SPEECH TO SIGN LANGUAGE TRANSLATION

Under the Esteemed Guidance Of

Dr. Ch. Meher Babu, M.Tech, Ph.D

Associate Professor

Submitted by

IV CSE(AI&ML) (Batch-13)

Ms. K.MAHALAKSHMI (20KT1A4223)

Ms. SARAH CRYSOLITE (20KT1A4249)

M. K B V SUBRAMAYESWARA REDDY (21KT5A4201)

ABSTRACT

The development of a Speech to Sign Language Translator (SSLT) aims to bridge the communication gap between hearing and deaf communities by automatically translating spoken language into sign language. The SSLT can also uses in the case of news ai reader to represent the news in sign language for the deaf people. The existing approach can predict a sign image of a language to the given input audio. The limitation in this approach is that we cannot get the sign for sentences, questions or wishes like “How are you?”, “Your looking good” or “Nice to meet you”. we cannot get all these in an image or 2D/3D avatars. Some of developed models has not given appropriate results for the given input So, to overcome this problem we have proposed a model which can take audio and generate sign video or gifs. It seeks to provide real-time translation of spoken language into sign language gestures, thereby bridging communication gaps between the hearing impaired and the hearing community. A speech recognition module, a linguistic rule-based translation module, a sign language dictionary, and a Tkinter-based model are all essential components. The system aims to improve accessibility and inclusivity by offering an intuitive mode of communication that promotes social integration and comprehension.

INTRODUCTION

Speech to Sign Language Translation (SSLT) is a technology that helps people who use spoken language communicate with those who use sign language. For people who are deaf or hard of hearing, sign language is important for talking and understanding others. But when someone speaks to them, it can be hard to understand without an interpreter. SSLT solves this by turning spoken words into sign language in real-time. It uses advanced technology like machine learning and computer vision to understand spoken words and translate them into sign language gestures accurately. While SSLT has made progress, there are still challenges, like making sure the translation is accurate and fast, and handling the complexity of sign language.

LITARATURE SURVEY

In [1], Jashwanth Peguda *et. al.*, built SSLT using Mozilla's common voice and MultiIndicMT.They used speech recognition technologies and wavelet-based transformations, such as MFCC with GMM, to turn the speech or voice into text. Later, the text should be pre-processed using NLP techniques, namely LSTM encoders and decoders. The direct translation method is used to translate the text character by character to the matching signs of the sign language.

Pankaj Sonawane *et. al.*, [2] built the SSLT utilizing Microsoft Kinect for Xbox 360 and recorded motion capture data for all basic Sign Language motions. They translate the speech into text and save the output as a composite string. The string is supplied to the text parser, which may separate the words into alphabets. Each alphabet was processed independently, and they had constructed a user interface to transform the speech to signs.

In [3], Rahul Solleti *et.at.*, implemented SSLT with AR glasses, which incorporates voice transcription using OpenAI tools and NLP approaches. voice translation utilizes recurrent neural networks (RNN) that may be trained to translate text to a sign language animation. Sequence-to-sequence models with encoder-decoder architectures are particularly well adapted to the text-to-sign translation job. The sign output can be represented using 2D/3D animation libraries or avatar motions that work with technologies such as Unity3D. Mapping the model outputs to control avatar movements creates an intuitive depiction of the sign language translation and Augmented Reality (AR). Glasses that display translated sign language avatars in an augmented reality interface can improve comprehension for deaf users.

Rudrabha Mukhopadhyay *et. al.*, and colleagues generated their own dataset in [4] that included text, audio, and video samples. The multitasking transformer network, which includes a speech encoder, posture generator/decoder, and text decoder, has been utilized by them. The input speech segments were matched to the sign pose sequences using the cross- model discriminator.

Ritika Bharti *et. al*., and her team used Google API and natural language processing (NLP) to produce an automatic SSL conversion in [5]. They have developed a three-stage methodology that uses the Google APIs to convert speech to text: On the converted text, NLP techniques like tokenization, lemmatization, and stemming should be used along with speech feature extraction, matching, and speech to text modelling. When the text and the visual sign word library match, they retrieve the videos, integrate them in accordance with Ti, and display the finished product. This project has been tested both offline and online.

In [6], Kajal Jadhav *et. al.*, developed Speech to ISL, which takes voice from a microphone and converts it into text using Google API. Then they use rule-based and statistical-based translation, with a set of 153 rules, to extract the sign sequence from the textual data. If the sign sequence and sign description match, the sign animation image is produced. The program achieves a 40 per cent delay reduction without impacting the efficiency of the interpretation process.

In [7], Rishin Tiwari *et. at.*, The system utilizes Pyaudio, SPHINX, and Google speech recognition API to convert voice input into text. Following this, the text is converted into Indian Sign Language output, which is displayed on the screen of the machine in the form of a series of

images or motioned video using various Python libraries for image processing. These combined technologies enable the system to accurately recognize and translate spoken language into text format, which is then further processed to display the corresponding Indian Sign Language on the screen, ensuring effective communication for the hearing and speech disabled.

An advancement in speech to sign language translation using a 3D avatar animator, or ES2ISL, was proposed by Bhavinkumar Devendrabhai Patel et al., [8]. In this system, real- time microphone data was obtained and then transformed into text using a Google Cloud speech recognizer. Following conversion, the text should be provided to the input parser, and data preprocessing techniques such as lemmatization and stemming were used. After using HamNoSys, SigML is used to generate a sign avatar. Using the SigML model, the avatar displays sign gestures based on the text.

The SSLT was implemented by Ananya Verma *et. al*., [9] using Direct translation systems, ViSiCAST translator, HamNoSys, and SiGML models. HamNoSys provides standardized notation for sign languages, enabling accurate sign gestures. They are also used for text-to- sign gestures, matching text or speech with appropriate sign gestures and producing an output with an intuitive user interface.

Liju Thomas et al. developed voice to American sign language in [10], utilizing NLTK, Google voice Recognition API, and Extensible Embeddable Language (EEL) to translate text into associated sign motions. the process of creating a series of pictures in sign language based on the spoken input, making it possible to communicate with hearing-impaired people more effectively and simply.

S.no	Author	Title	Dataset	Algorithm	Merits	Demerits	Accuracy
1.	Jashwanth Peguda	Speech to sign language translation for Indian Languages.	Data was taken from common voice by Mozilla and multiindicMT	Wavelet-based Transform-ation , LSTM's	Model supports up to six Indian Regional languages which are Telugu, Hindi, Tamil, Marathi, Malayalam, and Kannada.	The model generates the sign of hand gestures.	91
2.	Pankaj Sonawane	Speech to sign language (ISL) translation	Sign Data was captured from Microsoft Kinect	Speech-to-text API and Text tokenizers used.	Text input layout	Limited gesture library.	91
3.	Rahul Solleti (2023)	A Machine Learning framework and method to translate speech to sign language translation for AR glasses	Data was taken from real-time voice.	Open Ai tool for speech transcription, RNN	It can make suitable text inputs to sign gestures sequentially.	This experiment produces avatars in the absence of face expressions	--
4.	Rudrabha Mukhopadhyay	Towards Automatic Speech to Sign Language Generation	They have created own dataset which have videos, text, speech.	Multi-tasking transfor--mers Network, Cross-Model Discriminator	Model learns to generate continuous sign pose sequences in an end-to-end manner.	The Multi-tasking transfor--mers makes complexity in training of model.	53.33
5.	Ritika Bharti	Automatic Speech to Sign language conversion using google API and NLP	American Sign Language Video Dictionary database from Handspeak	Open Ai tool for speech transcription, NLP techniques	Can be used online and offline.	No alternative for output if the given input isn't in the database	Offline-74 Online-90

6.	Kajal Jadhav	Speech to ISL Translation	Real-time data through microphone	Rule based and statistical-based Translation.	Reduction in delay. System records partial results every 100 millisecond s.	The project may face limitations in accurately capturing the nuances of sign language, including facial expressions and body movements.	--
7.	Rishin Tiwari	Audio to sign converter	ISL dataset used which have English words and Sign gestures.	Google API, NLP technique's	Google Speech-to-Text, with its powerful neural network models, enables the recognition of over 120 languages	This experiment produces hand gestures in the absence of face expressions.	--
8.	Bhavinku mar Devendra bhai Patel	ES2ISL: An Advancement in Speech to Sign Language Translation using 3D Avatar Animator	ISL dataset which consists videos, 1113 English words	HamNoSys, SiGML	It is capable of generating output within 1 second per conversion , saving valuable processing time.	The system may face challenges in handling complex sentences or specialized vocabulary that are not included in the existing database.	72
9.	Ananya Verma	Converting Voice Signal to Visual Indian Sign Language (ISL)	Live speech form microphone	HamNoSys, SiGML, Direct Translation Systems	Two way communic ation.	Expressions change to signify the negation and interrogatio n in the sentence.	--

10.	Liju Thomas	Audibly: Speech to American Sign Language converter	Live speech form microphone	Extensible Embeddable Language (EEL), Google Speech Recognition API, NLTK,	NLP methods to find root words of phrases and remove filler	Limited Vocabulary	95

Design and Proposed Methodology:

The SSLT can also be used as an AI reader to depict news in sign language for deaf individuals. The Speech to Sign Language Translation (SSLT) model takes speech as input and converts it into text using Google Speech-to-text conversion API tools such as Sphinx and the Google Cloud API. NLP algorithms are applied to the transformed text, and video or gif data of sign language is stored and we can use the Tkinter Framework. If the converted text and the video in the database match, it retrieves the data and generates the video or gifs of sign gestures for the necessary input.

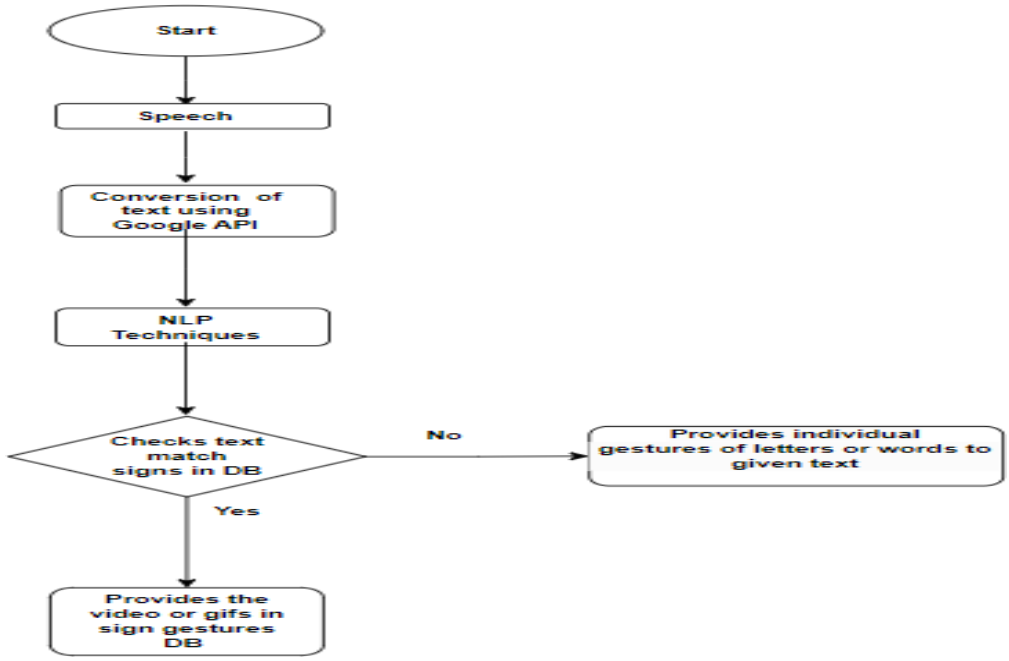

Fig 1: Design of Proposed Methodology

RESULTS AND DISCUSSION

The SSLT can convert real-time speech's from the microphone into text and, using the text, get videos of sign gestures from the ISL database. We used flask to implement the website. By clicking the "translate" button, we can record our voice and the sign gestures video will be produced based on the speech. The accuracy is 80 percent.

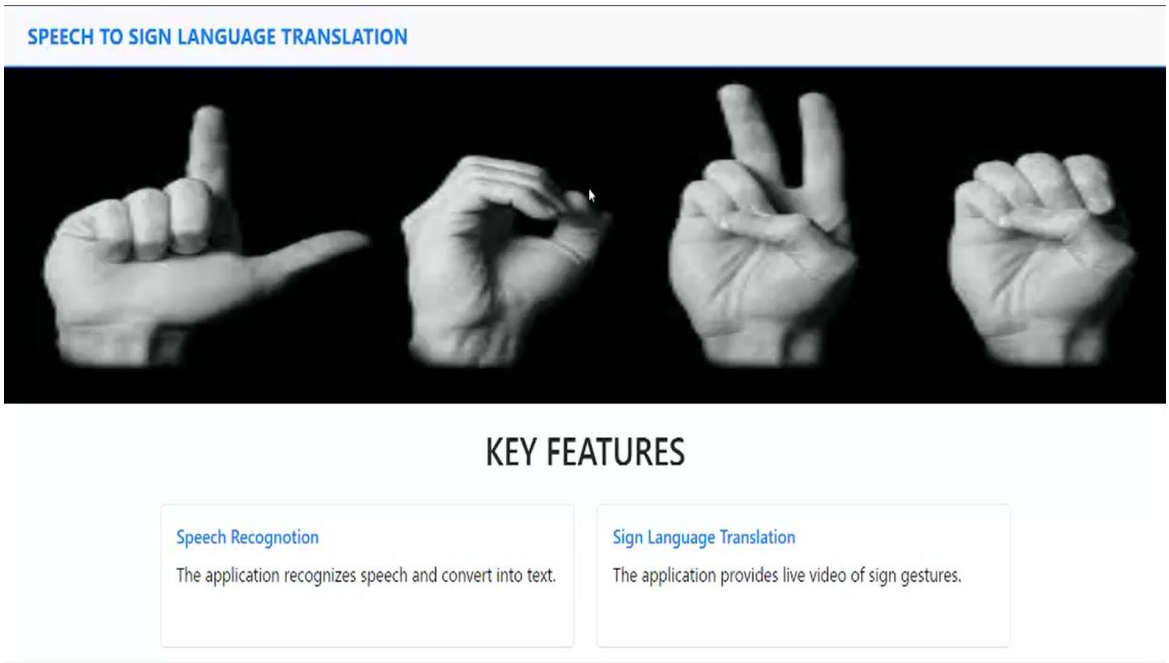

Fig 2: Home page1

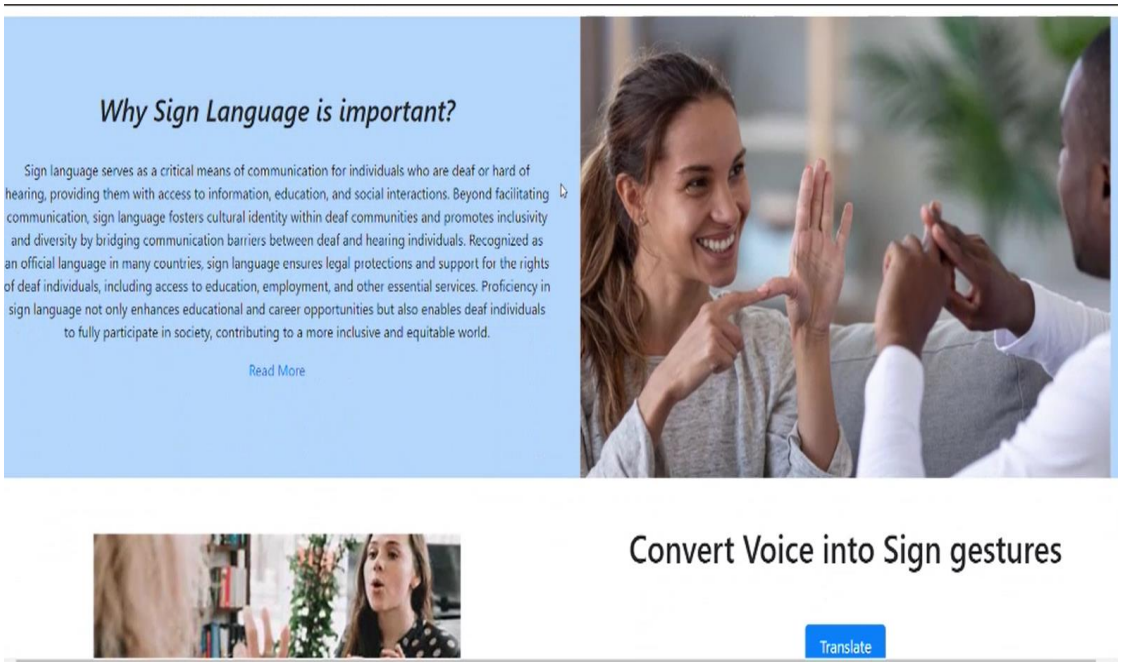

Fig 3: Home page2

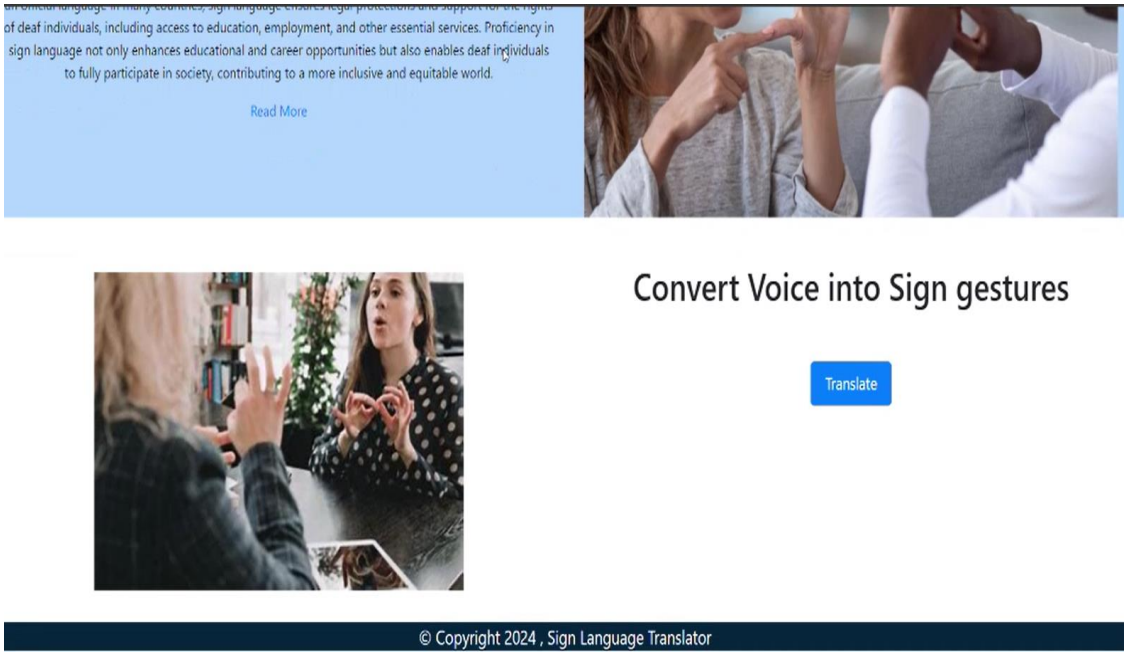

Fig 4: Convert voice by clicking "translate" button

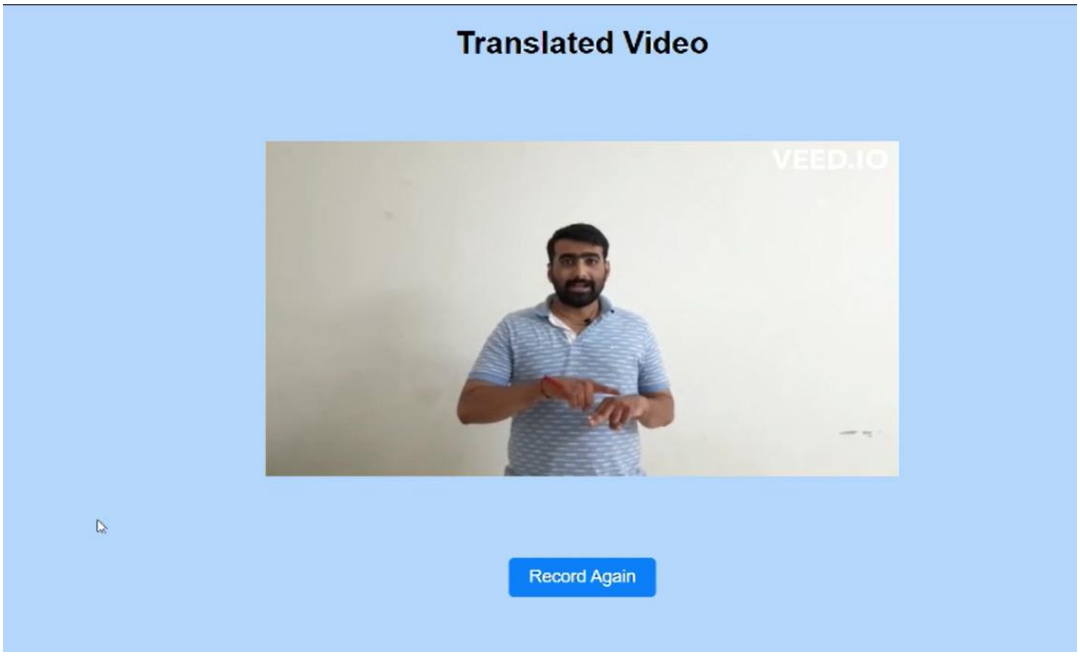

Fig 5: Sign gesture video

Conclusion

People who are deaf or dumb use sign language to express what they wish to say. For those who are deaf or dumb, it serves as a fundamental form of communication; without it, interactions with regular people may be challenging. In order to enable interaction with deaf people, the main goal of this study is to design an automated software system that translates voice to sign language gestures.

References

[1] Jashwanth Peguda, V Sai Sriharsha Santosh, Y Vijayalata, Ashlin Deepa R N, Vaddi Mounish, "Speech to Sign Language Translation for Indian Languages," 2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS) ,doi:10.1109/ICACCS54159.2022.9784996

[2] Pankaj Sonawane,Karan Shah, Parth Patel, Shikhar Shah, Jay Shah ,"Speech To Indian Sign Language (ISL) Translation System," 2021International Conference on Computing, Communication, and Intelligent Systems(ICCCIS), doi: 10.1109/ICCIS51004.2021.9397097

[3] Rahul Solleti," A Machine Learning framework and method to translate speech to real-time sign language for AR glasses", DOI: https://www.doi.org/10.58257/IJPREMS32125

[4] Parul Kapoor, Rudrabha Mukhopadhyay, Sindhu B Hegde, Vinay Namboodiri1, C V Jawahar," Towards Automatic Speech to Sign Language Generation"DOI:https://doi.org/10.48550/arXiv.2106.12790

[5]RitikaBharti,SarthakYadav,SouravGupta,and Rajitha Bakthula," Automated Speech to Sign language Conversion using Google API and NLP",DOI:http://dx.doi.org/10.2139/ssrn.3575439

[6] Kajal Jadhav,Shubham Gangdhar,Viraj Ghanekar," SPEECH TO ISL (INDIAN SIGN LANGUAGE) TRANSLATOR ", International Research Journal of Engineering and Technology (IRJET),Volume: 08 Issue: 04 | Apr 2021.

[7] Rishin Tiwari,Saloni Birthare, Mr. Mayank Lovanshi," Audio to Sign Language Converter",DOI:https://doi.org/10.22214/ijraset.2022.47271

[8] Bhavinkumar Devendrabhai Patel, Harshit Balvantrai Patel, Manthan Ashok Khanvilkar, Nidhi Rajendrakumar Patel, Thangarajah Akilan,"ES2ISL: An Advancement in Speech to Sign Language Translation using 3D Avatar Animator",DOI:10.1109/CCECE47787.2020.9255783

[9] Ananya Verma, Jeremiah Nunes, Dr M Aruna," Converting Voice Signal to Visual Indian Sign Language (ISL) ",DOI:http://dx.doi.org/10.2139/ssrn.4391302

[10] Abhirami M.S, Achint P.V, Liju Thomas, Salma Salam,"Audibly: Speech to American Sign Language converter",DOI:10.36227/techrxiv.20154068.v1