Voice To Sign Language Conversion

Pushpalatha S. Nikkam[1], Varsha Jadhav[2], Aishwarya Ulhas Naik[3], Rahul Hiremath[4]

[1]Assistant Professor, Department of Information Science and Engineering, SDM College of Engineering and Technology, Dharwad, Karnataka, India

[2]Assistant Professor, Department of Information Science and Engineering, SDM College of Engineering and Technology, Dharwad, Karnataka, India

[3]Student ,Department of Information Science and Engineering, SDM College of Engineering and Technology, Dharwad, Karnataka, India

[4]Student ,Department of Information Science and Engineering, SDM College of Engineering and

---

ABSTRACT

True incapacity can be seen as the inability to speak, where individuals with speech impairments struggle to communicate verbally or through hearing. To bridge this gap, many rely on sign language, a visual method of communication that uses hand gestures. Although sign language has become more widespread, interaction between those who sign and those who don't can still pose challenges. As communication has grown to be an essential part of daily life, sign language serves as a crucial tool for those with speech and hearing difficulties. Recent advances in computer vision and deep learning have significantly enhanced the ability to recognize gestures and movements. While American Sign Language (ASL) has been thoroughly researched, Indian Sign Language (ISL) remains underexplored. Our proposed approach focuses on recognizing 4972 static hand gestures representing 24 English alphabets (excluding J and Z) in ISL. The project aims to build a deep learning-based system that translates these gestures into text using the "Google Text-to-Speech" API, thereby enabling better interaction between signers and non-signers. Using a dataset from Kaggle and a custom Convolutional Neural Network (CNN), our method achieved a 99% accuracy rate.

***Key Words*:**
Convolutional Neural Network; Google text to speech API; Indian signing.

1. INTRODUCTION

Voice to Sign Language conversion is a significant advancement in assistive technology aimed at bridging the communication gap between hearing individuals and those with speech or hearing impairments. While sign language provides an effective means of communication for the deaf and mute community, most non-signers find it challenging to understand or use it. This often leads to isolation or dependency for those who rely solely on sign language. With the growing importance of communication in daily life, there is a strong need for tools that can translate spoken language into sign language in real time.

Our goal is to develop a voice-to-sign language system specifically for ISL, which can listen to spoken English, translate it into text, and display the corresponding ISL hand gestures. This approach can play a crucial role in making conversations more inclusive, especially in educational, social, and professional environments where hearing-impaired individuals often face challenges.

2. BODY OF THE PAPER

2.1 METHODOLOGY

The methodology followed in this project consists of a series of structured steps to ensure accurate translation of spoken language into sign language. Initially, data collection is carried out by gathering a diverse dataset of audio recordings and corresponding sign language video samples. Once collected, the data undergoes preprocessing, which includes noise reduction, normalization, and segmentation to improve accuracy. The core translation process employs automatic speech recognition (ASR) models and natural language processing (NLP) techniques to convert spoken input into structured sign language grammar. Machine learning algorithms map the processed text to appropriate sign language symbols, ensuring context-aware translation. Key features are then extracted, focusing on gestures and expressions used in sign language, to generate accurate visual representations. Finally, animated sign avatars or synthesized hand gestures are used to present the translated output effectively, enhancing accessibility for individuals with hearing impairments.

2.2 TECHNOLOGY USED

The SignVoice system integrates several modern technologies to enable efficient and accurate translation of spoken language into sign language gestures. At the core of the system are deep learning models, which are responsible for analyzing audio input and converting it into structured sign language representation. These models are trained on labeled datasets to distinguish phonetic patterns and match them with appropriate signs, providing a high level of translation precision. To enhance accuracy, speech-to-text processing techniques are employed to convert spoken words into textual data, allowing the model to focus on relevant linguistic features and minimize errors. Gesture synthesis is facilitated through motion capture and animation techniques that help generate fluid and natural sign language expressions, ensuring accessibility and effective communication for individuals with hearing impairments..

RESULTS AND DISCUSSIONS

In a machine learning-based speech-to-text-to-sign language conversion system using Indian Sign Language (ISL), the process begins when a user speaks a sentence such as “Where is the hospital?”. The system first employs automatic speech recognition (ASR) to convert the spoken words into text, resulting in: “Where is the hospital?”. Next, natural language processing (NLP) is used to restructure and simplify the sentence according to ISL grammar, which often removes auxiliary words and adjusts word order. In ISL, the phrase is transformed into: “HOSPITAL WHERE”, with facial expressions like a raised eyebrow indicating a question. This processed gloss (structured sign representation) is then sent to a sign language rendering module, where an animated avatar performs the signs using ISL-specific hand gestures, body movements, and facial expressions. The final visual output ensures that individuals who are deaf or hard of hearing receive the spoken message through clear, culturally appropriate ISL communication, enhancing accessibility and interactivity.

3. CONCLUSION

The voice to sign language conversion system offers a powerful solution for bridging the communication gap between hearing individuals and those with speech or hearing impairments. By leveraging speech recognition and deep learning, the system efficiently converts spoken input into corresponding sign language videos, enabling more inclusive and accessible interactions. Whether by matching full sentences or breaking them into individual words, the approach ensures accurate and user-friendly communication

REFERENCES

[1] Girish L , Harshitha C “Sign Language Recognition using Machine Learning”Dept of Computer Science and Engineering, Shridevi Institute of Engineering and Technology, India, Tumkur-572106 Volume 05 Issue 01, FEB 2023

[2] Sim, S., Lee, E. C., & Kim, H. Y. (2020). DeepASL: Enabling Ubiquitous and Non-Intrusive Word and Sentence-Level Sign Language Translation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020.

[3] Pal, U., Talukdar, T., & Kumar, S. (2018). Sign Language Recognition Using Recurrent Neural Networks. IEEE Transactions on Neural Networks and Learning Systems, 2018.

[4] Chen, X., Zhang, G., & Liu, W. (2019). A Multimodal System for Continuous Sign Language Recognition. IEEE Transactions on Multimedia, 2019.

[5] P. Sudhakaran, P. M, A. Kataria, V. Sindhu and R. Mathur, "Converting Voice Signal to Visual Indian Sign Language," 2024 Ninth International Conference on Science Technology Engineering and Mathematics (ICONSTEM), Chennai, India, 2024, pp. 1-6, doi: 10.1109/ICONSTEM60960.2024.10568884.