SIGNIFY: SPEECH TO SIGN LANGUAGE TRANSLATOR

[1]Anhad Chaube, [2]Khushboo Patil, [3]Neha Sawant, [4]Tejaswi Waghrulkar
Progressive Education Society's Modern College of Engineering, Pune

Guided By:
Dr. S.D Deshpande
HOD, Department of Information Technology,Modern College of Engineering, Pune

---

Abstract: Language plays an important role in the communication of ideas, thoughts, and information to others. Hearing-impaired people can understand our thoughts using a language called as sign language. Every country or region can have a different sign language which is completely based on their native language. In our research paper, our main focus is on Indian Sign Language. The Indian sign language is widely used by speaking and hearing communities in India While communicating with others, one of the most vital factor is listening. It may happen that the other party is not able to hear or grasp what you are talking about. This situation is faced by almost every hearing-impaired person in our society. Thus, we introduced our project Signify: Speech to Indian Sign Language translation system which can reduce the gap in communication between hearing-impaired people and society. The proposed system takes audio and text as input and matches it with the videos present in the database. It then results in corresponding sign movement videos based on the grammar rules of Indian Sign Language; if not found, it then goes through the processes of tokenization and lemmatization. The technology behind the system is natural language processing which makes the system with tokenization, parsing, lemmatization, and part-of- speech tagging.
Keywords: Indian Sign Language; natural language processing; tokenization; lemmatization; parsing;Signify

---

Introduction:
Every country has their own sign language which is entirely based on their native language. It is not easy for us to communicate when we know that the other person is not able to hear and understand what we are trying to say. Even some people with sufficient hearing abilities tend to ignore or avoid communication with those who cannot hear, and for those it becomes more difficult. To have the skill to talk to those who cannot hear cannot only bridge the gap between those two persons but also help in the exchange of information and new thoughts which could encourage these people to contribute to the development of technology. Every person can contribute in making unknowns into knowns and impossible possible.

Indian Sign Language provides facility for people to create an inclusive society in which people with disabilities have equal chances for growth so that they can live productive, safe, developed and dignified lives. In India's hard-of-hearing community, Indian Sign Language (ISL) is widely used. However, ISL is not utilized to teach students with hearing disabilities in deaf schools. Teachers are not trained to use ISL in their classrooms. Indian Sign language is not included in any of the teaching and learning materials.

The parents of hearing disabled children are often unaware of sign language's value in bridging communication gaps. ISL interpreters are very high demand at institutes and other locations where hard- of-hearing and hearing individuals communicate, yet In India there are only about 300 licensed interpreters.

Objectives of ISL:

- To train people to use Indian Sign Language (ISL).

To educate hearing disabled people and pursue their dreams.

- To encourage students with hearing disabilities in primary, intermediate, and higher education to use Indian Sign Language as a form for communication.
- To educate and train various groups, such as government officials, teachers, professionals, community leaders, and the public, on Indian Sign Language and how to use it.
- To promote Indian Sign Language in collaboration with hard-of-hearing groups and other institutions working on disabilities.

Problem Statement:

The main purpose of project is to take user input and convert it to sign language. Using Natural Language Processing (NLP) are trying to classify the text/speech into small part. Then searching those words from database. At the end displaying the appropriate sign video to the user. In this project we have:

1. Speech recognition & converting into text.
2. Statements and converting them into sign language.
3. If no words are found in database convert it into letters one by one.

Sign language is a language that makes use of guide communique methods including facial expressions, hand gestures and bodily actions to bring facts. This task makes use of videos for particular phrases combined to translate the text language into sign language. Speech impaired people use hand signs and gestures to speak. Normal human beings face problem in understanding their language. For this reason there is a need of a gadget which acknowledges the different signs, gestures, and conveys the statistics to the deaf humans from regular human beings. It bridges the gap among physically challenged people and everyday humans. Our approach offers the bring about minimum time span with most precision and accuracy in assessment to other current processes.

Literature Survey:

A two-way conversation device is usually recommended in the paper[1] (Tewari.Y, June 2021) but the authors are most effective capable of convert 26 alphabets and three characters with an accuracy charge of 99.Seventy eight% the usage of CNN fashions. The authors only advise that destiny paintings must be conducted inside the field of herbal language processing to convert speech into sign language.

Within the paper [2] (Kunjomon J, November 2019), the authors advocate a gadget that converts sign language into English and Malayalam. The authors of the paper advise the use of an Arduino Uno, which makes use of a couple of gloves to understand the signs and symptoms and translates the signs from ISL to the preferred language. The device is beneficial, because it acknowledges -hand and motion signs and symptoms.

The Indian sign Language interpreter provided within the paper [3] (Gangadia D, 2020) makes use of hybrid CNN fashions to stumble on multiple signal gestures after which is going directly to expect the sentence that the consumer is making an attempt to gesture by way of the usage of herbal language processing techniques. The system is able to gain 80–95% accuracy underneath various conditions.

In some other study, the HSR model is used by the authors in converting ISL signs into text. The HSR version gives a bonus over RGB-based totally models, however this system has an accuracy starting from 30 to a hundred% relying upon the illumination, hand position, finger role [4] (Shangeetha R.Okay, December 2014) etc.

The authors of [5] (Sawant S.N, might also 2014) paper propose a gadget that recognizes 26 ASL signs and symptoms and converts them into English textual content. They use main factor evaluation to come across the signs in MATLAB. The ASL to signal language synthesis tool makes use of VRML avatars and plays them the use of a BAP player. The major trouble with the machine is that many complicated actions are not viable using the modern VRML avatars. As an example, touching the hand to any a part of the frame is not feasible in the modern-day device [6] (Tzovaras D, 2005).

In another have a look at noted in [7] (Soanre B, 2021), one video-based totally signal language translation device converts signs and symptoms from ISL, BSL, and ASL with an standard accuracy of ninety two.4%. The software utilizes CNN and RNN for the actual-time reputation of dynamic signs. The gadget then converts the signs into textual content after which uses text–speech API to offer an audio output to the consumer. The authors of any other paper first use the Microsoft Kinect 360 digital camera to seize the motion of the ISL signs and symptoms. A harmony engine is used to show the Blender 3-d animation created with the aid of the authors. Despite the fact that the gadget can correctly convert phrases into sign language, it isn't capable of convert terms/more than one words into ISL.

Existing System:

Despite the fact that sign language is used internationally to bridge the space of conversation for listening to or speech impaired which depend mostly on sign language for each day communique, there are not efficient fashions that convert text to Indian sign language. There's a loss of proper and effective audio visible assist for oral communication. While tremendous progress has already been made in computer reputation of signal languages of different countries but a totally restrained work has been finished in ISL computerization. Work accomplished up to now in this subject has been lots greater focused on American sign language (ASL) or British signal language, however for Indian sign language, structures which have been evolved are only a few.

Proposed System:

The proposed system presented on this paper is a actual-time audio to Indian sign Language conversion system so one can help hearing-impaired people to speak easily with other hearing people. The system comprises specifically six components:

1. Audio-to-textual content conversion if the enter is audio.
2. The tokenization of English text into words.
3. Parsing the English text into word shape trees.
4. The reordering of sentences based totally on Indian signal Language grammar policies.
5. Using lemmatization at the side of component-of-speech tagging so that synonyms of words or the basis form of a phrase may be used if the precise phrase is not gift within the database.
6. Indian sign Language video output.

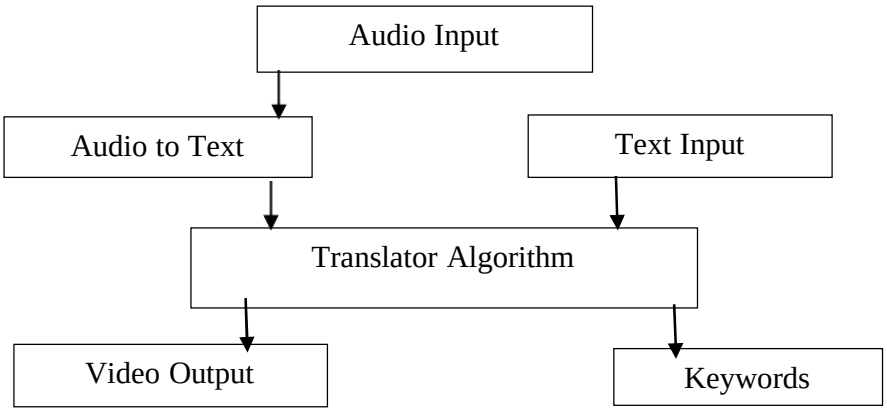
The overall efficiency of the gadget is advanced, as it splits the phrase into letters. If the video for the corresponding phrase isn't always present inside the database then it indicates output letter through letter so that it does not skip any word. Every other unique characteristic of the system is that it is able to apprehend phrases in the sentence and display sign language movies similar to the phrase if present in a dataset as opposed to word by way of word. Hence, it will increase the horizon and literacy of the device.

System Architecture:

Figure 1 shows the system architecture. The user has an option to enter the input either by text or by audio. The input is processed through the natural language model designed by the authors and keywords are given as the output. If the text within the keywords contains phrases or a combination of multiple words for which the sign language video is present in the database, then those videos are shown for such cases otherwise, the keywords are tokenized further into words or letters.

About Dataset

We had used the https://islrtc.nic.in/ to download the video clips of each and every word of Indian sign language. The database is sourced from Indian Sign Language Dictionary 3rd edition developed by the Indian Sign Language Research and Training Centre(ISLRTC), Department of Empowerment of persons with disabilities, Ministry of social justice and empowerment, Government of India. It contains dictionary of videos from numerous fields like agriculture, academics, accessories, vocabulary etc. We will like to maintain an unfiltered input that covers a wide range of words.



Methodology:

NLTK is the heart of the audio to Indian sign Language conversion system, as it is the most powerful open-source NLP library which is used to assist with human language records. Text processing is achieved using NLTK, which involves various steps, consisting of tokenization, the elimination of stop words, lemmatization, parse tree technology, part-of speech (POS) tagging, etc.

Step 1: Tokenization

Tokenization is the system of splitting text into a listing of phrases also referred to as tokens. NLTK has a module named tokenize () that is further divided into two kinds, i.e word tokenize to break up a sentence into a series of phrases and sentence tokenize to cut up a paragraph right into a list of sentences.

Step 2: Stop word removal

Stop words are a listing of very common but much less informative words that can be ignored. As an instance—her, me, has, itself, he, so, too, they, them, and so forth. On account that they're not so essential inside the sentence, they may be removed from the sentence. This can improve the overall performance ofthe system.

Step 3: Parsing

Parsing is the syntax analysis section wherein it exams whether or not the string obtained after tokenization belongs to proper grammar or not. Parsing allows to alter the textual content based at thetarget language's grammar shape.

Step 4: Lemmatization

The system of reworking the inflected kinds of a word to its root-based totally dictionary shape is known as lemmatization. This root-based dictionary shape of a phrase is referred to as a lemma. The significance of this step lies in ISL, as it requires a root phrase. To test the consequences of lemmatization, we examine its effects via sample sentences. As an example "He became gambling and eating on the

identical time". The outcomes in table show that lemmatization on my own isn't always enough to provide accurate root phrases, as it does no longer take into consideration the context of the sentence. To conquer this trouble, part-of-speech tagging comes into the picture.

Table1:Word-to-root-word conversion using lemmatization

Word	Lemma
He	He
Was	Was
Playing	Playing
And	And
Eating	Eating
At	At
Same	Same
Time	Time

Step 5: Part-of-Speech tagging
POS tagging refers back to the process of labeling words with exceptional constructs of English grammar, such as adverbs, adjectives, nouns, verbs, prepositions, and so on. POS is a group of a list of tuples wherein the primary part of the tuple is the word itself and the second component is a tag that identifies whether the word is an adjective, verb, noun, and so on. To check whether element-of-speech tagging can improve the consequences acquired after lemmatization, we analyze the equal pattern sentence used above in lemmatization. The effects in desk 6 show that when integrating component-of-speech tagging with lemmatization it offers a correct base form of a word, which in flip improves the accuracy of phrase to base phrase conversion.
Consequently, the mixture of part-of-speech tagging and lemmatization is used in our proposed machine to beautify the accuracy of our gadget.

Word	Lemma(after POS tagging)
He	He
Was	be
Playing	Play
And	And
Eating	Eat
At	At
Same	Same
Time	Time

Table 2: word to root word conversion the usage of POS tagging and lemmatization

Thus, the combination of part-of-speech tagging and lemmatization is used in our proposed system to enhance the accuracy of our system.

Technologies Used:

Speech recognition

The live speech is received as enter from the microphone of our system. This is finished the usage of the Python package deal PyAudio. PyAudio is a Python package that is used to file audio on a ramification of structures. The audio consequently received is transformed into text the use of Google Speech Recognizer API. It's far an API that allows to convert audio to textual content via incorporating neural network models.

Within the input layout of giving the audio file, the acquired audio is translated into text via using this Google Speech Recognizer. For lengthier audio documents, the audio is split into smaller chunks primarily based at the occurrence of silence. The chunks are then exceeded into the Google Speech Recognizer to successfully convert into text.

HTML, CSS, & JavaScript

HTML is used to provide the *basic structure* of the system, which is enhanced and modified by other technologies like CSS and JavaScript. **CSS** has been used to control *presentation, formatting, and layout*. **JavaScript** is used to control the *behavior* of different elements.

Django

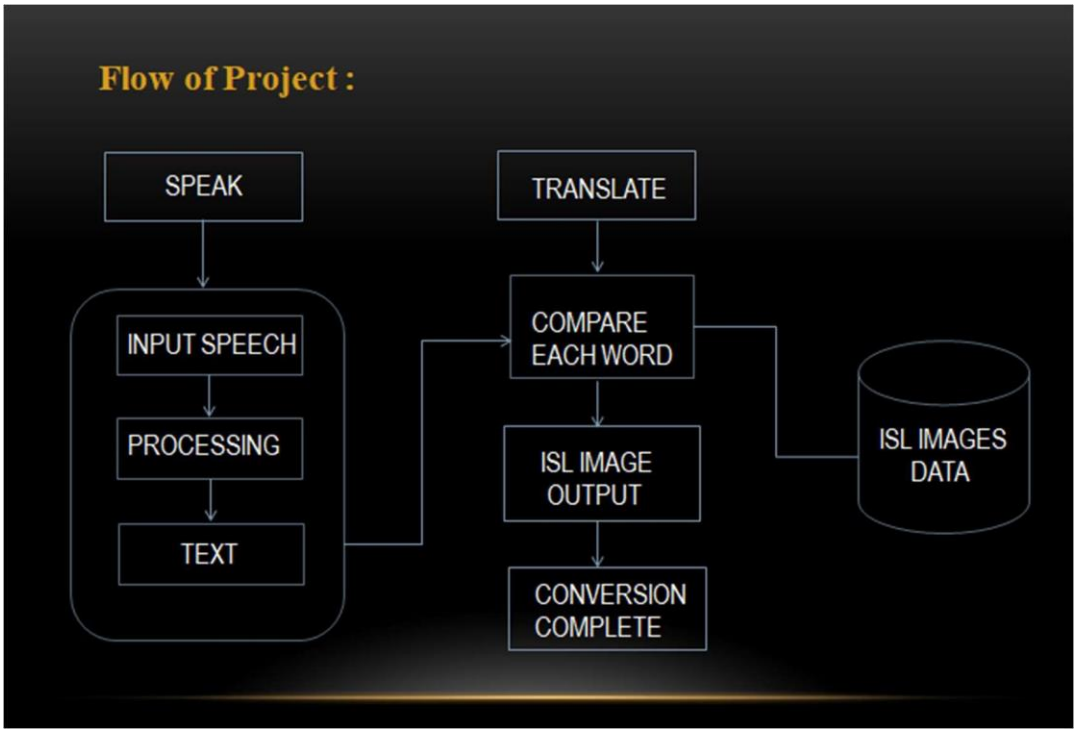
Django is a web application framework written in Python programming language. It is based on MVT (Model View Template) design pattern. The Django is very demanding due to its rapid development feature. It takes less time to build application after collecting client requirement.

By using Django, we can build web applications in very less time. Django is designed in such a manner that it handles much of configure things automatically, so we can focus on application development only. HttpResponse (source code) provides an inbound HTTP request to a Django web application with a text response. This class is most frequently used as a return object from a Django view.

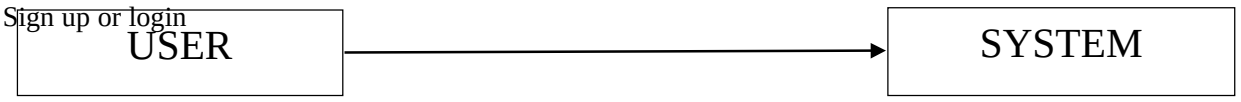
Text to sign Language

The device iterates via each phrase within the processed textual content sentence which is acquired from the previous step and searches the corresponding sign language video sequences inside the nearby system. If the word is located, the system shows the output as a video series. If the phrase is not found within the local system, then it splits the phrase into letters, in keeping with letter the sign video clips are play.

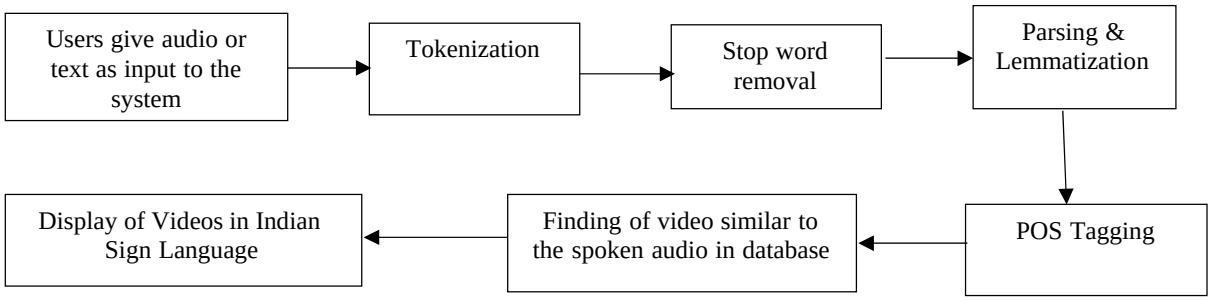
Flow Diagram:



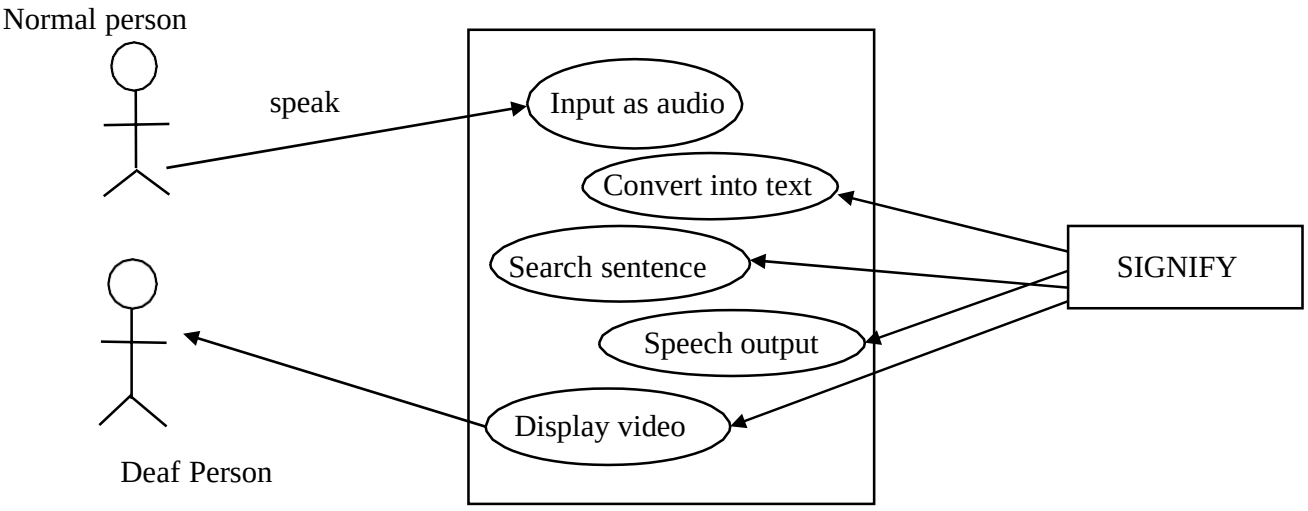
DFD L0 Diagram:



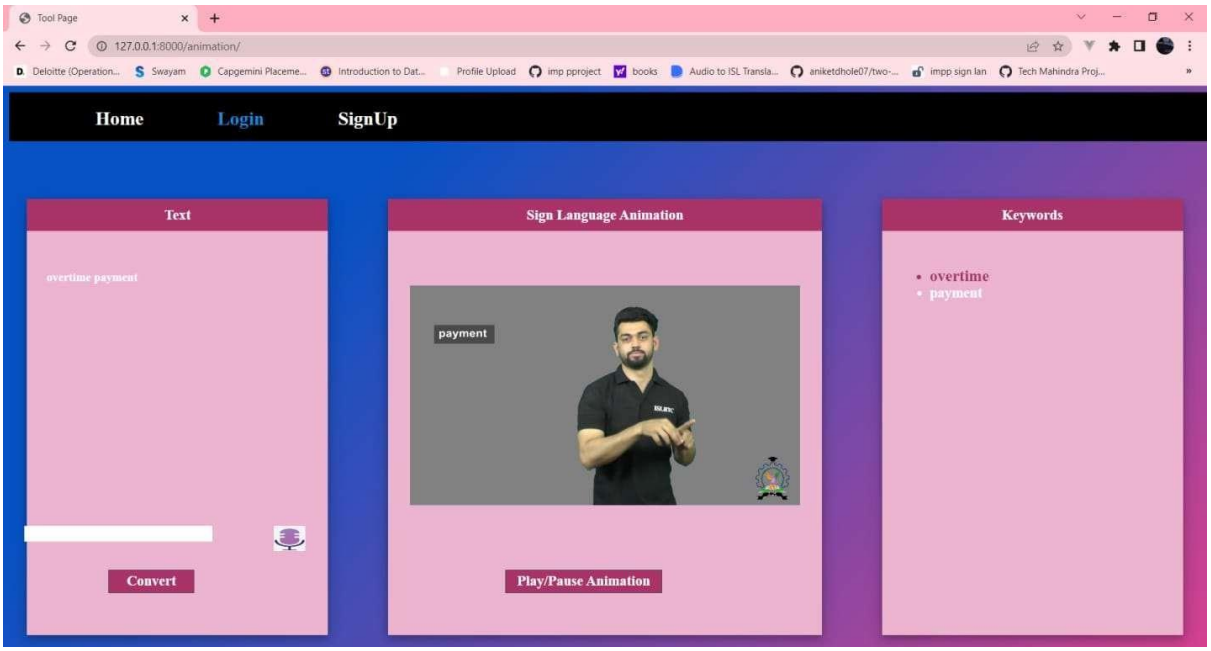
DFD L1 Diagram:



Use Case Diagram:



Results:



Conclusions

Through this paper, we have presented a user-friendly audio/text to Indian Sign Language translation system specially developed for the hearing- and speaking-impaired community of India. The main aim of the system is to bring a feeling of inclusion among the hearing-impaired community in society. The system does not only help the person who is suffering from a disability but would also be beneficial for the hearing people who want to understand the sign language of a hearing-impaired person so that they can communicate with them in their language. The core of the system is based on natural language processing and Indian Sign Language grammar rules. The integration of this system in areas such as hospitals, buses, railway stations, post offices, and even in video conferencing applications, etc., could soon be proved a boon for the hearing-impaired community in India. A significant section of the Indian society suffers from hearing and speech impairment. This population uses Indian Sign Language as their primary mode of communication. Due to the difficulty in learning and understanding the meaning and context of written texts, sign language is preferred. Sign language involves the usage of hands, lip movements and expressions in order to communicate words, emotions and sounds. The proposed system provides an efficient method to aid communication between an individual with hearing and speech impairment. It is a field that has had little development over the years particularly in successful implementation in Python programming language. The system would improve access to information for the hearing-impaired population of the country like India. Moreover, the system can also act like an educational tool to learn ISL. Here, we have attempted to create a model that will allow people with disabilities to express themselves distinctly, which will help them blend with the rest of the world without any difficulty. Our proposed model will successfully convert the given input audio into animated video. Many improvements along this route can be made as and when the ISL Dictionary grows.

Future Scope

- In future, the proposed approach will be tested against unseen sentences. Furthermore, machine translation approach will be studied and implemented on parallel corpora of English and ISL sentences. The ISL corpus will be used for testing ISL sentences and the performance will be evaluated with evaluation parameters.

- This could enable sign language users to access personal assistants, to use text- based systems, to search sign language video content and to use automated real-time translation when human interpreters are not available. With the help of AI, automated sign language translation systems could help break down communication barriers for deaf individuals.
- Various front-end options are available such as .net or android app, that can be used to make the system cross platform and increase the availability of the system.
- The system can be extended to incorporate the knowledge of facial expressions and body language too so that there is a complete understanding of the context and tone of the input speech.
- A mobile and web based version of the application will increase the reach to more people.
- Integrating hand gesture recognition system using computer vision for establishing 2-waycommunication system.
- We can develop a complete product that will help the speech and hearing impaired people, and thereby reduce the communication gap.
-

References:

1. Tewari Y, Soni P, Singh S, Turlapati, M.S, Bhuva A. Real Time Sign Language Recognition Framework for Two Way Communication. In Proceedings of the International Conference on Communication, Information and Computing Technology, ICCICT, Mumbai, India, 25–27 June2021.

2. Kunjumon, J, Megalingam, R.K. Hand gesture recognition system for translating indian sign language into text and speech. In Proceedings of the 2019 International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, 27–29 November 2019; pp. 14–18

3. Gangadia D, Chamaria V, Doshi V, Gandhi J. Indian Sign Language Interpretation and Sentence Formation. In Proceedings of the 2020 IEEE Pune Section International Conference, Pune 2020, Pune, India, 16–18 December 2020; pp. 71–76

4. Shangeetha R.K, Valliammai V, Padmavathi S. Computer vision based approach for Indian sign language character recognition. In Proceedings of the 2012 International Conference on Machine Vision and Image Processing, MVIP, Coimbatore, India, 14–15 December 2012; pp. 181–184

5. Sawant S.N, Kumbhar M.S. Real time sign language recognition using PCA. In Proceedings of the 2014 IEEE International Conference on Advanced Communication, Control and Computing Technologies, ICACCCT 2014, Ramanathapuram, India, 8–10 May 2014; pp. 1412–1415.

6. Papadogiorgaki M, Grammalidis N, Tzovaras D, Strintzis M.G. Text-to-sign language synthesis tool; Text-to-sign language synthesis tool. In Proceedings of the 2005 13th European Signal Processing Conference, Antalya, Turkey, 4–8 September 2005.

7. Sonare B, Padgal A, Gaikwad Y, Patil A. Video-based sign language translation system using machine learning. In Proceedings of the 2021 2nd International Conference for Emerging Technology, INCET, Belagavi, India, 21 May 2021.