

Speech to Sign Language Translator Interface

Jalaj¹, Kavya Shukla² and Vivek Srivastava³

¹⁻³ABES Engineering College, Ghaziabad, India

Email: jalajsc@gmail.com, shuklakavya925@gmail.com, viveksrivastava@abes.ac.in

Abstract—Speech is vocalised form of human communication to express their thoughts and feelings with others. Speech allows individuals to convey thoughts, ideas, emotions, and information to others using spoken language. It is a fundamental aspect of human interaction and can vary significantly across different languages, dialects, and individual speaking styles. However, even while communicating people find it hard to communicate because of different languages and dialects. People with hearing impaired also find it difficult to understand others. This creates a barrier in communication. Sometimes lengthy speech is also hard to follow due to pace, pronunciation etc. Speech recognition is the interdisciplinary study to interpret and understand spoken language into text or commands that a machine can understand and process.

This research paper presents the design, development, and implementation of an application capable of transcribing audio files into text and then to Sign Language using Python programming language. The application utilizes graphical user interface (GUI) components and asynchronous techniques to facilitate audio file selection, transcription from a microphone, and language customization.

Index Terms— Natural Language Processing, Speech Recognition, Text Summarization, Machine Learning, Sign Language.

I. INTRODUCTION

Speech is the most important mean to communicate thoughts with others. Speech Recognition (SR) is a mechanism to identify what the speaker is saying and converting it to text using SR algorithms and converting it to Sign Language So that hearing impaired people can understand. Speech Recognition process helps the machine to recognise different speech based on different words and phrases. The hearing disabled people have hard times communicating with others. They can't express themselves to others as not everyone is not familiar with sign language. The purpose of this project is to make a program which can help disabled people in communication. Deaf people can communicate easily by visual methods so this system can help them to communicate easily.

In recent years, advancements in technology, particularly in the field of Natural Language Processing (NLP), have opened up new possibilities for enhancing communication accessibility for the deaf and hard of hearing community. NLP techniques enable the conversion of spoken language into sign language, offering a valuable tool for improving the communication capabilities of individuals who primarily communicate using sign languages.

In India, where there's a notable gap between the population of individuals with hearing and speech impairments and the accessibility of sign language interpreters, it becomes especially vital to develop a solution. Our research

aims to bridge this divide by designing a translator facilitating seamless communication between non-sign language users and those who rely on sign language, eliminating the necessity for extensive sign language training.

The objective of this research paper is to explore the application of NLP in converting spoken English into an ideal form of sign language. By leveraging NLP algorithms and techniques, we aim to develop a system that facilitates seamless Effortless communication between sign language users and non-users. This system holds the potential to bridge the communication gap between individuals with auditory impairments and the broader community, thereby fostering inclusivity and accessibility in communication.

Furthermore, the proposed system serves as a learning tool for individuals interested in learning sign language, offering a comprehensive and efficient means of acquiring proficiency in Indian Sign Language (ISL). By providing a platform for both communication and education, our system aims to empower individuals with auditory impairments to participate fully in social, educational, and professional settings.

Overall, this research paper seeks to contribute to the advancement of communication accessibility and inclusivity for individuals who are hard of hearing or speech impaired. Through the integration of NLP technology and sign language interpretation, we aim to create a more inclusive society where everyone has the opportunity to communicate and connect effectively, regardless of their auditory abilities.

The "Audio to Sign Language Converter" is an innovative technology that aims to bridge the communication gap between the deaf and hearing communities. Using advanced automatic speech recognition (ASR), natural language processing (NLP), and sign language animation techniques, this system transcribes spoken language into text, comprehends its linguistic nuances, and translates it into sign language, complete with dynamic animations that convey gestures and facial expressions. With real-time processing capabilities and adaptability to regional sign language variations, it empowers deaf individuals to engage in fluid, meaningful conversations, access education, and participate fully in various aspects of life. This transformative technology not only enhances accessibility but also promotes inclusivity and understanding among diverse communities, illustrating the potential of AI to address real-world challenges and foster social inclusion.

This paper aims to detail the construction and functionality of an application built using Python and customtkinter library, designed specifically for audio-to-text transcription tasks.

A. Objective

To bridge the communication gap in India, where sign language interpreters are scarce, this research proposes an NLP-powered translator that converts spoken English to Indian Sign Language (ISL) and vice versa. This system empowers the deaf and hard-of-hearing community to express themselves clearly, access information readily, and potentially even learns ISL, while allowing hearing people to converse with them without needing formal sign language training. This project fills a crucial gap in ISL translation technology, promoting accessibility and inclusivity for all.

This system tackles spoken English to sign language translation in stages: first, highly accurate speech recognition converts spoken words to text. Then, natural language processing analyzes the text, grasping its meaning and context. Next, translation models bridge the gap between written language and sign language, considering the specific grammar and syntax of the target sign language. Finally, the system generates dynamic sign animations or videos that accurately portray the translated signs, including gestures and facial expressions, all in real-time for smooth conversation.

B. Proposed System

The GUI architecture revolves around three core components:

- Left Frame: Functions as a container showcasing project details and hosting control buttons.
- Textbox: Displays transcribed output from both audio sources and second textbox contained processed text.
- Right Frame: Animated result output of Sign Language.

The application's design emphasizes user-friendliness and efficiency. The utilization of Python's Speech Recognition module simplifies transcription tasks while offering a clear and intuitive interface for users to interact with the transcription functionalities.

We have used customtkinter to display the GUI of selecting audio file, microphone, audio selector buttons for controlling GUI. Speech Recognition library to convert Speech to Text is used. We have also used OpenCV to output the Sign Language video on the Screen. Natural Language Processing is used to remove the filler words like 'is', 'am', 'are', 'was', 'were', etc. Words that are in plural form or adjective will be converted to root word which will be easy and helpful in conversion of text to sign language.

TABLE I: LITERATURE SURVEY TABLE

S.No.	Year of publishing	Title	Keywords	Algorithm Discussion	Conclusion
1	2022 [1]	A proposed artificial intelligence-based real-time speech-to-text to sign language translator for South African official languages for the COVID-19 era and beyond: In pursuit of solutions for the hearing impaired	COVID-19, South Africa, artificial intelligence, hearing impaired, machine learning, sign language, speech, text, translation.	This system has been created to convert every word received as input into sign language, specifically translating based on the Indian sign language for this project.	This technology serves to close the communication divide between people with normal hearing and those who are hearing-impaired.
2	2022 [2]	Translating Speech to Indian Sign Language Using Natural Language Processing	Indian Sign Language; natural language processing; tokenization; lemmatization; parsing	Natural Language Processing	This system is based on natural language processing the grammar rules of Indian Sign Language. Its incorporation into various domains such as hospitals, buses, railway stations, post offices, and even video conferencing applications represents a significant integration.
3	2021 [3]	Speech To Sign Language Translator For Hearing Impaired	Speech to Sign Language Translation, Speech Recognition, Indian Sign Language, Natural Language Processing	Automated Language Translation and Computational Linguistics.	The paper successfully attains its objective by creating a system capable of speech recognition, text conversion, and displaying Indian sign language.
4	2023 [4]	Implementation of Video and Audio to Text Converter	Video to text, Audio to text, Python, Tkinter.	A graphical user interface (GUI) using tkinter can enhance the accuracy of an audio-to-text converter. This converter accurately transcribes audio into text.	This program provides a basic, minimally operational example.
5	2021 [5]	SPEECH TO TEXT CONVERSION AND SENTIMENT ANALYSIS ON SPEAKER SPECIFIC DATA	Speech Recognition, Sentiment Analysis , Natural Language Processor.	Speech Recognition Module, Sentiment Analysis Module, Google Search API, Python Speech Recognition module are used.	This research introduces a universal system that utilizes user-provided audio input to analyze spoken content by automatically converting speech into text.
6	2019 [6]	Speech to text conversion and summarization for effective understanding and documentation	Feature extraction, Natural language processing, Natural language toolkit, Speech recognition, Text summarization	Google API, NLP (Natural Language Processing)	The objective of the proposed research is to minimize the time and labor involved in manually documenting lengthy speeches during an event.
7	2020 [7]	Sign Language Production using Neural Machine Translation and Generative Adversarial Networks	Motion capturing data, deep learning, Neural machine translation	Neural Machine Translation, RNN based machine translation methods	Introduced the inaugural system for translating spoken language into sign language via video. The predictions made by the NMT network effectively address the MG, leading to reliable text-to-pose translations.
8	2021 [8]	Speech to Indian Sign Language Translator	Computational Linguistics, Lemmatizing, Natural Language Processing, Python, Sign Language, Web-scraping	Natural Language Processing, Speech to text Conversion, Parsing, Stemming and Lemmatizing	In order to maintain consistent video quality, incorporating animated characters to portray sign language can be beneficial. Given the slight variations in Indian Sign Language (ISL) dialects across different regions of India, the system can be adapted accordingly to better serve the needs of the hearing and speech impaired individuals in those areas.

9	2021 [9]	Audio to Sign Language Translation	Machine Learning, Natural Language Processing, Audio-To-Text, Indian Sign Language.	Tokenization, Lemmatizer, Eliminator, mapping word	A Machine learning based automated system for audio to sign language conversion is proposed in this paper.
---	----------	------------------------------------	---	--	--

C. Methodology

We propose a novel speech-to-sign language converter with precise functionality, integrated with a customtkinter-based GUI for user interaction. Its benefits include straightforward execution, an interactive interface, and an accurate conversion system, among other features.

Module Description:

- Audio to Text conversion
- Pre-processing of Text
- Porter Stemming Algorithm
- Text to Sign Language Convertor
- Graphic generation
- GUI output display

Audio to Text Conversion:

In this we are using Speech Recognition library of python. While some speech recognition systems like PocketSphinx operate offline, their installation process can be complex, demanding multiple dependencies. In contrast, Speech Recognition library stands out as a more user-friendly option, requiring an active internet connection for operation.

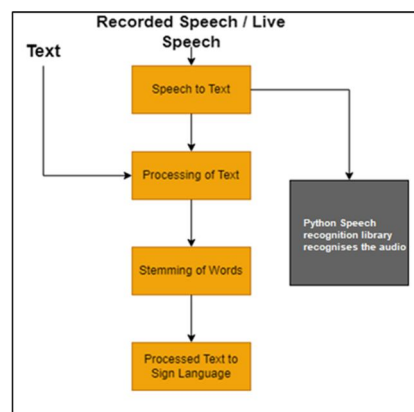


Figure 1: Speech to text

Pre-processing of Text:

Filler words, often used to bridge gaps in sentences, tend to be less meaningful. English has over 30 filler words that add little context. Eliminating these words can enhance sentence clarity and efficiency, saving time in the process.

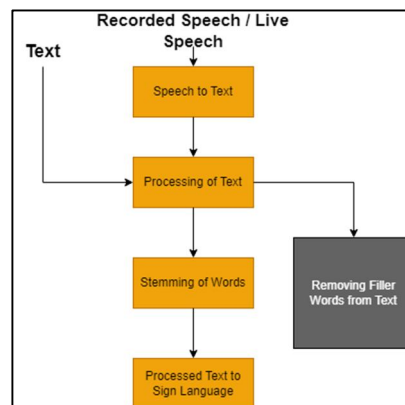


Figure 2: Removing filler words

Porter Stemming Algorithm:

This algorithm excels in removing morphological and inflexional endings from English words. Employing the Porter Stemming Algorithm, system efficiently eliminates common suffixes and prefixes, arriving at the root or original word. For instance, it condenses "true," "truly," and "trusty" to the root word "truth." This stemming process contributes to a faster search for sign language equivalents of given words, ultimately optimizing the time required for such tasks.

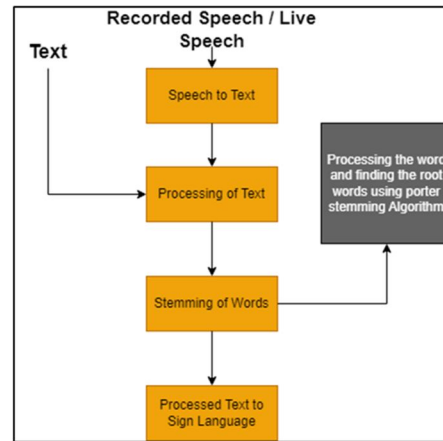


Figure 3: Stemming of words

Text to Sign Language:

We utilize OpenCV to transform text into Sign Language, modifying the text according to specific criteria. The system examines each word within the processed text obtained from the prior step, searching for corresponding sign language image sequences stored locally. Upon identifying a match, the system presents the output as a series of animated videos, utilizing Python's OpenCV module.

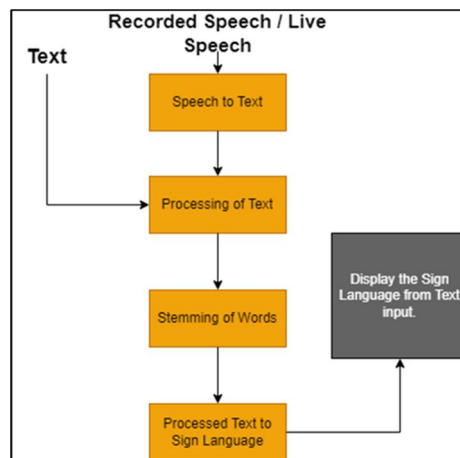


Figure 4: Display of Sign Language gestures

Graphic Generation:

Using blender 3D tool, we created animated gesture representing character for sign movements of each words. Process of creating animation:

1. Open Blender3D
2. Choose an animation character
3. Add the action and animation
4. Export as mp4

GUI output display

The GUI comprises three primary components: the left frame with control buttons, a textbox and a frame of Sign Language. The left frame acts as a contain, showcasing the project name and housing the control buttons.

The textbox serves to exhibit the output text derived from audio sources. The right box contain Sign Language animation derived from Text as source.

II. RESULTS

The application's design emphasizes user-friendliness and efficiency. The utilization of Python's SpeechRecognition module simplifies transcription tasks while offering a clear and intuitive interface for users to interact with the transcription functionalities.

Figure 6 to 8 shows the GUI of the system developed.

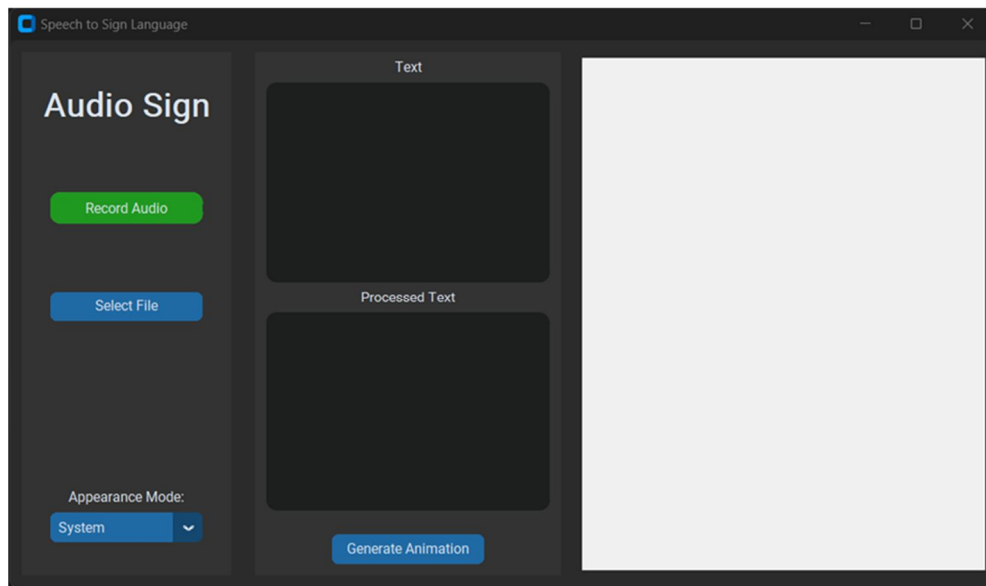


Figure 5. Main GUI of the system

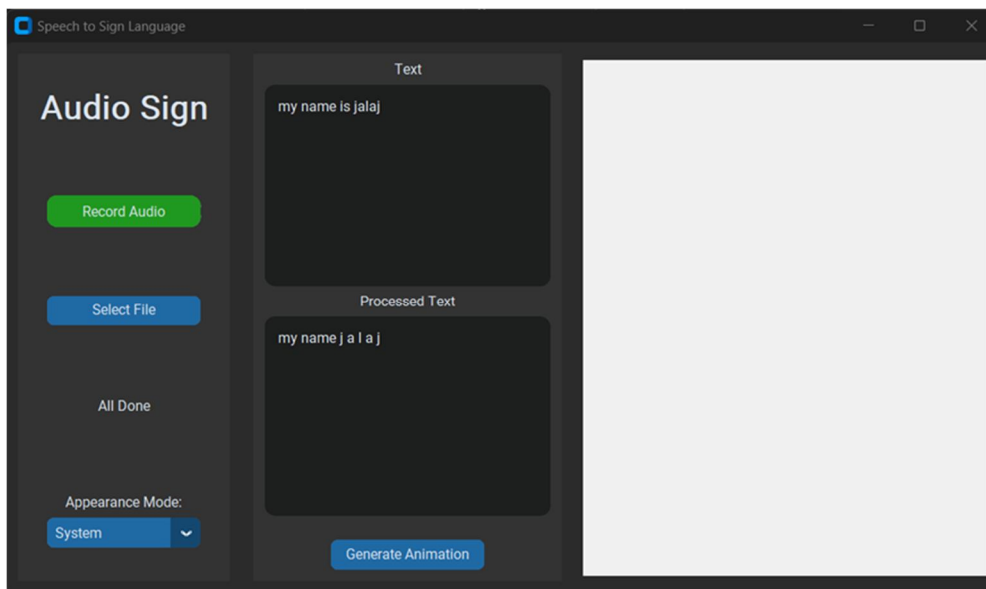


Figure 6. Speech to text and processing of text

Figure 7 shows speech to text using speech recognition library and processing of text using stemming and lemmatization of text to form a processed text.

The database contains all the animated videos for each word and output will be sequentially put in the output canvas. As example, phrase “My name is Jalaj” is given input from speech. This Program processes the text and pops each separate word with not useful word eliminated as “my name j a l a j”.

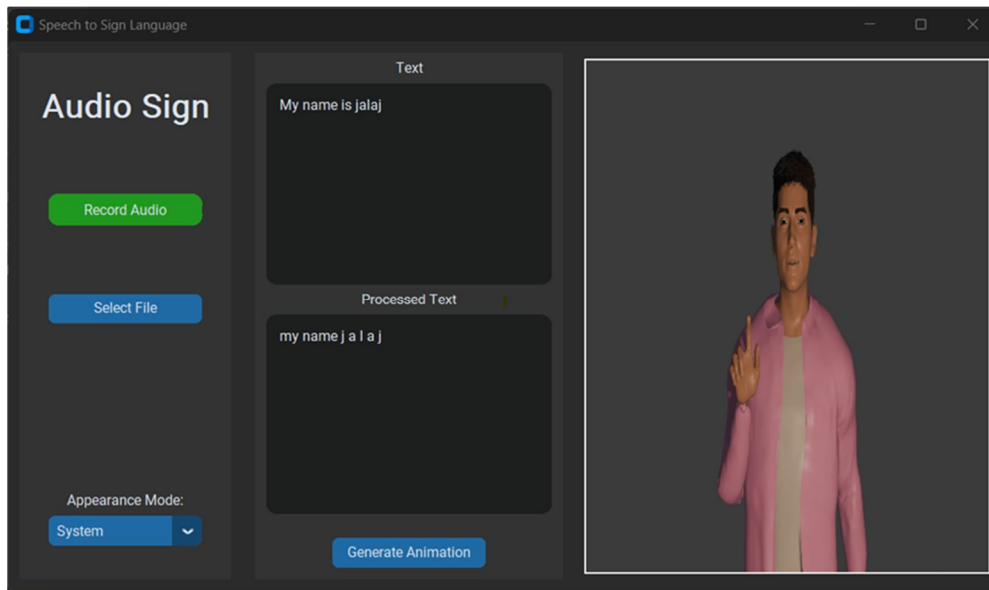


Figure 8: Sign Language gesture from processed text

Video for “my”, “name” is in the database so it won’t be separated. Figure 8 shows the output animation. The Application works perfectly converting Speech or Text to Sign Language which are Images of given word in Sign Language.

III. CONCLUSION

This project offers a solution to the challenges encountered by individuals with hearing and speech impairments through the integration of speech recognition and image processing techniques using Python programming. Speech recognition technology involves computers translating voice signals into related text or commands by identifying and understanding spoken words.

Sign Language plays a pivotal role in fostering an inclusive society, granting individuals with disabilities equitable opportunities for growth and development, enabling them to lead productive, secure, and dignified lives. However, India currently has only 304 licensed interpreters.

An easy to use user interface program is presented in this paper. The NLTK tool makes it so much easier to make machine understand human language. Natural language toolkit makes translation easier and faster. There is a high likelihood that these systems will revolutionize the personal experiences of millions of individuals with speech and hearing impairments.

FUTURE WORK

Future endeavours involve creating a system that can convert sign language along with speech simultaneously. These systems can revolutionize the world. News channels can use that system to show news to hearing impaired masses.

ACKNOWLEDGMENT

We owe special debt of gratitude to our Professor Mr. Vivek Srivastava Department of Computer Science & Engineering for guiding us for her constant support and guidance throughout the course of our work. We also take the opportunity to acknowledge the contribution of Professor (Dr.) Divya Mishra, Head, Department of Computer Science & Engineering for her full support and assistance during the development of the project.

REFERENCES

- [1] Madahana, M., Khoza-Shangase, K., Moroe, N., Mayombo, D., Nyandoro, O., & Ekoru, J. (2022). A proposed artificial intelligence-based real-time speech-to-text to sign language translator for South African official languages for the COVID-19 era and beyond: In pursuit of solutions for the hearing impaired. *South African Journal of Communication Disorders*, 69(2), 915.

- [2] Sharma, P., Tulsian, D., Verma, C., Sharma, P., & Nancy, N. (2022). Translating Speech to Indian Sign Language Using Natural Language Processing. *Future Internet* 2022, 14, 253.
- [3] Ezhumalai, P. (2021). Speech to sign language translator for hearing impaired. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, 12(10), 1913-1919.
- [4] Dr. M. Saraswathi, VVSV Ronit and S Sai Pranav. (2023) Implementation of Video and Audio to Text Converter. *International Journal of Research Publication and Reviews*, Vol 4, no 5, pp 1204-1208 May 2023
- [5] Nikhil Jain, Many Goyal, Agravi Gupta and Vivek Kumar. 2021. SPEECH TO TEXT CONVERSION AND SENTIMENT ANALYSIS ON SPEAKER SPECIFIC DATA. *International Research Journal of Modernization in Engineering Technology and Science* Volume:03/Issue:06/June-2021
- [6] A, Vinnarasu & Jose, Deepa. (2019). Speech to text conversion and summarization for effective understanding and documentation. *International Journal of Electrical and Computer Engineering (IJECE)*. 9. 3642. 10.11591/ijece.v9i5.pp3642-3648.
- [7] Stoll, S., Camgoz, N. C., Hadfield, S., & Bowden, R. (2020). Text2Sign: towards sign language production using neural machine translation and generative adversarial networks. *International Journal of Computer Vision*, 128(4), 891-908.
- [8] Kulkarni, Alisha & Kariyal, Archith & Dhanush, V & Singh, Paras. (2021). Speech to Indian Sign Language Translator. 10.2991/ahis.k.210913.035.
- [9] Kumaresan, G., Venkat, R. D., & Yamini, G. (2021). Audio to Sign Language Translation. *i-manager's Journal on Pattern Recognition*, 8(1), 31.