International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 05 | May - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

SPEECH TO SIGN CONVERSION USING NATURAL LANGUAGE PROCESSING

Amrutha B L[1], Harshitha H M[1], Nisarga G P[1] , Sindhu S R[1], Mrs. G V Sowmya[2]

[1]*UG students, Dept. of ISE, JNNCE, Shivamogga, India.*
[2]Assistant Prof., Dept. of ISE, JNNCE, Shivamogga, India.

---***---

Abstract - Speech to sign conversion is one essential piece of technology that helps the hearing community and the hearing impaired communicate. This paper presents a novel approach utilizing Natural Language Processing (NLP) techniques for speech to sign conversion. The system that is being suggested leverages advanced NLP algorithms to transcribe spoken language into text, which is thereafter converted to sign language. The process involves several key steps: speech recognition and sign language synthesis. The synthesized sign language gestures are generated based on a comprehensive database of sign language vocabulary and grammar rules. The proposed speech to sign conversion system demonstrates promising outcomes concerning accuracy, efficiency, and usability. By harnessing the power of NLP, this technology possesses the capacity to significantly enhance communication accessibility for the hearing-impaired community, facilitating seamless interaction n a number of fields, such as education and healthcare and social communication.

Key Words: Speech Recognition, Natural Language Processing, Tokenization, Lemmatization, Parsing.

1.INTRODUCTION

Sign language serves as a vital bridge for communication among the deaf and hard-of-hearing community, particularly in countries like India, where millions experience hearing impairment. With approximately 6.3% of the population facing hearing challenges, sign language emerges as a crucial visual language system, embodying grammar, syntax, and vocabulary essential for effective communication. The complexity of sign language lies in its diverse forms, including isolated and continuous sign languages, each enabling communication at different levels of complexity[3]. Moreover, sign language reflects local culture, leading to regional variations such as American Sign Language (ASL) and Indian Sign Language (ISL), highlighting the rich linguistic diversity within the deaf community.

Advancements in technology, particularly in Natural Language Processing (NLP) and computer vision, have paved the way for innovative solutions like speech-to-sign language conversion systems. These systems aim to facilitate seamless communication between individuals using spoken language and those relying on sign language, offering promising avenues for inclusive communication platforms. In India, the development of such systems holds significant potential, given the vast population with hearing impairments[1].

The process of converting speech to sign language involves intricate steps, including speech recognition, semantic parsing, and sign language synthesis. Through the integration of NLP techniques, these systems can accurately transcribe spoken language into text and then translate it into sign language gestures, leveraging deep learning models for enhanced accuracy and adaptability. Furthermore, the use of motion capture technologies adds a visual dimension to the translation process, enhancing the accessibility and usability of sign language for both the hearing impaired and those learning sign language[6].

Efforts to develop comprehensive sign language datasets and parallel corpora further contribute to the advancement of speech-to-sign language conversion systems, providing robust training data for machine learning algorithms. The convergence of linguistic research, technological innovation, and cultural understanding holds immense promise for improving communication accessibility and inclusivity for the deaf and hard-of-hearing community[11].

2. Related Work

This section consists about systems that translate sign language how well. Here's a detailed explanation of this section.

Sign Language Translation Systems:

The movement of signs can be accurately produced based on spoken input. In reference [1], researchers developed a system for translating speech into Spanish Sign Language (SSL) by employing a speech recognition component alongside a natural language translator. Reference [2]aims to develop a comprehensive translation system focusing on English audio input converted to English text, parsed into grammatical structures adhering to Indian Sign Language (ISL) rules. It addresses the limited availability of systems translating text or audio to ISL. The system involves modules for grammar representation, stop word removal, stemming, lemmatization, and synonym replacement. Unlike existing systems, it prioritizes converting sentences into ISL grammar for more effective communication with users.

Reference [3] outlines an ongoing project to develop a neural machine translation system that converts text into sign language, presented through a 3D avatar. This innovation aims to enhance access to information for the deaf community, fostering social, educational, and employment inclusion. The project addresses four key aspects: defining a written representation of sign language, building a parallel corpus, creating a neural translation model, and visually rendering the translation with a signing avatar. While initially focused on Brazilian Portuguese to Brazilian Sign Language, the concepts are adaptable to other language pairs.

An automated system[7] for Indian Sign Language (ISL) education is developed due to limited research and tools available. ISL lacks written form, making sign animation essential. The system converts ISL words into HamNoSys, then into SiGML for animation using JA SiGML URL APP. Existing tools like eSIGNEditor lack ISL support, necessitating this custom solution. The system bridges the gap in ISL education by enabling efficient sign language animation for the Indian context.

The authors[9] tackle the challenge of machine translating written English text to American Sign Language (English/ASL) by proposing a statistical machine translation system. The work address existing system limitations and propose a novel approach that involves creating an artificial corpus based on grammatical dependencies rules due to the lack of sign language resources. The system utilizes a parallel corpus for statistical machine translation, employing IBM alignment algorithms enhanced with Jaro-Winkler distances to optimize the training process. A decoder is then implemented to translate English text to ASL using a proposed transcription system based on gloss annotation. Evaluation is conducted using the BLEU metric to assess translation quality.

The paper[10] introduces an architecture leveraging prosodic information from speech audio and semantic context from text to generate sign pose sequences. It incorporates an additional task of predicting Facial Action Units (FAUs) to capture facial expressions crucial in sign language communication. Models are trained on an Indian Sign Language dataset comprising videos with audio and text translations. Evaluation using Dynamic Time Warping (DTW) and Probability of Correct Keypoints (PCK) scores shows superior performance compared to previous models when combining prosody, text, and FAU prediction.

3. Methodology

The proposed System consists of different phases in the framework.

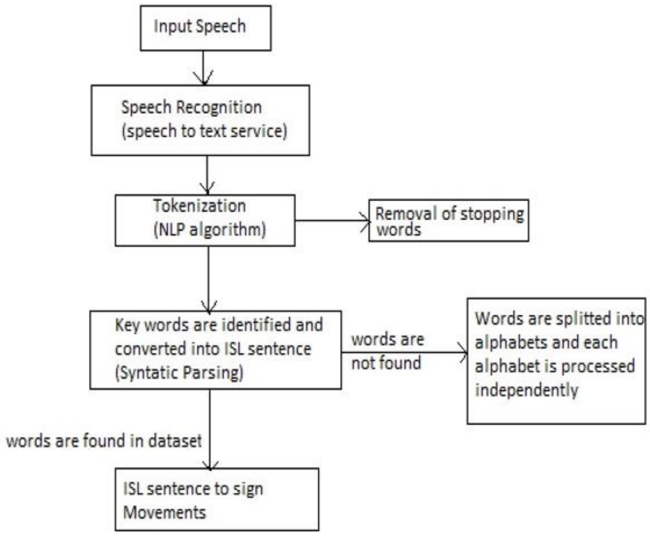


Fig -1: Block Diagram of Proposed Methodology

The Fig 3.1 represents the steps in the implementation of the proposed project as follows:

3.1 Speech to English Sentence Conversion:

Speech recognizer module is used for recognizing the given input speech and then converting it into text. Speech recognition starts by taking the sound energy produced by the person speaking and converting it into electrical energy with the help of a microphone. It then converts this electrical energy from analog to digital, and finally to text.

3.2 Conversion of English Sentence into ISL Sentence:

NLTK (Natural Language Toolkit) is a Python library used for natural language processing tasks such as tokenization, tagging, parsing. Tokenization and tagging are fundamental steps in many NLP pipelines. Tokenization: NLTK offers several tokenizers for breaking text into tokens (words, punctuation, etc.). The tokenizer used as function which splits text into words based on whitespace and punctuation. Tagging: After tokenization, NLTK can assign part-of-speech (POS) tags to each token.

3.3 Sign Movement Generation:

Deep learning models are utilized for image processing during the querying of a database to locate a specific word. If the word is present in the database, an associated image is chosen for integration into a video clip. Conversely, if the word is not found, alphabetic or character images are selected as alternatives.

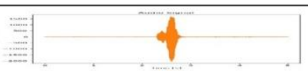
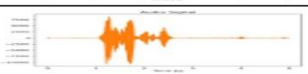
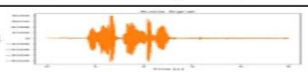
4. Results

The presented analysis focuses on the sign language learning system comprises the sections: sign database, speech recognition results, sign movement generation from ISL sentences.

Sign database: We established a sign language database comprising 1000 words commonly used in daily International Sign Language (ISL) communication. These words include everyday regular usages in terms such as "home", "welcome", "sorry", "rain", "agriculture", "wind". Around 30 frequently used words were collected for each letter of the alphabet. The vocabulary was derived from unique words in ISL. In instances where certain words lack specific sign movements, hand gestures representing the 26 letters of the alphabet are used.

Speech recognition results: We utilized the "SpeechRecognition" Python library to perform speech recognition tasks, converting English audio recordings or files into text format. Our service accepts speech or audio files in .wav format with varying sampling frequencies. In our setup, the audio files have a sampling frequency of 4.1 KHz. The speech recognition module provides outputs for both isolated words (representing discrete speech) and complete sentences (representing continuous speech). We present the results in Table 1. The time in seconds is depicted on the X-axis, while the Y-axis illustrates the signal's amplitude.

Table 1: Transcription of speech signals into text for diverse speech types.

Speech Type	Speech Signal	Output Text
Discrete		Hello
Continuous		Flowers are beautiful
Continuous		He was playing the game

The proposed model translates text between English and International Sign Language (ISL). The translation process involves converting English sentences into ISL and representing ISL sentences as signs. To evaluate the system's performance, the generated sentences were split into training and testing sets, with an 80:20 ratio respectively. The Word Error Rate (WER) was calculated for the input word. This metric measures the accuracy of text processing and the generation of International Sign Language (ISL) sentences. The WER value, which is derived from the Levenshtein distance (edit distance function), was determined to be 21.8 for our project system.

After translating the English sentence into Indian Sign Language (ISL), we proceeded to generate the sign movement for the ISL sentence. Frames were captured for each meaningful word and then combined to generate a continuous sign gesture video. The proposed project boasts an accuracy rate of 98%.

5. Conclusion

The Speech to Sign project has successfully developed a system capable of translating spoken language into sign language, thereby fostering better communication accessibility for the Deaf and hard of hearing community. Through robust speech recognition and sign language generation, our system demonstrates promising accuracy and usability, marking a significant step forward in overcoming communication barriers.

Moving forward, future efforts will focus on enhancing the system's accuracy, expanding its vocabulary and gesture variability, implementing real-time interaction capabilities, and ensuring accessibility across various devices and platforms. These advancements will further improve the inclusivity and effectiveness of the Speech to Sign system, facilitating smoother communication between individuals who use spoken language and those who use sign language.

REFERENCES

1. Debashis das Chakladar, Pradeep Kumar, Shubham Mandal and Partha Pratim Roy, ”3D Avatar Approach for Continuous Sign Movement Using Speech/Text”, applied sciences, MDPI, vol. 11, no. 3439, https://doi.org/ 10.3390/app11083439.
2. Jashwanth Peguda, V Sai Sriharsha Santosh, Y Vijayalata , Ashlin Deepa R N, Vaddi Mounish, “Speech to Sign Language Translation for Indian Languages”, 8th International Conference on Advanced Computing and Communication Systems (ICACCS), 2022, pp. 1131 1135
3. Pankaj Sonawane, Karan Shah, Parth Patel, Shikhar Shah, Jay Shah, “Speech To Indian Sign Language (ISL) Translation System”, 2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), IEEE xplore, 2021, pp. 92-96.
4. José Mario De Martino, Ivani Rodrigues Silva, Janice Gonçalves Temoteo Marques, Antonielle Cantarelli

Martins, Enzo Telles Poeta, Dener Stassun Christinele, João Pedro Araújo Ferreira Campos, "Neural machine translation from text to sign language", Universal Access in the Information Society, Springer, https://doi.org/10.1007/s10209-023-01018-6.

5. María Villa-Monedero , Manuel Gil-Martín , Daniel Sáez-Trigueros , Andrzej Pomirski and Rubén San-Segundo , "Sign Language Dataset for Automatic Motion Generation" , Journal of Imaging ,MDPI, 9, 262. https://doi.org/10.3390/ jimaging91202622023,2023
6. Devesh Chaman Verma, Pratibha Sharma and Nancy Nancy, "Translating Speech to Indian Sign Language Using Natural Language Processing", Future Internet, MDPI, vol.14, no. 9: 253, https://doi.org/10.3390/fi140902532022, 2022.
7. Khushdeep Kaur and Parteek Kumar , "HamNoSys to SiGML Conversion System for Sign Language Automation" , Twelfth International Multi-Conference on Information Processing 2016 (IMCIP), 2016, pp. 794 – 803.
8. Mohammed Faisal, Mansour Alsulaiman, Mohamed Mekhtiche, Bencherif Mohamed Abdelkader, Mohammed Algabri, Tariq Bin Saleh Alrayes, Ghulam Muhammad, Hassan Mathkour, Wadood Abdul, Yousef Alohali, Muneer Al-Hammadi, Hamdi Altaheri , Taha Alfakih , "Enabling Two-Way Communication of Deaf Using Saudi Sign Language", IEEE Access, vol. 11, pp. 135423-135434.
9. Achraf, O., Jemni, M. , "Designing high accuracy statistical machine translation for sign language using parallel corpus: case study English and American sign language", Journal of Information Technology Research https://doi.org/10.4018/JITR.20190 40108 , vol.12(2), pp. 134–158 , 2019,
10. Mounika Kanakanti, Shantanu Singh, Manish Shrivastava, "MultiFacet: A Multi-Tasking Framework for Speech-to-Sign Language Generation", International Conference On Multimodal interaction, ACM digital Library, 2023, https://doi.org/10.1145/3610661.3616550
11. Stephanie Stoll1, Necati Cihan Camgoz, Simon Hadfield, Richard Bowden, "Text2Sign: Towards Sign Language Production Using Neural Machine Translation and Generative Adversarial Networks", International Journal of Computer Vision, Springer, pp. 891-908, 2020, https://doi.org/10.1007/s11263-019-01281-2
12. Yuvraj Grover, Riya Agarwal, Deepak Sharma, Prashant K. Gupta, "Sign Language Translation System for hearing/Speech Impaired People: A Review", International Conference on Innovative Practices in Technology and Management (ICIPTM), 2021, pp 10-14
13. López-Ludeña, V.; San-Segundo, R.; Morcillo, C.G.; López, J.C.; Muñoz, J.M.P. Increasing adaptability of a speech into sign language translation system. Exp. Syst. Appl. 2013, 40, 1312–1322.