



International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 14, Issue 6, June 2025

A Real-Time Voice to 3D Sign Language Conversion

E. Thanmai Reddy¹, T. Sathwik Reddy², Y. Saketh Reddy³, Mr. G. Prasad⁴

U.G. Student, Department of Information Technology, ACE Engineering College, Ankushapur, Telangana, India¹⁻³

Assistant Professor, Department of Information Technology, ACE Engineering College, Ankushapur, Telangana, India⁴

ABSTRACT: The Audio to Sign Language Converter is an innovative Python-based system designed to bridge the communication gap between the hearing and hearing-impaired communities. It facilitates real-time translation of spoken language into sign language using a combination of Natural Language Processing (NLP), computer vision, and advanced machine learning techniques. The system leverages Google Speech Recognition to transcribe audio input with high accuracy, which is then processed for sign language translation. To visually represent sign language, the system employs 3D animated avatars that perform corresponding sign gestures. This approach ensures more natural, expressive, and easily understandable visual feedback for users who rely on sign language. OpenCV is integrated for gesture mapping and processing, enhancing the system's ability to interact in real-time. One of the key features of this solution is its support for multiple regional languages, enabling accessibility across diverse linguistic populations. This multilingual capability makes the system inclusive and practical for use in multicultural environments.

KEYWORDS: Speech Recognition, Natural Language Processing (NLP), 3D Sign Language Animation, Real-Time Translation, Assistive Technology, Hearing Impairment, Gesture Mapping, Multilingual Support, Computer Vision, Python Programming

I. INTRODUCTION

Communication is a fundamental human need, and language plays a vital role in conveying thoughts, emotions, and information. However, for individuals who are hearing-impaired, traditional auditory communication methods pose significant challenges. Sign language serves as the primary medium of communication for the deaf and hard-of-hearing community. Despite its effectiveness within this group, communication with individuals who do not understand sign language—such as those in public services, education, or healthcare—is often limited. This creates a barrier that can lead to social exclusion and lack of access to essential services.

The Audio to Sign Language Converter addresses this challenge by acting as a bridge between spoken language and sign language through the use of advanced technology. The system listens to audio input from a speaker, processes it using Natural Language Processing (NLP) and Google Speech Recognition, and then translates the recognized speech into 3D animated sign gestures using computer vision and OpenCV. This system can be applied in numerous domains such as education, where it can help instructors communicate with hearing-impaired students; healthcare, to assist medical professionals in explaining diagnoses or treatment options; and public services, where it can aid in delivering announcements and instructions clearly to all.

II. LITERATURE SURVEY

1. Sahu, S., Dash, S., & Samal, S. (2020). Real-Time Sign Language Conversion Using Deep Learning Techniques. *International Journal of Advanced Research in Computer Science*, 11(6), 214-219.

Developed a real-time sign language conversion system using deep learning. Their system focused on converting spoken language into sign language gestures, emphasizing accessibility through real-time translation. The system used Python-based libraries and models to identify speech input and output the equivalent sign gestures.

2. Nguyen, H. T., Nguyen, T. V., & Dang, T. H. (2021). Real-Time Sign Language Conversion Using Convolutional Neural Network. In Proceedings of the International Conference on Advanced Computational Intelligence (ICACI), 188-193.

Nguyen et al. further enhanced the field by introducing a real-time conversion model using CNNs. Their system processed live audio inputs, converted them to text using speech recognition, and subsequently transformed that into corresponding sign gestures. The research confirmed the potential of CNNs for fast and accurate gesture mapping from audio input.

3. Zhang, Y., Liu, Y., & Wu, J. (2020). Audio-to-Sign-Language Translation Based on Attention Mechanism. In Proceedings of the International Conference on Artificial Intelligence and Computer Science (ICAICS), 50-54.

Proposed an innovative audio-to-sign-language translation model based on attention mechanisms. This system used Python and deep learning frameworks to focus selectively on important parts of the input audio sequence, thereby improving the accuracy and fluidity of sign generation. The integration of the attention mechanism helped maintain semantic continuity, which is crucial for interpreting spoken language in sign format.

III. METHODOLOGY

The proposed system for the "Real-Time Voice to 3D Sign Language Conversion" system involves a modular, end-to-end pipeline that translates spoken language into animated sign gestures. It begins with capturing real-time audio input through a microphone, which is then processed using the Google Speech Recognition API to convert speech into text. The transcribed text undergoes Natural Language Processing (NLP) to tokenize, lemmatize, and extract contextual meaning. Based on this processed text, the system performs gesture mapping using a predefined sign language dictionary or dataset. Each word or phrase is matched to a corresponding gesture. These gestures are then rendered into realistic 3D animations using tools like Blender or OpenCV to simulate accurate sign language. The animated output is displayed through a user-friendly interface, allowing real-time interaction. The system also supports multilingual input and ensures fast response times for fluid communication. Overall, this methodology integrates speech processing, NLP, and 3D animation to provide an inclusive, accessible solution for hearing-impaired users.

Architecture:

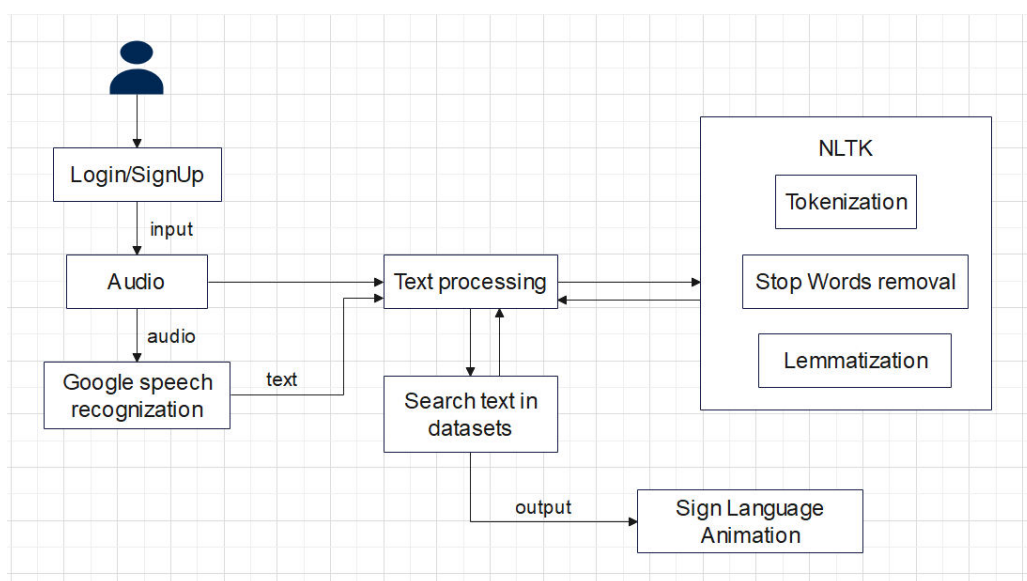


Fig.1 Architecture

IV. EXPERIMENTAL RESULTS

Welcome Interface

The main GUI welcome screen of the application. It displays the title and an option to proceed, indicating the launch point for the user to start the conversion process.

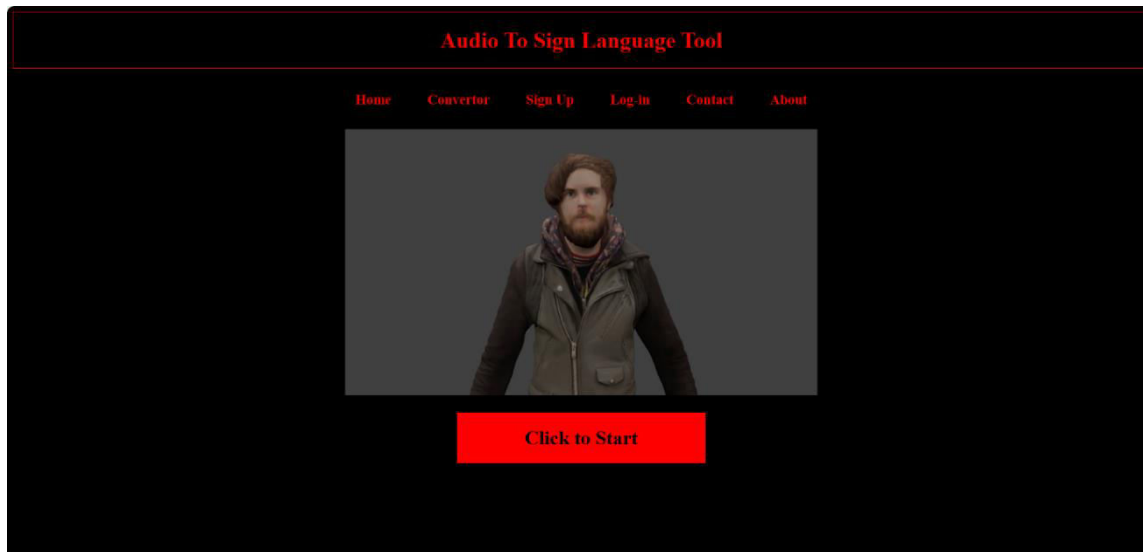


Fig.2 Welcome Interface

Sign Up

The registration screen for new users to create an account. It collects basic details like name, email, username, and password, and stores them for future login.

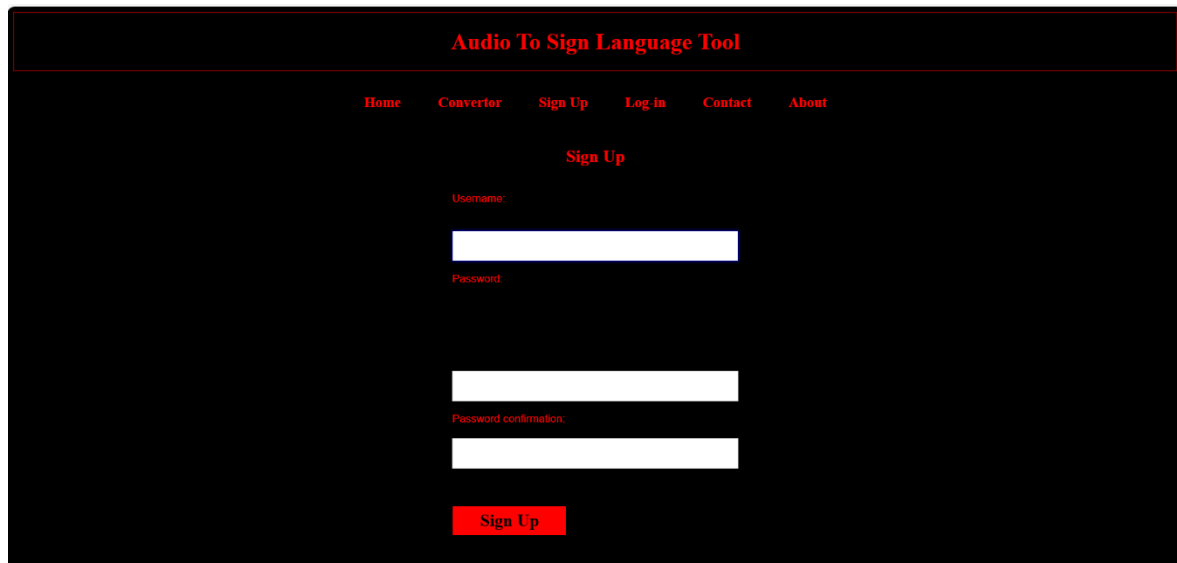


Fig.3 Sign Up

Login

This screen allows existing users to securely access the system by entering their registered username/email and password. It ensures that only authorized users can use the application.

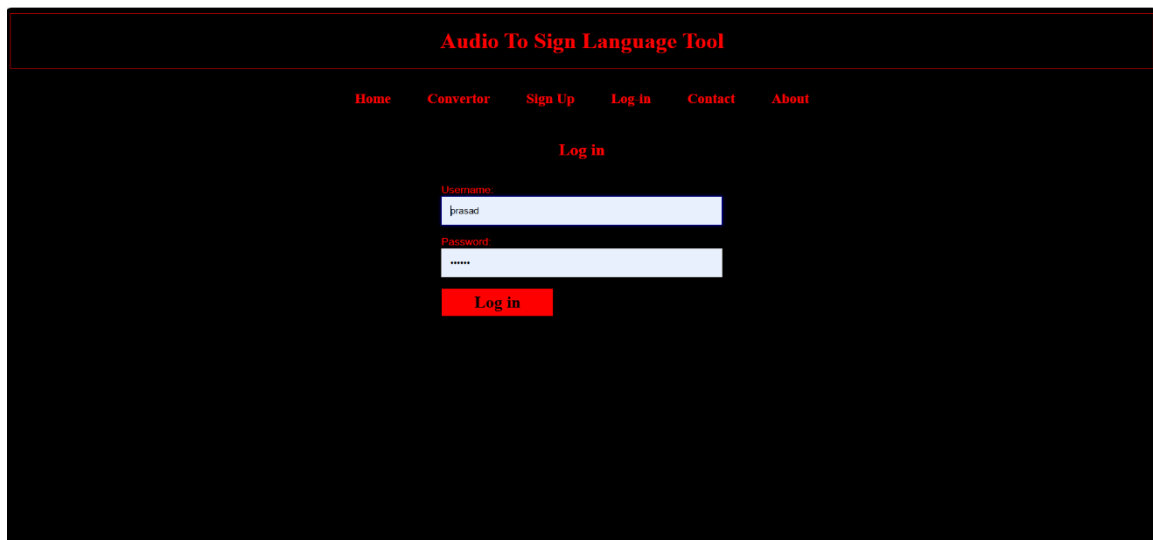


Fig.4 Login

Audio Input Interface

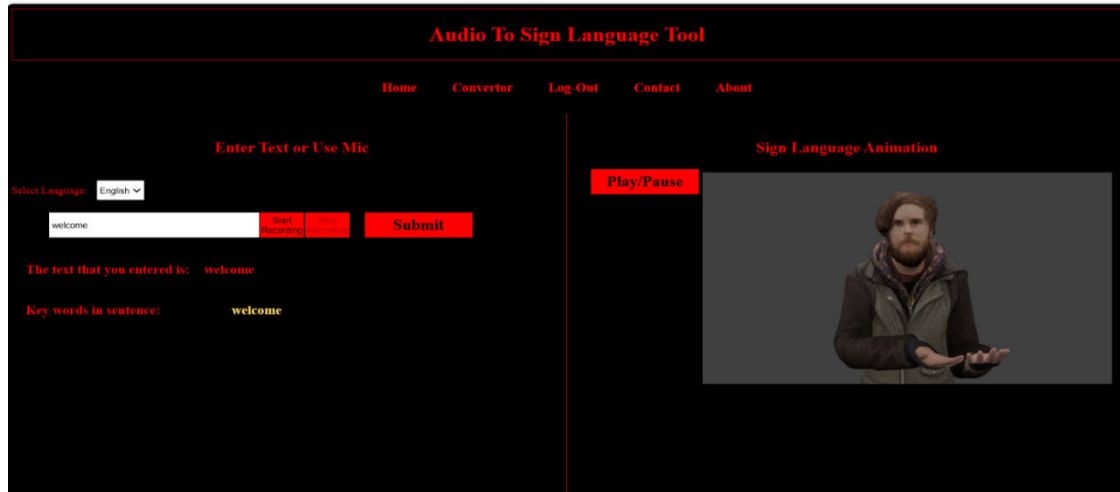
This screen allows users to either record audio in real time or can give it in the text in multiple languages. It acts as the primary input phase for capturing spoken language.



Fig.5 Audio Input Interface

Speech-to-Text Conversion, Text Processing and Mapping

The system processes the audio input and converts it into text using a speech recognition engine. This step verifies that the spoken words are correctly transcribed. The converted text is analyzed and mapped to its corresponding sign language gestures. This intermediate step ensures linguistic context is preserved before visual rendering.



8.6 Sign Language Animation 1

V. CONCLUSION

The Audio to Sign Language Converter is a meaningful step toward bridging the communication gap between hearing and hearing-impaired individuals. By combining speech recognition, natural language processing, and 3D animation, the system translates spoken language into sign language in real time. It offers accessibility across multiple languages and provides a user-friendly interface for effective communication. This project not only demonstrates the practical use of technology for social good but also promotes inclusivity in key areas such as education, healthcare, and public services. With further enhancements, it has the potential to become a widely adopted tool for inclusive communication.

Future improvements can include adding support for more regional languages, expanding the gesture library for complex sentences, and using deep learning models for better context understanding. The project also opens doors for mobile and web-based versions, making it easier to reach users anywhere. With advancements in AI and computer vision, the system can be made smarter and more human-like. Overall, this project sets the foundation for building assistive technologies that ensure no one is left behind in the digital era.

REFERENCES

- [1]. Huang, Y., Zhang, M., Huang, H., & Zhu, J. (2021). Audio-to-Sign Language Translation Based on Deep Learning. In Proceedings of the International Conference on Artificial Intelligence and Computer Vision (AICV), 128-133.
- [2]. Li, H., Zhang, Z., & Wu, S. (2020). Audio-Visual Sign Language Recognition Based on Deep Learning. In Proceedings of the IEEE International Conference on Multimedia and Expo Workshops (ICMEW), 1-6.
- [3]. Sharma, A., & Singhal, R. (2019). Audio to Sign Language Translation using Deep Learning. In Proceedings of the International Conference on Artificial Intelligence and Data Engineering (AIDE), 197-203.
- [4]. Haldar, S., Chaudhuri, S., & Kumar, S. (2020). Real-time Speech to Indian Sign Language Conversion using Deep Learning. In Proceedings of the International Conference on Computational Intelligence and Data Science (ICCIDS), 1-
- [5]. Padole, S. S., & Bansode, S. S. (2021). Sign Language Conversion from Audio using Deep Learning. In Proceedings of the International Conference on Advanced Computing and Intelligent Engineering (ICACIE), 213-218.
- [6]. Xu, J., Li, Y., Yang, J., & Wang, L. (2018). Speech-to-Sign Language Translation using Deep Learning. In Proceedings of the International Conference on Intelligent Computing and Internet of Things (ICIT), 78-83.
- [7]. Zhang, Y., Li, C., & Song, J. (2019). Sign Language Recognition from Speech using Deep Learning Models. In Proceedings of the International Conference on Electronics, Communications and Control Engineering (ICECC), 63-67.
- [8]. Rana, A., Gupta, R., & Dixit, A. (2020). Real-Time Sign Language Recognition using Deep Learning. In Proceedings of the International Conference on Advances in Computing, Communication, & Automation (ICACCA), 133-139.
- [9]. Ghosh, A., & Das, D. (2018). Speech to Sign Language Conversion using Deep Learning Techniques. In Proceedings of the International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), 1465-1469.
- [10]. Laddha, S., Mathur, S., & Bhagat, P. (2020). Audio to Sign Language Translation using Deep Learning Approach. In Proceedings of the International Conference on Inventive Research in Computing Applications (ICIRCA), 829-833.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details