless speech-to-sign translation through AI-driven automation

Keywords: Speech-to-Sign Language Translation, Assistive Communication System, Artificial, Intelligence in Accessibility, Real-Time Speech Recognition, Natural Language Processing(NLP), Animated Avatar Communication, Sign Language Generation

Corresponding Author: *K.Murugan*
Assistant Professor / CSE, Hindusthan Institute of Technology
Coimbatore, Tamil Nadu, India
Mail: murugan.k@hit.edu.in

INTRODUCTION

Communication is a fundamental human necessity that enables social interaction, knowledge sharing, and inclusive participation in daily activities. However, individuals with hearing and speech impairments often face significant challenges in communicating effectively with the general population. Sign language serves as the primary mode of communication for many deaf and mute individuals, but a large portion of society lacks the knowledge required to understand sign language gestures. This communication barrier can lead to social isolation, limited educational opportunities, and reduced access to essential services for differently-abled individuals. As technology continues to evolve, there is an increasing need for intelligent assistive systems that can facilitate seamless interaction between hearing and speech-impaired communities and the rest of society.

Recent advancements in Artificial Intelligence, Natural Language Processing, and Computer Vision have created new opportunities to develop automated translation systems capable of bridging communication gaps. Speech recognition technology has achieved significant improvements in accurately converting spoken language into text, while animation and graphical rendering technologies have enabled the realistic visualization of human gestures and expressions. By combining these technological developments, it is possible to create real-time systems that translate spoken language into sign language representations, thereby enhancing accessibility and inclusivity in communication. Despite the availability of various assistive communication tools, many existing systems rely on pre-recorded videos, static images, or manual interpretation, which may not provide real-time responsiveness or contextual accuracy. Additionally, traditional sign language learning and interpretation methods can be time-consuming and may not be practical in fast-paced environments such as classrooms, public services, healthcare facilities, and workplaces. These limitations highlight the necessity for intelligent automated platforms that can process live speech input, understand linguistic context, and generate expressive sign language output dynamically.

To address these challenges, this project introduces an AI-Generated Real-Time Speech-to-Sign Language Translation System using Animated Avatars. The proposed system captures spoken input through a microphone and applies advanced speech recognition algorithms to convert it into textual form. The processed text is then analyzed using Natural Language Processing techniques to extract meaningful linguistic patterns and map them to corresponding sign language gestures. These gestures are visually represented through a 3D animated avatar capable of performing dynamic hand movements and facial expressions, ensuring natural and effective communication. The system further enhances user interaction through an intuitive graphical interface that supports multilingual speech input, real-time translation feedback, and educational assistance for sign language learning. By integrating artificial intelligence-based speech processing, gesture mapping algorithms, and avatar animation technologies, the proposed platform aims to provide an efficient, scalable, and inclusive communication solution. This research contributes to the development of assistive technologies that promote accessibility, reduce communication barriers, and support the social integration of hearing and speech-impaired individuals in modern digital society.

LITERATURE SURVEY

Chen et al. (2025) introduced an intelligent communication system that integrates Large Language Models (LLMs) with structured data representations to interpret user instructions and generate meaningful outputs. Their approach demonstrates how advanced language models can understand natural language inputs and produce context-aware responses. This concept is highly relevant to speech-to-sign translation systems, where accurate interpretation of spoken language is essential. However, their system relies heavily on large computational resources and may face challenges in real-time responsiveness and handling ambiguous linguistic inputs.

Sun et al. (2025) developed a real-time multimodal sensing system that processes live input data and produces rapid analytical outputs using machine learning models. Their work highlights the importance of real-time processing and fast response systems, which are crucial in assistive communication technologies. Although their approach achieves high speed and efficiency, its performance may vary depending on environmental conditions and input quality, similar to challenges faced in speech recognition systems.

Khan et al. (2025) proposed an AI-based smart assistant designed to interact with users through natural language and provide automated responses for assistance and guidance. The system demonstrates how conversational AI can enhance user interaction and accessibility. This concept aligns with speech-to-sign translation platforms, where natural language understanding plays a key role. However, the system depends on a well-structured knowledge base and may struggle with complex or context-dependent queries.

Hasan et al. (2024) presented an autonomous assistive system that integrates computer vision, AI algorithms, and interactive modules to support user communication and monitoring tasks. Their system utilizes real-time processing and intelligent decision-making to improve system efficiency. While effective, limitations include challenges in noisy environments and dependency on hardware capabilities, which are also relevant concerns in speech-to-sign systems.

Xu et al. (2024) introduced an embodied AI system capable of understanding natural language instructions and performing context-aware actions through a "think-observe-execute" framework. This research emphasizes the importance of combining language understanding with real-time execution, which is essential for generating accurate sign language gestures from speech input. However, issues such as data dependency, computational cost, and potential bias remain challenges.

Shaik et al. (2023) reviewed the role of Artificial Intelligence in real-time monitoring and user interaction systems. Their study highlights the effectiveness of AI in improving personalization, automation, and predictive analysis. These insights are applicable to assistive communication systems, where personalization and adaptive learning can enhance translation accuracy. However, challenges such as data privacy and system integration must be addressed.

Zhang et al. (2023) developed a real-time sensing and signal processing system capable of extracting meaningful information from dynamic inputs. Their work demonstrates the importance of robust signal processing techniques in handling real-world data variations. This is particularly relevant for speech recognition modules, where background noise and signal distortion can affect performance.

Warmbein et al. (2023) explored the adoption of intelligent assistive technologies and identified factors such as user acceptance, training requirements, and system usability as critical for successful implementation. Their findings emphasize that beyond technical performance, user experience and accessibility play a vital role in the effectiveness of assistive communication systems.

Sun et al. (2022) proposed a vision-based real-time processing system that extracts meaningful features from visual inputs and converts them into interpretable outputs. Their approach highlights the role of computer vision in gesture recognition and visual communication systems. However, performance may vary based on environmental conditions such as lighting and motion.

Huang et al. (2022) introduced an AI-enabled assistive platform that integrates multiple sensors and real-time processing modules to support remote interaction and monitoring. Their system demonstrates how multimodal integration improves system functionality and user interaction. However, it relies on stable operating conditions and high-quality input data.

Rohmetra et al. (2021) examined AI-driven real-time interaction systems and highlighted their role in improving accessibility and reducing communication barriers. Their work emphasizes the importance of reliable AI models, data privacy, and system robustness in real-world deployments.

PROPOSED SYSTEM

The proposed AI-Generated Real-Time Speech-to-Sign Language Translation System is designed as an intelligent assistive communication platform that enables effective interaction between hearing individuals and people with hearing or speech impairments. The primary objective of the system is to provide a centralized digital solution that can capture live speech input, process linguistic information using Artificial Intelligence techniques, and convert it into meaningful sign language gestures represented through animated avatars. In many real-world situations such as classrooms, hospitals, public service centers, and workplaces, communication barriers often arise due to the lack of common understanding between spoken language users and sign language users. Traditional interpretation methods depend on human interpreters, prerecorded gesture videos, or static visual aids, which may not provide real-time responsiveness, contextual accuracy, or scalability. These limitations create challenges in ensuring inclusive communication and accessibility for differently-abled individuals.

The proposed system addresses these challenges by integrating real-time speech recognition, natural language processing, gesture mapping algorithms, and avatar-based animation within a unified intelligent platform. The system captures spoken language through a microphone and applies advanced speech recognition models to convert audio signals into textual form instantly. The generated text is then processed using Natural Language Processing techniques to analyze sentence structure, identify keywords, and understand contextual meaning. This processed linguistic information is mapped to corresponding sign language gestures using a gesture translation engine that references a structured sign language database.

To provide effective visual communication, the translated gestures are dynamically rendered through a 3D animated avatar capable of performing hand movements, body gestures, and facial expressions in a natural and synchronized manner. This avatar-based visualization approach enhances user engagement and improves comprehension compared to static gesture images or prerecorded clips. The system also incorporates multilingual speech support, enabling users to communicate in different spoken languages while receiving sign language output in real time. By combining speech analytics with intelligent gesture generation, the proposed system ensures accurate, responsive, and inclusive communication support.

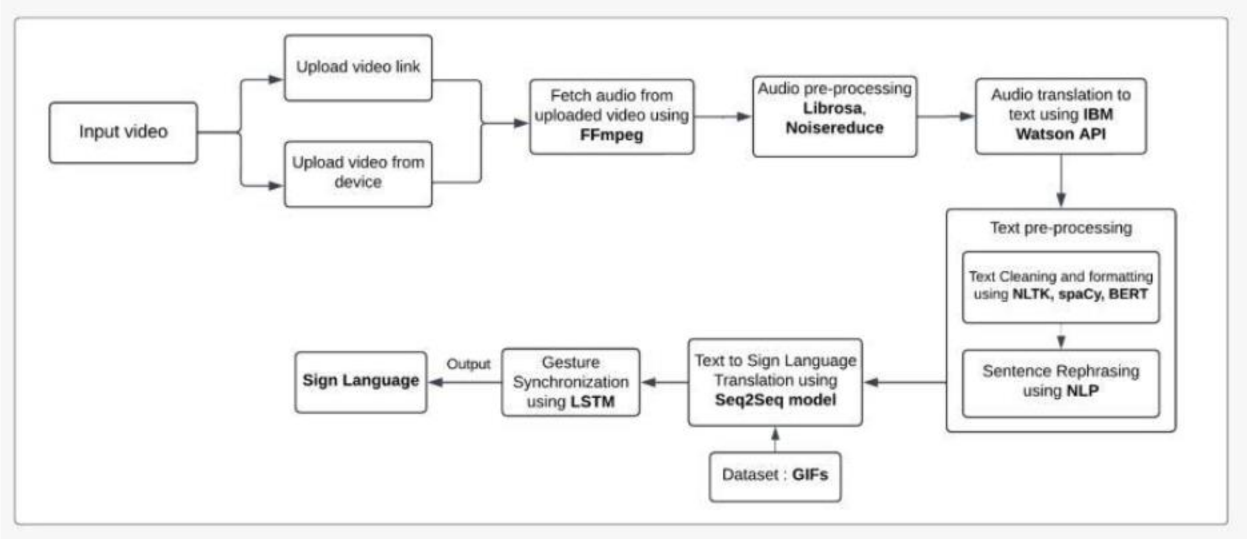


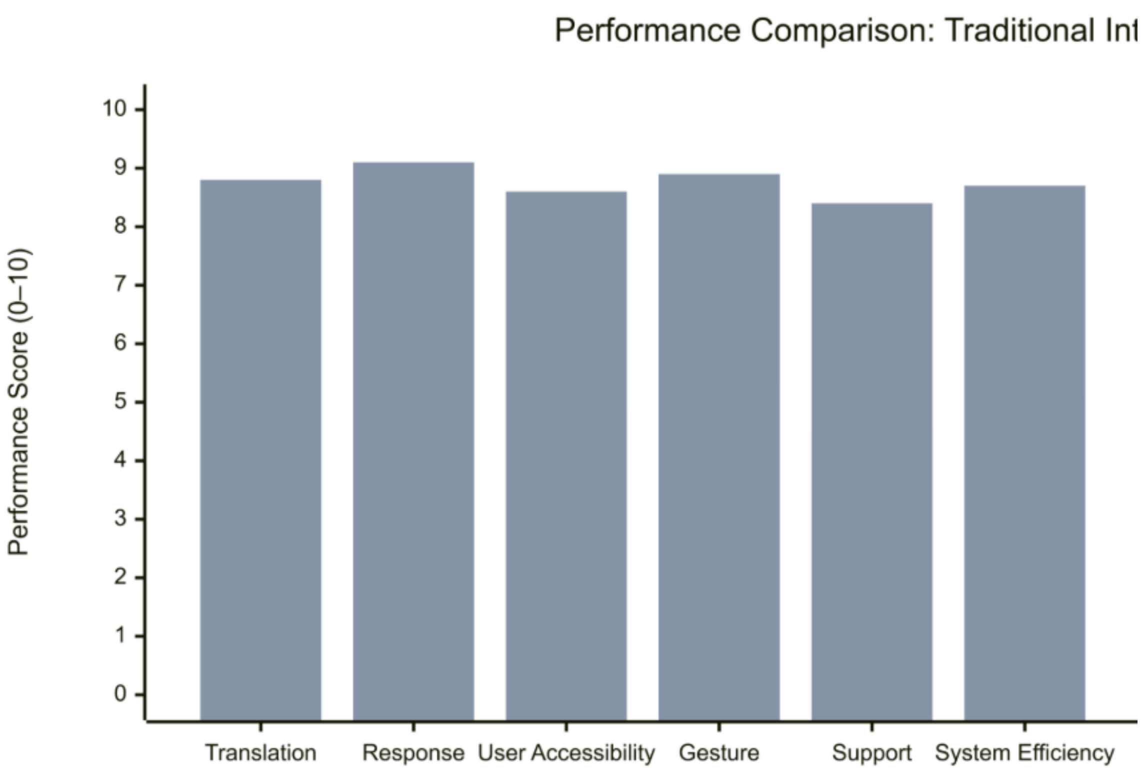
Figure 1: SYSTEM ARCHITECTURE OF AI-GENERATED REAL-TIME SPEECH-TO-SIGN LANGUAGE TRANSLATION USING ANIMATED AVATARS

RESULTS AND DISCUSSION

The successful implementation of the AI-Generated Real-Time Speech-to-Sign Language Translation System demonstrates significant improvements in assistive communication, accessibility, and real-time interaction between hearing individuals and people with hearing or speech impairments. The developed platform provides an integrated digital environment that combines speech recognition, natural language processing, gesture mapping, and animated avatar visualization within a unified architecture. By bringing these components together, the system reduces dependency on manual interpretation methods and enables users to communicate more efficiently through automated sign language generation.

One of the key outcomes of the proposed system is its ability to accurately convert live speech input into textual data using advanced speech recognition algorithms. During system testing, the speech processing module successfully captured spoken sentences through a microphone and generated corresponding text outputs in real time. This real-time speech-to-text conversion significantly improved communication speed compared to traditional interpretation approaches. The integration of Natural Language Processing techniques further enhanced system performance by enabling contextual understanding of sentences, keyword extraction, and grammatical structuring before gesture translation. This linguistic processing ensured that the generated sign gestures were meaningful and relevant to the spoken input.

Another important result observed during the evaluation phase was the effective performance of the gesture mapping engine and animated avatar visualization module. The system successfully mapped processed textual information to corresponding sign language gestures stored in the sign language database. These gestures were dynamically rendered through a 3D animated avatar capable of performing synchronized hand movements and basic facial expressions. The avatar-based representation improved user comprehension and engagement compared to static gesture images or prerecorded videos. Users were able to visually interpret the translated gestures clearly, demonstrating the practical applicability of avatar-driven sign communication.



The platform also incorporated a user-friendly graphical interface that displayed real-time translation results and gesture animations in a structured manner. This interface enabled users to monitor translation progress, interact with the system, and utilize learning support features for understanding sign language gestures. Additionally, the system supported multilingual speech input, allowing users to communicate using different spoken languages while receiving sign language output. This feature improved the inclusivity and scalability of the platform for diverse communication environments such as educational institutions, public service centers, and workplaces.

From a data management perspective, the integration of a database module ensured secure storage of user sessions, speech text logs, and gesture translation records. This functionality allowed users to retrieve previous communication sessions and analyze translation performance over time. The automated report generation feature further enhanced system usability by enabling the creation of structured documentation that can be used for research analysis, learning assessment, or assistive communication evaluation.

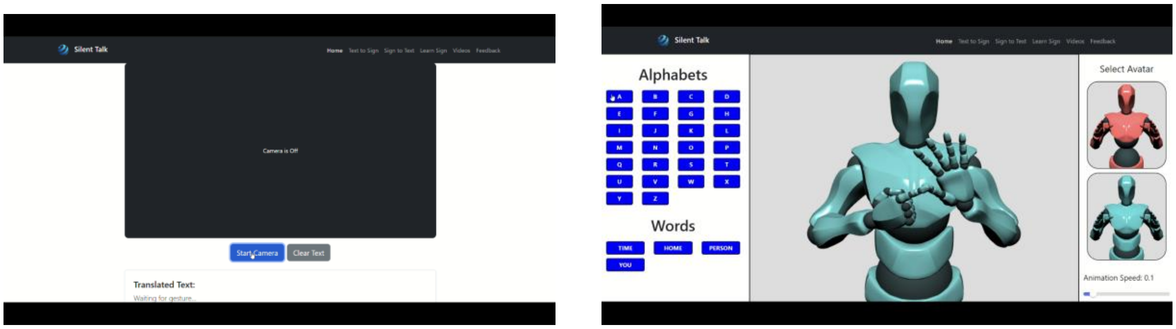


Figure 2 Output of the proposed System

CONCLUSION

The proposed AI-Generated Real-Time Speech-to-Sign Language Translation System successfully demonstrates an intelligent and technology-driven approach for improving communication accessibility between hearing individuals and people with hearing or speech impairments. By integrating speech recognition, natural language processing, gesture mapping techniques, and animated avatar visualization, the system provides an efficient and automated solution for converting spoken language into meaningful sign language gestures. This unified platform simplifies the communication process by eliminating the need for manual interpretation and enabling real-time translation in a user-friendly digital environment. One of the major strengths of the developed system is its ability to process live speech input and generate corresponding visual sign outputs instantly. The real-time translation capability significantly enhances interaction speed and ensures that users can communicate more naturally in dynamic environments such as educational institutions, workplaces, healthcare centers, and public service settings. The integration of Natural Language Processing techniques further improves translation accuracy by enabling contextual understanding of spoken sentences before gesture generation. This ensures that the displayed sign gestures convey appropriate meaning rather than simple word-by-word translation. The use of animated avatars for gesture visualization also plays a crucial role in improving user comprehension and engagement. Compared to traditional assistive tools that rely on static images or prerecorded gesture videos, the avatar-based approach provides dynamic hand movements and basic facial expressions, creating a more realistic and interactive communication experience. In addition, the inclusion of a structured graphical user interface allows users to monitor real-time translation results, access learning support features, and interact with the system efficiently. From a system management perspective, the integration of database storage and automated reporting functionalities enhances the overall usability and scalability of the platform. These features contribute to building a reliable assistive technology framework that can be expanded for large-scale deployment. Overall, the AI-Generated Real-Time Speech-to-Sign Language Translation System serves as a comprehensive assistive communication solution that combines artificial intelligence, speech processing, gesture synthesis, and avatar animation within a single architecture. The system contributes to promoting inclusive digital interaction, reducing communication barriers, and supporting the social integration of hearing and speech-impaired individuals. In future developments, the platform can be enhanced by incorporating advanced deep learning models, emotion-aware avatar expressions, mobile application integration, and cloud-based real-time processing. Such improvements will further strengthen the effectiveness of the system and expand its applicability in smart assistive communication environments.

REFERENCE:

1. Deepa, R., Karthick, R., Velusamy, J., & Senthilkumar, R. (2025). Performance analysis of multiple-input multiple-output orthogonal frequency division multiplexing system using arithmetic optimization algorithm. Computer Standards & Interfaces, 92, 103934.

2. Senthilkumar,Dr.P.Venkatakrishnan,Dr.N.Balaji, Intelligent based novel embedded system based IoT Enabled air pollution monitoring system, ELSEVIER Microprocessors and Microsystems Vol.77, June 2020

3. M. Muthalakshmi, N.Mythili, Gurkirpal Singh, R.Senthilkumar (2025). Innovative Approaches for Evaluating Sugarcane Quality: Utilizing Near-Infrared Spectroscopy to Forecast Brix, Pol, and Fiber Content in Commercial Agricultural Domains. Journal of Food Processing, Wiley, https://doi.org/10.1111/jfpe.70233

4. Senthilkumar Ramachandraarjunan, Venkatakrishnan Perumalsamy & Balaji Narayanan 2022, ‘IoT based artificial intelligence indoor air quality monitoring system using enabled RNN algorithm techniques’, in Journal of Intelligent & Fuzzy Systems, vol. 43, no. 3, pp. 2853-2868

5. N. Nagarani, M. Muthalakshmi , E. S. Vinothkumar and R. Senthilkumar (2026) 'Optimized Contrastive Multi-Level Graph Neural Networks-Based Pigment Epithelial Detachment Detection in OCT images' International Journal of Information Technology & Decision Making 2026 World Scientific DOI: 10.1142/S0219622026500343

6. Sanitha P C; Syed Nageena Parveen; Shaik Thaherbasha; M. Shanmugapriya; T. Kalaivani; R. Senthilkumar, Transparent Nutrition: An Explainable AI-based Diet Tracking System for Preventing Nutrition-Related Disorders. 2025 3rd International Conference on Intelligent Cyber Physical Systems and Internet of Things (ICoICI) DOI:10.1109/ICoICI65217.2025.11252549

7. T. Jayasri; M.R. Archana Jenis; P.B. Aswathy; S. Manoranjitham; Christo George; R. Senthilkumar Identity-First Defense in Zero Trust Security Architecture to Protect Cyberspace 3rd International Conference on Intelligent Cyber Physical Systems and Internet of Things (ICoICI) DOI:10.1109/ICoICI65217.2025.11254505

8. J. Uthayakumar; Swapna; A. Ravikumar; S. Sreeraj; R. Senthilkumar; Babu Pandipati AI-Driven Water Resource Management Systems 2025 2nd International Conference on Computing and Data Science (ICCDS) **DOI:** 10.1109/ICCDS64403 .2025.11209318

9. R.Swathiramya; V.V.Karthikeyan; P.Sumathi; Sruthy K V; Afreen Hussain; R.Senthilkumar Multimodal Machine Learning Models for Intelligent Interpretation of Text, Image and Audio Inputs 2025 5th International Conference on Emerging Research in Electronics, Computer Science and Technology (ICERECT) **DOI:**10.1109/ICERECT65215.2025.11377322

10. Srinju.M; Dr.V.Dhanasekaran; S. Guruprasath; Dr.K.Edison Prabhu; K.J Godlin Debby; Dr.R.Senthilkumar AI-Based Recommendation System for Weight Management Using User Feedback and Health Metrics 2025 5th International Conference on Emerging Research in Electronics, Computer Science and Technology (ICERECT) **DOI:** 10.1109/ICERECT65215.2025.11379842

11. D. Chakladar, P. Kumar, S. Mandal, P. P. Roy, M. Iwamura, and B.-G. Kim, "3D Avatar Approach for Continuous Sign Movement Using Speech/Text," Applied Sciences, vol. 11, no. 8, pp. 1–13, Apr. 2021

12. R. Jain, "Converting YouTube Video to American Sign Language Translation Using Convolution Neural Network and Video Processing," iJournals: International Journal of Software & Hardware Research in Engineering, vol. 10, no. 5, 2022.

13. H. Kotha, S. D. Ponugoti, and V. Krishnan, "Audio to Sign Language Using NLTK," Journal of Emerging Technologies and Innovative Research (JETIR), vol. 10, no. 6, pp. e404–e408, Jun. 2023.

14. R. Bharti, S. Yadav, S. Gupta, and B. Rajitha, "Automated Speech to Sign Language Conversion using Google API and NLP," in Proc. Int. Conf. on Advances in Electronics, Electrical & Computational Intelligence (ICAEEC), Jun. 2019.

15. Ghosh and R. Mamidi, "English To Indian Sign Language: Rule-Based Translation System Along With Multi-Word Expressions and Synonym Substitution," in Proc. 19th Int. Conf. on Natural Language Processing (ICON), New Delhi, India, Dec. 2022, pp. 123–127

16. N. Mehta, S. Pai, and S. Singh, "Automated 3D Sign Language Caption Generation for Video," Universal Access in the Information Society, vol. 19, no. 3, pp. 725–738, 2020, doi: 10.1007/s10209-019-00668-9

17. L. T. Nguyen, F. Schicktanz, A. Stankowski, and E. Avramidis, "Automatic generation of a 3D sign language avatar on AR glasses given 2D videos of human signers," in Proc. 1st Int. Workshop on Automatic Translation for Signed and Spoken Languages (AT4SSL), Virtual, Aug. 2021, pp. 71–81.

18. S. Stoll, N. C. Camgoz, S. Hadfield, and R. Bowden, "Text2Sign: Towards Sign Language Production Using Neural Machine Translation and Generative Adversarial Networks," International Journal of Computer Vision, vol. 128, no. 4, pp. 891–908, Jan. 2020

19. H. Monga, J. Bhutani, M. Ahuja, N. Maida, and H. Pande, "Speech to Indian Sign Language Translator," in Recent Trends in Intensive Computing, M. Rajesh et al., Eds. IOS Press, 2021, pp. 55–60. doi: 10.3233/APC210172.
20. U. S. Prasad, K. A. Anuradha, E. Siddartha, S. S. Reddy, and N. P. Reddy, “Audio/Text to Sign Language,” International Journal of Creative Research Thoughts, vol. 11, no. 5, pp. 1450–1456, May 2023.

21. Khare, V. Kulkarni, and A. Upadhayay, “A Collaborative Augmented Reality System Based on Real Time Hand Gesture Recognition,” Global Journal of Computer Science and Technology, vol. 11, no. 23, pp. 1–5, Dec. 2011.

22. Pathak, A. Kumar, Priyam, P. Gupta, and G. Chugh, “Real-Time Sign Language Detection,” Intl. J. for Modern Trends in Science and Technology, vol. 8, no. 1, pp. 32–37, 2022.

23. L. Goyal and V. Goyal, "Text to Sign Language Translation System: A Review of Literature," Intl. J. Synthetic Emotions, vol. 7, no. 2, pp. 20–28, Jul.–Dec. 2016.

24. Z. Yu, S. Huang, Y. Cheng, and T. Birdal, "Sign Avatars: A Large-Scale 3D Sign Language Holistic Motion Dataset and Benchmark," in Proc. English to Indian Sign Language Rule-Based Translation System, May 2023.

25. S. Kumar, K. Saraswat, and S. Prasad, "Audio Extraction from Video," International Journal of Research in Engineering and Science (IJRES), vol. 11, no. 3, pp. 45–49, May 2023.

26. Pandipati and R. Praveen Sam, "Speech to Text Conversion Using Deep Learning Neural Net Methods,"Turkish J. Computer and Mathematics Education, vol. 12, no. 5, pp. 2037–2042, 2021.

27. S. Sasavade, T. Sutar, K. Baral, and D. Kambale, "Extract the Audio from Video by Using Python," Intl. Research J. Engineering and Technology, vol. 10, no. 6, pp. 120–123, Jun. 2023.

28. R. Zuo, F. Wei, Z. Chen, B. Mak, J. Yang, and X. Tong, "A Simple Baseline for Spoken Language to Sign Language Translation with 3D Avatars," in Proceedings of the European Conference on Computer Vision (ECCV), 2024