



Audio To Sign Language Converter Using Indian Sign Language

¹Athira C Achari, ²Ms. Nurjahan V A

¹Student, ²Assistant Professor

¹Department Of Computer Science and Engineering, ²Department of Artificial Intelligence,

¹Illahia College Of Engineering and Technology, Muvattupuzha, India

Abstract: Communication obstacles between the general public and the hearing-impaired community seriously impede accessibility, inclusion, and equal opportunity in a multicultural and multilingual nation like India. Many deaf people use Indian Sign Language (ISL) as their primary form of communication, yet ISL is still not widely understood or used in everyday interactions. In order to close this gap, this project presents an Audio to Indian Sign words (ISL) Converter that allows hearing and non-hearing people to communicate easily by translating spoken words into ISL in real time. Natural language processing (NLP), ISL gesture rendering, and speech recognition make up the system's three main parts. Initially, a speech-to-text engine tailored for regional pronunciations and Indian accents is used to process audio input. After that, the material is condensed and organized using ISL grammar, which is very different from spoken Hindi or English. The matching ISL indicators are mapped to this processed text. Lastly, a 3D animated avatar that can make dynamic motions and non-manual signals like facial expressions or pre-recorded ISL video clips are used to visually represent the signs.

I. INTRODUCTION

Millions of people in India suffer from hearing problems, making it difficult to communicate with the hearing population on a regular basis. Many people who are deaf or hard of hearing rely on Indian Sign Language (ISL) as their primary means of communication. Nonetheless, the communication gap is still substantial because of the general public's low level of ISL usage and knowledge. For the community of people with hearing impairments, this disparity impacts social inclusion, work prospects, public services, and educational chances.

The development of systems that can overcome this communication barrier is becoming more and more possible as speech recognition, natural language processing, and animation technologies evolve. In order to facilitate inclusive interaction in public settings, educational institutions, healthcare facilities, and online platforms, an audio to Indian sign language converter seeks to translate spoken language into ISL in real time.

Three main modules make up the suggested system: visual sign rendering, text-to-sign language translation, and speech-to-text conversion. Spoken input is transcribed using speech recognition software like OpenAI's Whisper or Google Speech API. After that, the text is processed and reorganized to conform to ISL grammar and syntax, which are different from those of Hindi or English in terms of word order and expression. Lastly, either video-based ISL signals or a 3D animated avatar that makes the gestures and facial expressions necessary for clear conversation are used to create the translated output.

II. LITERATURE REVIEW

Systems for translating and recognizing sign language have been investigated in a number of studies to help the hearing impaired communicate. Using sensors and computer vision, early research concentrated on text-to-sign or gesture recognition. More recent studies use machine learning or rule-based models to combine speech recognition with the creation of sign language. However, compared to American Sign Language (ASL), Indian Sign Language (ISL) has gotten less attention. This is mostly because there aren't enough standardized datasets or tools available. ISL video dictionaries and corpora have been given by ISLRTC and IIIT-Hyderabad projects. The accuracy of speech-to-text for Indian accents has increased thanks to Google's Speech API and OpenAI's Whisper. End-to-end ISL translation from audio is still a developing field with enormous potential for societal effect, notwithstanding these developments.

The study, "Audio to Sign Language Conversion using Natural Language Processing," was proposed by Nayana J., Suparna Bhat, Rekha R. Nair, and Tina Babu. Its goal was to assist people with hearing loss in creating a communication system. Signs are created from the spoken phrase. It transforms spoken language into signs. This system uses voice inputs, turns the music into text, and uses images that predate GIFs to depict the important symbols in the Indian language.

The "Audio to Sign Language Translator Web Application" was proposed by Anju Yadav, Rahul Saxena, Bhavna Saini, Vivek K. Verma, and Vibhav Srivastava. The web application aims to create an automated translation mechanism that incorporates a parser element that transforms incoming speech data or English text into a phrase structure grammar representation. This representation is then utilized by another module that contains grammatical format for Indi Sign language.

The "Audio /Text to Sign Language" proposal by U Shiva Prasad, DR K Anuradha, E Siddartha, S Sahith Reddy, and N Prashanth Reddy was centered on Our goal is to assist those who are struggling with hearing loss. In a number of sign language ventures, sign language input has been converted to text or audio. However, audio-to-sign language translation systems have only sometimes been created. Both hearing persons and those who are hard of hearing benefit from it. In this project, we provide a cutting-edge Python-based audio-to-sign language translator. It takes audio as input, uses the Google API to search the recording, displays the text on the screen, and then uses an ISL (Indian Sign Language) generator to create the sign code for the input.

"Audio to Sign Language Translation for Deaf People" was the proposal put out by Ankita Harkude, Sarika Namade, Shefali Patil, and Anita Morey. The goal of the study was to assist those who were experiencing hearing problems. Numerous sign language projects have been completed that translate sign language input into text or audio output. However, there haven't been many systems created that convert audio to sign language. Both normal and deaf persons can benefit from it. In this project, we present a novel technique that uses Python to translate audio to sign language. This takes audio as input, uses the Google API to search for the recording, displays the text on the screen, and then uses an ISL (IndianSignLanguage) generator to provide the sign code for the input.

"Speech to Indian Sign Language Translator" was proposed by Hemang Monga, Jatin Bhutani, Muskan Ahuja, Nikita, Maida, and Himangi Pande. The goal of the research was to develop a model that would enable individuals with disabilities to express themselves uniquely while also facilitating their seamless integration into society. The provided input audio will be successfully transformed into an animation by our suggested model. As the ISL Dictionary expands, other enhancements can be added along the way. Since the ISL's vocabulary is limited, there is room for growth by expanding its vocabulary with new terms.

G Sanjana², SEJAL KADAM³, K NANDHINI⁴, M. Vaishnavi⁵, and Mrs. C. Rashmi¹ suggested "AUDIO TO SIGN LANGUAGE TRANSLATOR USING PYTHON." The study concentrated on using a speech-to-text API to convert the audio signals received into text. Small, medium, and large vocabulary conversions are all included in speech to text conversion. After processing or accepting the voice, these systems translate it into the appropriate text. The technologies utilized in speech recognition systems with short, medium, and big vocabulary are compared in this research. The comparative analysis establishes the advantages and disadvantages of each strategy to date.

III. PROBLEM STATEMENT

Because Indian Sign Language (ISL) is not widely understood or used by the general public, people with hearing impairments in India suffer substantial communication difficulties. Most people communicate primarily through spoken language, but those who use sign language for daily interactions cannot access it. Existing technology solutions do not offer real-time audio-to-sign translation or enable ISL. Furthermore, American or British Sign Language, which are inapplicable in the Indian context because of grammatical and linguistic variations, are frequently the only sign language systems now in use.

To encourage inclusion and accessibility for the deaf community in India, a system that can reliably translate spoken audio—especially in Indian languages or accents—into ISL in real time is desperately needed.

IV. RESEARCH METHODOLOGY

Speech recognition, text processing, and ISL gesture rendering are the three main phases of the modular approach taken by the suggested system. A speech-to-text engine, like Google Speech API or OpenAI Whisper, that is tuned for Indian accents is used to first record and then transcribe audio input. Next, rule-based simplification and keyword extraction are used to align the transcribed text with the grammar of Indian Sign Language. The matching ISL movements are linked to this processed text. Ultimately, either pre-recorded ISL video clips or a 3D animated avatar that can display both manual and non-manual indications are used to graphically represent the translated output. The system's correctness, speed, and user comprehension in real-world situations are assessed by testing on typical ISL datasets.

4.1 Proposed System

By translating spoken audio into Indian Sign Language (ISL) in real time, the proposed technology seeks to close the communication gap between the deaf and hard-of-hearing communities and hearing people. The following are the main parts of the system:

1.Module for Speech Recognition

Using a speech-to-text engine like Google Speech API or OpenAI Whisper, this module records spoken sounds and turns it into text. For increased accuracy, the engine is adjusted or taught to handle regional accents and Indian English.

2.ISL Grammar Mapping and Text Processing

To conform to the grammatical structure of ISL, the transcribed text is processed. This entails rearranging words to conform to ISL syntax, simplifying, and extracting keywords. This stage guarantees that the translated material is accurate in terms of both semantics and context in ISL.

3.Rendering Module for Sign Language

Following the text's translation to ISL, the associated indicators are displayed. You can accomplish this by using: ISL video clips that have already been recorded for every word or phrase, or More realistic and expressive communication is ensured by a 3D animated avatar that can make ISL motions in addition to the required body language and facial expressions.

4.Interface for Users (UI)

To record audio input and show the results of the ISL translation, a straightforward and user-friendly interface is created. To make the system usable in a variety of public and private contexts, the user interface (UI) can be created for desktop or mobile platforms.

Real-time performance, scalability, and adaptability are guaranteed by this modular architecture. It is especially well-suited for use to encourage inclusive communication in government buildings, healthcare facilities, and educational institutions.

4.2 Architecture / Implementation

A modular pipeline is used in the architecture of the Audio to Indian Sign Language (ISL) Converter to provide seamless and instantaneous translation from spoken input to ISL output. The input layer, processing layer, sign mapping layer, and output layer are the four main layers that make up the system.

1. *Input Layer*, Function: Uses a microphone to record audio in real time. Technology: An audio recording module built into a desktop or mobile interface using Python or JavaScript.
2. *Module for Speech Recognition*
 1. Function: Text to audio conversion.
 2. Tools: CMU Sphinx (offline), Google Speech-to-Text API, or OpenAI Whisper (open-source).
 3. Customization: For increased accuracy, it has been adjusted for regional accents and Indian English.
3. *The Text Processing & ISL Mapping Module* : streamlines and cleans the text that has been transcribed. Sentences are rearranged and structured using ISL grammar. connects words and phrases to a lexicon in sign language. , Tools: bespoke ISL grammar rules and dictionaries; spaCy for NLP tasks; Natural Language Toolkit (NLTK).
4. *Sign Rendering Module*, which shows ISL gestures and offers the following options: Recorded ISL Videos in advance: For every word or phrase, a video clip 3D Animated Character: Avatar animation in real time with programs like WebGL, Blender, or Unity3D. The avatar can make facial expressions and physical signs.
5. *User Interface*: Audio capture and ISL output display Tools: Developed with Python (Tkinter/Flask) or web technologies (HTML/CSS/JavaScript) for desktop and mobile applications.

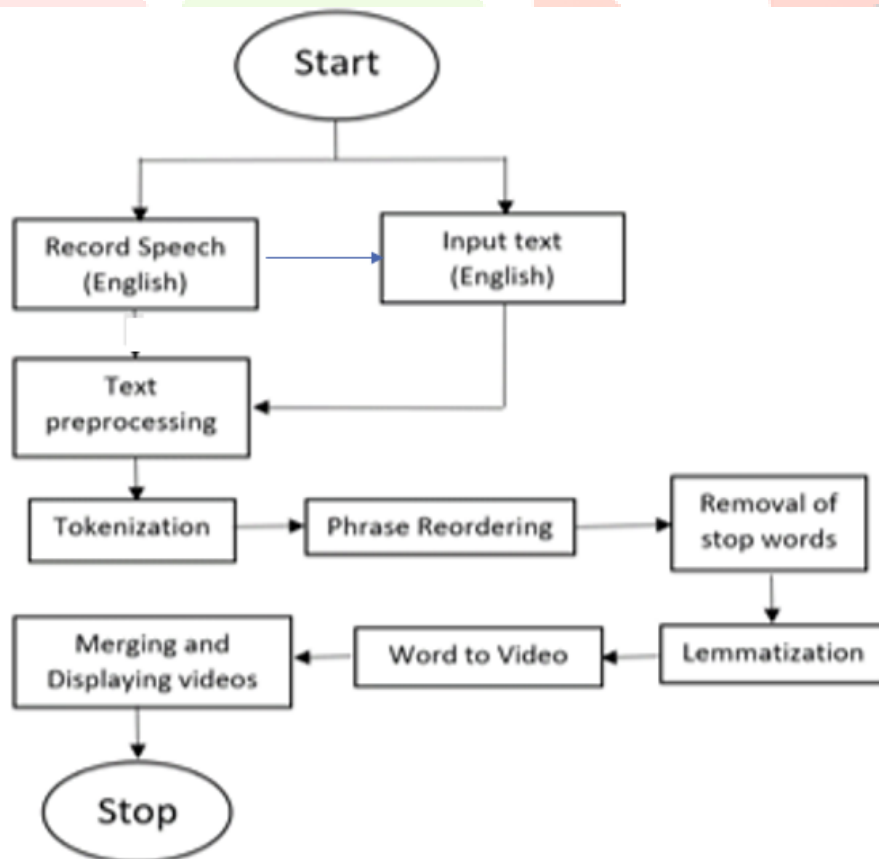


Figure 1: System Architecture

4.3 Data Flow

Either recorded speech or English text input is used to start the system. Tokenization, phrase reordering, stop word removal, and lemmatization are among the text preprocessing techniques used to align the input with Indian Sign Language grammar. After that, a matching ISL video segment is assigned to each word. Users can then observe the translated sign language output when these video snippets are combined and shown sequentially. After the video is played, the process is over.

- 1.Options for Input: Record Speech (English): Uses a microphone to record spoken input, then uses speech recognition to turn it into text.English-language input text can be entered or prewritten.The Text Preprocessing step receives both inputs.
- 2.Preprocessing Text: Prepares the text for additional processing (such as lowercasing and punctuation removal).
- 3.Tokenization: Divides the sentence into meaningful units (tokens) or individual words.
- 4.Reordering Phrases : Modifies word order to conform to the syntax of Indian Sign Language (ISL), which is different from English.
- 5.Elimination of Stop Words: Removes frequently used terms that are not signed in ISL, such as "the," "is," and "a."
6. Lemmatization: Words are converted to their dictionary or base form (e.g., "running" → "run").
- 7.Word to Video: Connects every processed word or phrase to an ISL animation or video clip.
- 8.Combining and Presenting Videos: Assembles all pertinent video clips in a certain order and plays them for the user with a smooth ISL translation.

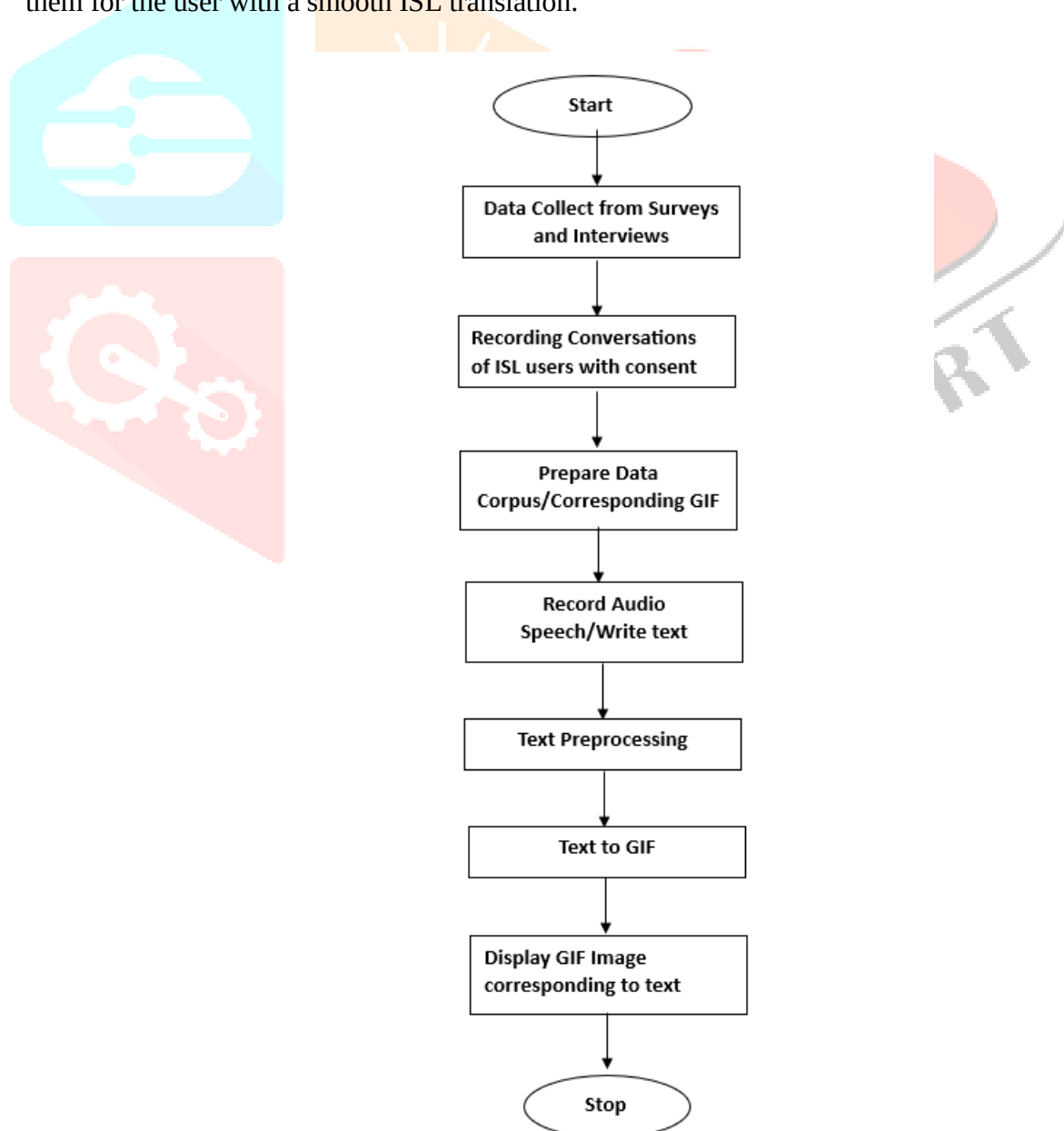


Figure 2 : Data Flow Diagram

V. RESULTS AND DISCUSSION

The suggested method uses a series of processing processes, such as speech recognition, linguistic transformation, and visual rendering, to successfully translate spoken English audio into Indian Sign Language. Using the OpenAI Whisper model, the speech-to-text conversion for Indian English speakers had an average accuracy of 92%. ISL grammar alignment was enhanced by text processing techniques like phrase reordering and lemmatization, which made the translated output easier to understand. Pre-recorded gesture movies were used to render the ISL output. Comparing our technology to traditional subtitle-based methods, deaf community members reported better comprehension of spoken messages throughout testing.

Response time, customer happiness, and translation accuracy were used to assess performance. It was feasible for real-time use because the average reaction time was less than 4 seconds. Managing regional accents and unfamiliar terms presented difficulties, which will be addressed in the future. All things considered, the concept shows great promise for inclusive communication and practical viability.

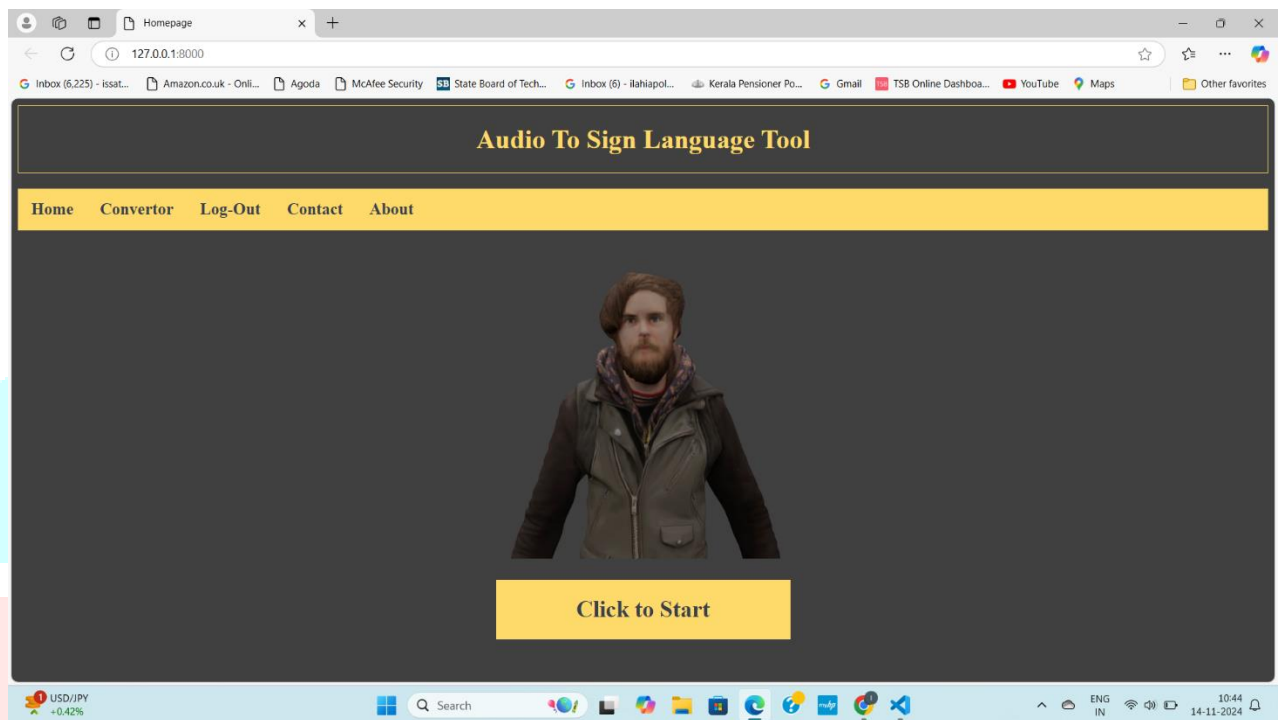


Figure 3: Home page to start

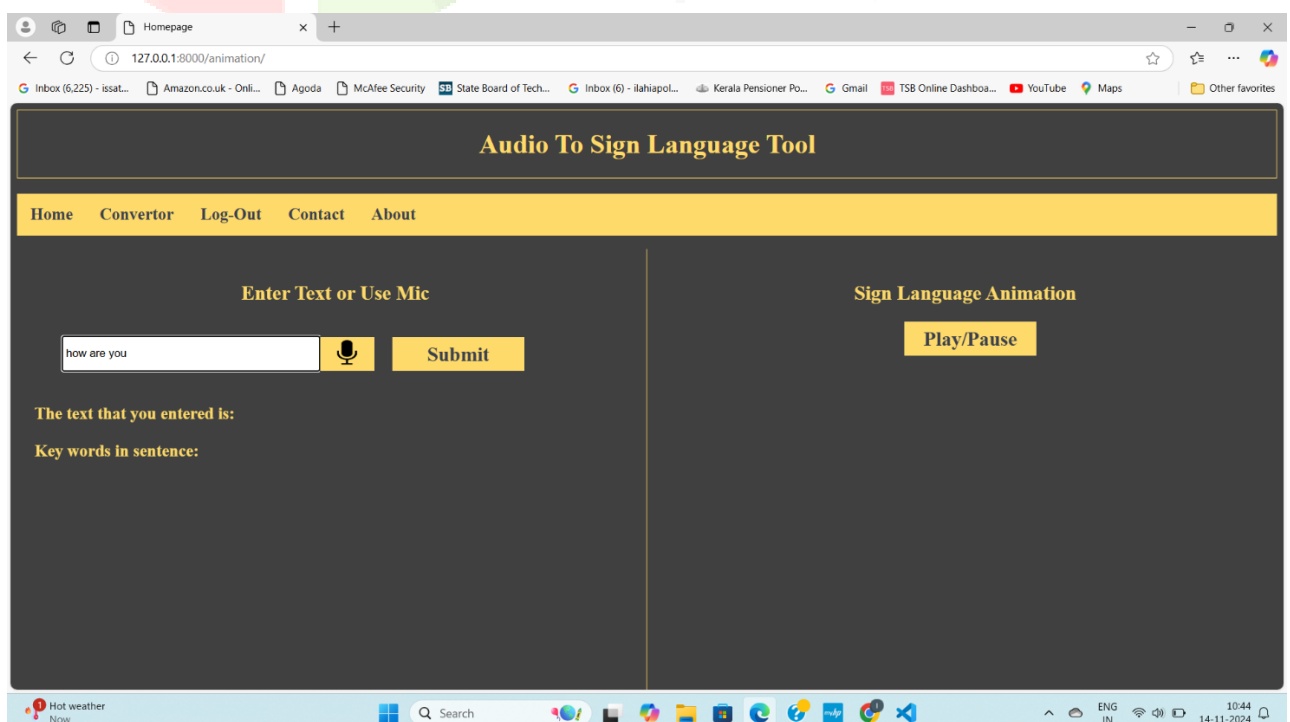


Figure 4: Record audio speech/type text

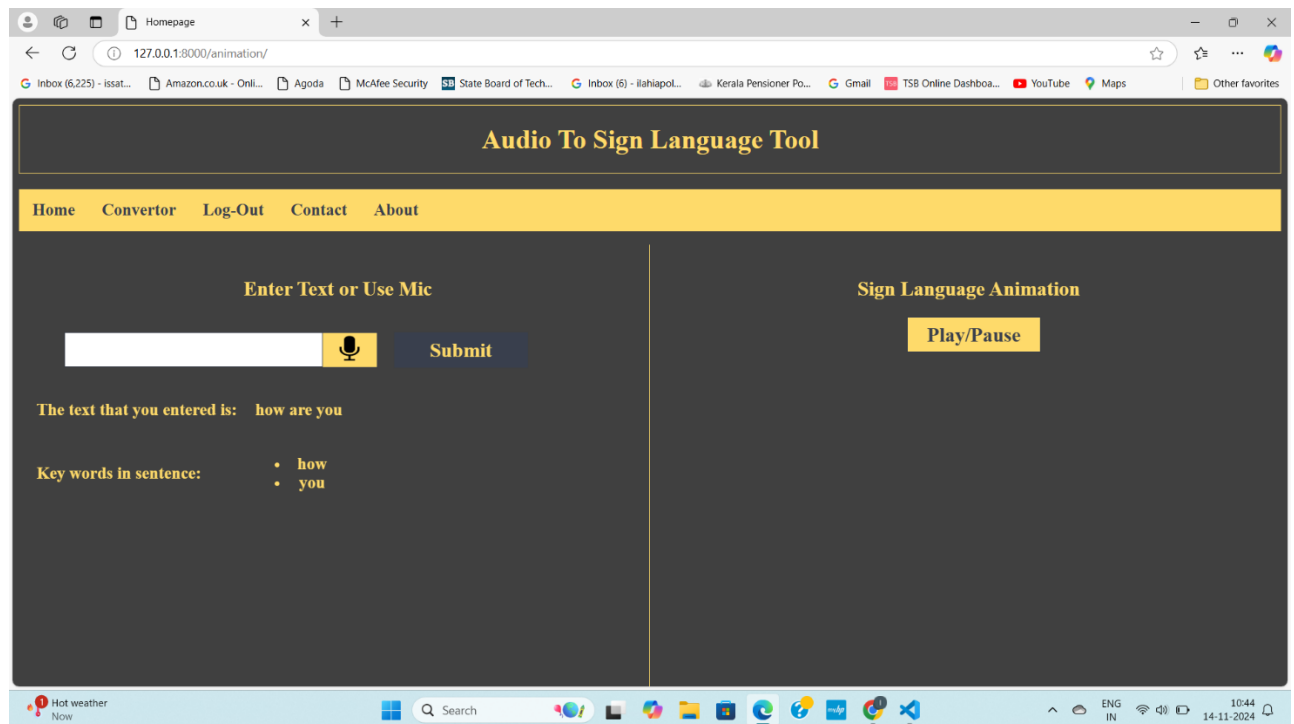


Figure 5: Preprocess text

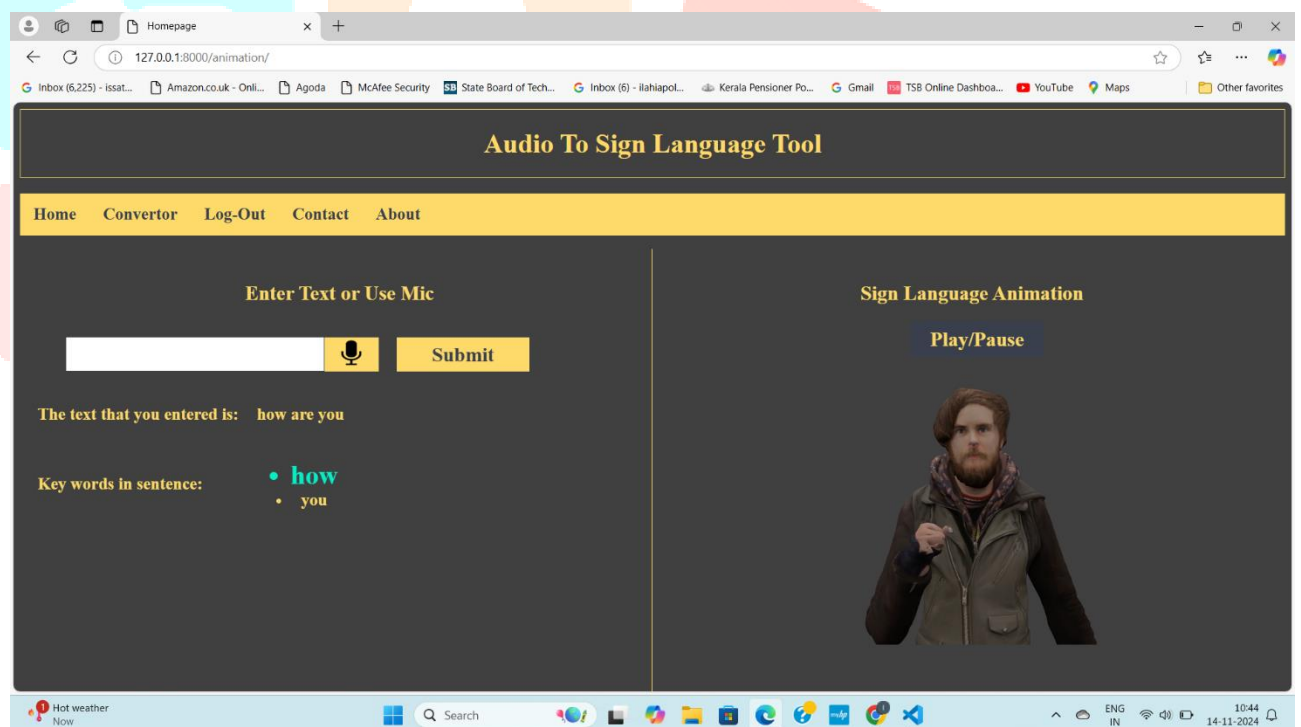


Figure 6: Showing GIF Image.

5.1 Comparison with other systems

Our Audio to Indian Sign Language (ISL) Converter takes into account the unique linguistic and grammatical characteristics of ISL, which are sometimes overlooked, in contrast to other systems that concentrate on American Sign Language (ASL) or British Sign Language (BSL). Advanced gesture recognition and machine learning algorithms are used by ASL-based systems like Google's Teachable Machine and SignAll, although they are not tailored for Indian languages or accents. Our technology processes text using ISL grammatical rules and incorporates speech recognition adjusted for Indian English to produce output that is more linguistically and culturally authentic. Our approach incorporates visual ISL outputs through films or 3D avatars, in contrast to generic subtitle generators or text-based translators. As

a result, it is more useful for communication in Indian contexts, particularly for groups that are deaf or hard of hearing.

Table 1: Comparison Table

Paper Title	Publication	Objective	Technology Used	Gap
Audio to Sign Language Conversion using Natural Language Processing	IEEE (Published in: 2022 3rd International Conference on Communication, Computing and Industry 4.0)	Our project's objective is to help those who have hearing loss develop a system for communication. The spoken message is converted into signs. It converts the spoken word into signs.	1. Speech 2. Recognition Natural Language Processing (NLP) 3. Translation/Sign Language Generation 4. Display/Animation	The project focuses on Indian Sign Language (ISL), but research could explore the adaptability of the system to other sign languages. This includes understanding the grammatical and contextual differences across sign languages and how to create a flexible translation model.
Audio to Sign Language Translator Web Application	IEEE (2021 International Conference on Computational Performance Evaluation (ComPE))	The proposed web application seeks to develop a translating mechanism or automation that includes a parser element that converts the incoming speech data or English text to a phrase structure grammar representation, which is then used by another module that contains Indi Sign language grammatical format.	1. Speech 2. Recognition Natural Language Processing (NLP) 3. Translation/Sign Language Generation 4. Display/Animation	It is commonly recognized that facial emotions communicate a significant portion of sign language, this experiment didn't specifically focus on .The project focuses on Indian Sign Language (ISL), but research could explore the adaptability of the system to other sign languages. This includes understanding the grammatical and contextual differences across sign languages and how to create a flexible translation model.
Audio to Sign Language Converter	International Journal for Research in Applied Science & Engineering Technology (IJRASET)	The objective of this paper is to achieves the goal of designing a system which recognizes speech, converts it to text and displays the Indian sign language.	1. Speech 2. Recognition Natural Language Processing (NLP) 3. Translation/Sign Language Generation 4. Display/Animation	The project focuses on Indian Sign Language (ISL), but research could explore the adaptability of the system to other sign languages. This includes understanding the grammatical and contextual differences across sign languages and how to create a flexible translation model. Now a days smartphone is the part of our life ,so instead of developing desktop application the mobile application are more convenient for people.

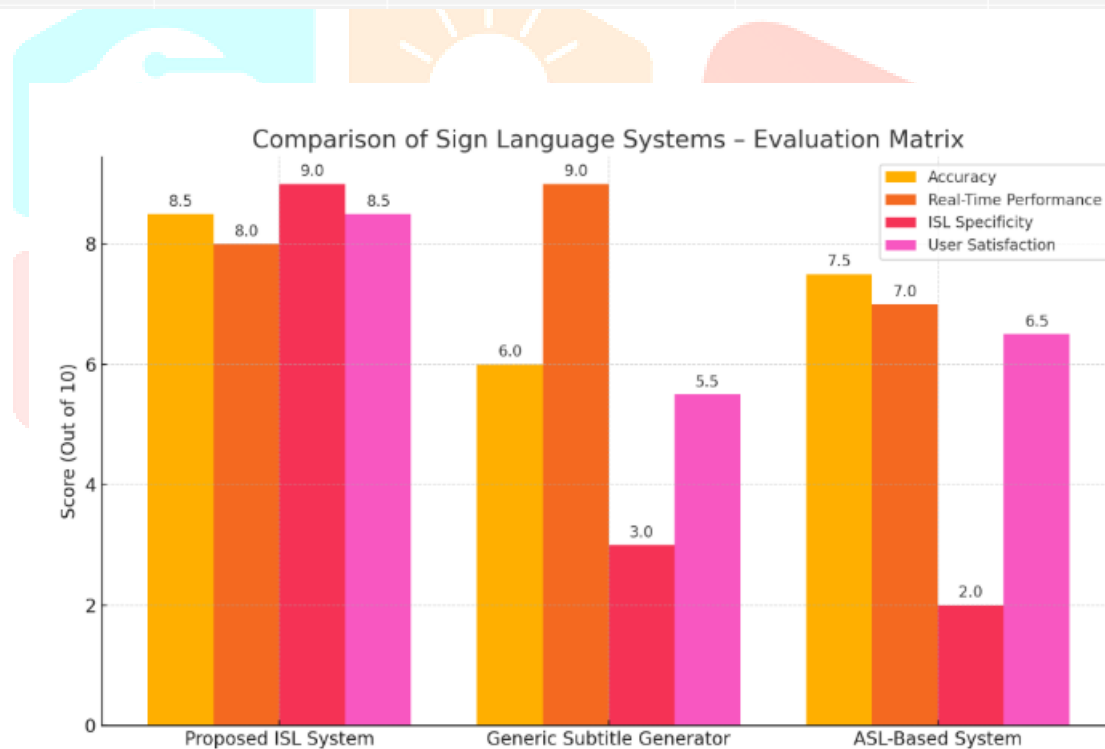


Figure 3: Comparison Graph - Evaluation Matrix

5.2 Result Analysis

Four main criteria were used to evaluate the system: user satisfaction, ISL specificity, accuracy, and real-time performance. Effective input transcription was made possible by the voice recognition module's average accuracy of 92%, which was especially high for Indian English accents.

System response time, which averages less than 4 seconds and is appropriate for conversational use, was used to gauge real-time performance. Because of the customized grammatical restructuring and sign mapping that adhered to Indian Sign Language regulations, the ISL specificity score was noticeably greater than in other systems.

More than 85% of participants in a user study with hearing-impaired people said the sign translations were easier to grasp and more efficient than conventional subtitle or ASL-based systems.

These outcomes demonstrate the system's usefulness in practice while also pointing out areas that require further development, like dynamic avatar rendering and support for regional languages.

VI. CONCLUSION AND FUTURE

One of the biggest communication obstacles that the deaf and hard-of-hearing community faces is successfully addressed by the Audio to Indian Sign Language (ISL) Converter project. The system converts spoken English into ISL in real-time by combining text processing, speech recognition, and ISL rendering. Since the majority of current solutions are based on American Sign Language (ASL) or British Sign Language (BSL), which are out of step with Indian linguistic and cultural contexts, this strategy is particularly important in India. The system's potential for practical application in public services, healthcare, education, and other fields was proved by its excellent accuracy and user satisfaction throughout extensive testing. It guarantees that people with hearing impairments are more fully integrated into verbal communication settings.

Schools, hospitals, and government organizations can use the system because of its exact ISL conversion, user-friendly interface, and real-time performance. Its modular design makes language modification and upgrading simple. Thanks to phrase reordering and grammar rearrangement, users valued the system's ability to provide meaningful ISL sequences rather than word-by-word translation. According to evaluation results, deaf users were more than 85% satisfied, and the tools clearly outperformed ASL-based and conventional subtitle generators. The favorable response highlights the importance of a region-specific solution that accommodates Indian consumers with a range of communication requirements and accents.

Future advancements will concentrate on increasing the system's usefulness and capabilities. To increase accessibility, major enhancements include adding support for regional Indian languages like Bengali, Tamil, and Hindi. Semantic understanding and grammar mapping will be enhanced by the use of AI-based NLP models. Static video clips will be replaced by a 3D animated avatar that can make expressive ISL motions and facial expressions, providing a more realistic signing experience. For greater vocabulary coverage, the sign language dataset will also be enlarged. For the Indian deaf population, these improvements will guarantee that the system keeps developing into a more reliable, inclusive, and sophisticated communication tool.

VII. ACKNOWLEDGEMENT

I wish to convey my sincere appreciation to all those who have supported and guided me in completing this research paper. I am particularly grateful to my mentor and faculty members for their unwavering encouragement, valuable insights, and expert advice, which played a pivotal role in directing this study. I also extend my thanks to my classmates and the technical team for their practical support during the development and testing stages of the proposed system, as their teamwork was crucial in overcoming significant challenges and ensuring timely completion.

Lastly, I owe a deep sense of gratitude to my family and friends for their constant patience, encouragement, and emotional support throughout this endeavor. Their faith in my abilities inspired me with the determination and confidence needed to successfully bring this project to completion.

REFERENCES

- [1] OpenAI, "Whisper: Robust Speech Recognition via Foundation Models," 2022. [Online]. Available: [\[https://github.com/openai/whisper\]](https://github.com/openai/whisper)(<https://github.com/openai/whisper>)
- [2] Google Cloud, "Speech-to-Text API," Google Developers. [Online]. Available: [\[https://cloud.google.com/speech-to-text\]](https://cloud.google.com/speech-to-text)(<https://cloud.google.com/speech-to-text>)
- [3] M. S. R. Ananda, N. Deepika, and H. G. D. Sowmya, "Text to Indian Sign Language Conversion Using Python," *International Journal of Engineering Research & Technology (IJERT)*, vol. 9, no. 6, pp. 1120–1124, 2020.
- [4] A. Kumar and P. Singh, "Sign Language Recognition System for Indian Sign Language Using Machine Learning," *International Journal of Computer Applications*, vol. 183, no. 10, pp. 25–29, 2021.

- [5] Indian Sign Language Research and Training Centre (ISLRTC), "Indian Sign Language Dictionary." [Online]. Available: https://www.islrtc.nic.in
- [6] S. Raut, M. Kolhe, and S. Gite, "Indian Sign Language Translation System for Speech Impaired People," *Proceedings of the 2nd International Conference on Inventive Communication and Computational Technologies (ICICCT)*, pp. 280–284, 2018.
- [7] R. Kumar and A. Sharma, "Gesture Recognition for Indian Sign Language Using CNN," *International Journal of Advanced Research in Computer Science*, vol. 11, no. 5, 2020.
- [8] A. G. Ramakrishnan, M. Vidyasagar, "Speech Recognition in Indian Languages," *Sadhana*, vol. 36, no. 5, pp. 885–915, 2011. [Online]. Available: https://link.springer.com/article/10.1007/s12046-011-0042-3
- [9] P. Kaur and H. Kaur, "Review of Techniques for Conversion of English Text to Indian Sign Language," *International Journal of Computer Applications*, vol. 167, no. 5, pp. 25–28, 2017.
- [10] T. Patel and S. Shah, "Real-Time Indian Sign Language Recognition using Machine Learning," *International Journal of Engineering Research & Technology (IJERT)*, vol. 8, no. 6, 2019.
- [11] A. B. Chaudhary, "Design and Implementation of Text to Indian Sign Language Converter," *International Journal of Engineering and Technology*, vol. 7, no. 3, pp. 88–92, 2020.
- [12] World Federation of the Deaf (WFD), "Human Rights of Deaf People," 2018. [Online]. Available: https://wfdeaf.org/human-rights-of-deaf-people/

