Audio To Sign Language Conversion

Sathvik M
School of Computer Science and Engineering
Vellore Institute of Technology
Chennai
sathvik.m2018@vitstudent.ac.in

Abstract—**Sign language is a visual language use by the deaf and dumb as theor mother tongue, this can used by a person who has difficulties speaking or by a person who can hear but cannot speak and also by normal people to communicate the hearing disabled people. Our project aims to bridge the gap between the deaf and dumb people and as well as the normal people with the advent of web applications, natural language processing.The main purpose of the project is to build an interface which accpets audio as input and converts them to respective sign language.This project builds upon the rules of Indian Sign Language and follows the ISL grammar.**

Keywords—Sign language, ISL, Deaf and Dumb

I. Introduction

The current indian population is 1.3 billion , in that the population which is vocally impaired and hearing impaired are more than 12.3 million and only 4.5 million can only afford to the schools for the deaf and dumb if available. There was a huge hit back to the whole educational institutions throughout India as well as other countries because of the recent pandemic situation that captivated the whole world. Thereby the mode of education was changed to online which was hard not only for studying but for teaching too. The ISL (Indian Sign Language) is used for communication and also for educational purposes but this has been hard for the dumb and deaf compared to the Oral based online mode of education because the communication is rather done by words than signs. The project helps in decreasing the hardships and increasing the quality of the education of living of this community by intervening and providing this community with a web page that is capable of translating english live audio to sign language and converting text to sign language as well

II. Literature Survey

i Neha Poddar,Shrushti Rao,Shruti Sawant, Vrushali Somavanshi: Study of Sign Language Translation using Gesture Recognition- This project is a small step towards helping physically challenged people and a lot more can be done to make the product more sophisticated, user friendly & efficient. Using more memory and powerful microprocessors, more languages can be covered. This project can be modified to make it compatible with mobile phones. We can increase the range of products by using more powerful trans-receiver modules.

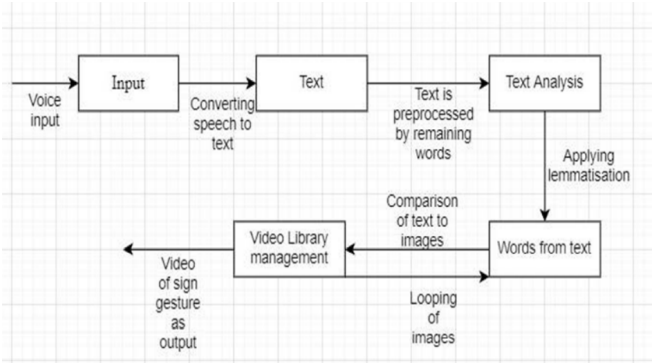
ii Pawan Kumar, Savita Khatri: Generating Indian Sign Language Text Using English/Hindi Text- In this paper, they have compared the existing models for the conversion of plain English/Hindi Text into the Indian Sign Language Text. Most of the systems which have been developed so far are meant to work only for a single domain. A unified system is needed for the deaf-mute community to help them interact with the normal people in their day to day lives.

iii Ankita Harkude, Sarika Namade, Shefali Patil, Anita Morey: Audio to Sign Language Translation for Deaf People- In this paper, they have implemented translation from input audio to sign language. It translates each and every letter in the audio message and it displays each sign language letter. But it is not efficient for future purposes.

III. Proposed Work

In this project, we will use the NLTK library and the JavaScript Web Speech API to create a web application that receives audio speech or text as input and translates it to equivalent Indian Sign Language for persons who are unable to talk or hear, as well as deaf people. First, we receive input in either audio or text format, as specified by the user. If the input is in the form of audio, we will use the JavaScript Web Speech API to transform it to text and show it on the screen. If the input is text-based, we will show it straight on the screen. The text input will then be processed using the NLTK library to break it into independent words without affecting the essential sense of the phrase, and the words will be shown on the screen. The animation that corresponds to the text will then be shown and played on the screen. We will use HTML5, CSS, and JavaScript to develop a responsive UI for improved communication in our online application. HTML5 is used to create our project's web page, CSS is used to make our web application's UI

more responsive, and JavaScript is used in the project for login/logout and sign in procedures. It also makes use of the Web Speech API to record audio speech and convert it to text representation. We will be utilizing the Lemmatization technique from the NLTK library to parse text. Lemmatization, in general, refers to doing things correctly by the use of vocabulary and morphological study of words. The ends of the words are deleted in this procedure to restore the base word, also known as Lemma. We will present it in video format as the final outcome. However, we still have to determine whether to create a film of ourselves doing ISL motions or to obtain information from the internet. The visual output will be matched with the processed text words and presented on the screen.



Modules Involved:

1. Getting input from User:
● Users can provide input in two ways. Users can use the mic symbol button to record the speech.
● Another way, users can type manually in the given text box.
● If User gives the input manually in the text box, the second module which is conversion of audio to text can be ignored.

2. Conversion of Audio into text using JavaScript Web Speech API:
● When User gives input in the form of audio, then the audio will be converted to text and displayed in the given text box.
● The audio can be recognized in various accents like American accent, British Accent or Indian Accent.
● According to the accent, the text will be recognized from the audio and displayed.

3. Text analysis by removing stop words:
● After recognizing the text from audio, the text will be analyzed using the NLTK library.
● In this module, the acquired text will be pre-processed by removing stop words from the text.
● Stop words are a group of words that are widely used in a language. Stop words in English include "a," "the," "is," "are," and others.
● Stop words are frequently used in Text Mining and Natural Language Processing (NLP) to exclude terms that are so widely used that they contain very little meaningful information.

4. Text mining by applying lemmatization:
● In this module, after removing stop words from the text, the pre-processed text will be processed using the lemmatization method which is available in the NLTK library.
● Lemmatization typically refers to doing things correctly by using a vocabulary and morphological analysis of words, with the goal of removing only inflectional ends and returning the base or dictionary form of a word, which is known as the lemma.
● The reason why we chose lemmatization over stemming is that Lemmatization accuracy is more as compared to Stemming, and Lemmatization would be recommended when the meaning of the word is important for analysis.

5. Display the corresponding sign language video:
● After the text mining process is finished, all of the processed words will be saved in a list, and the text of sign language videos will be compared to words in the list.
● Displaying the matching video to the user for the word in the list. There will be a play/pause button, and you will be able to pause and play the video whenever you like.

IV. Experimental Setup

Hardware Requirements:
- Minimum requirements:4GB of RAM and dual core processor
- Recommended requirements:8GB of RAM and quad core processor
- Microphone
- Speaker
- Internet connectivity

Software Requirements:
- Programming language:Python 2.7 or above
- Libraries required: NLTK, Django
- Programming IDE:Pycharm professional

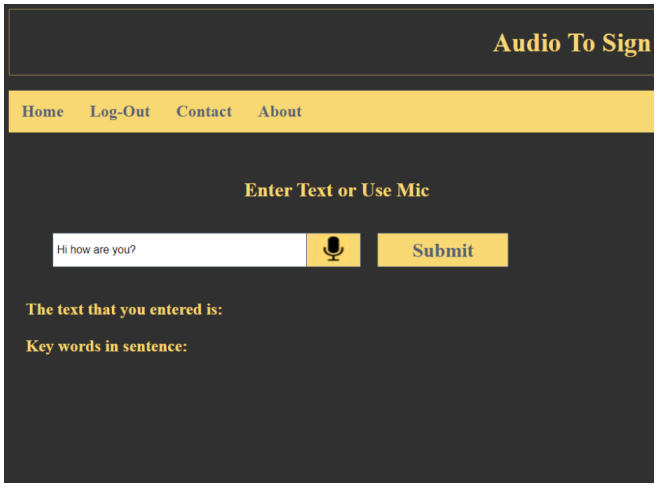
V. Implementation of the system

We are getting a live audio input using the Webkit speech recognition and then we are converting into the text form so that it will be helpful for us in doing the necessary processes. After getting the audio and converting into the text format , we need to break the words so that we can if they match with the dataset animation gif and display the word so in order to do that we need to break the text sentence , so for that we use POSTagging method which is a natural language processing concept after using this all the words in the sentence will get a specific tense to it and after doing it , the stop words are all removed then the maximum tense of the whole sentence is given to whole sentence. Then using the lemmatization process which is another

natural language processing concept, this basically splits the sentence into words and checks if there is a specific word gif , if there then provides it as output else it will further split the word into letters and will provide the output

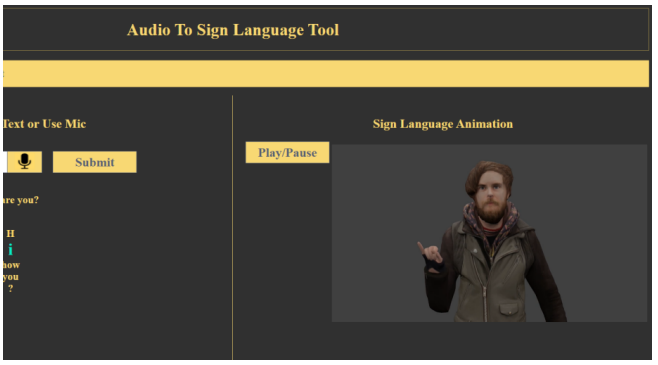
VI. Result and Discussion

The conversion of the live english audio is converted to the text



The output of the lemmatization process in which the respective splitting is done in accordance to the dataset that

The output of the project will be a clip of ISL gifs which is from a predefined database having video for most of the words and each every alphabest.



VII. Conclusion and Future Work

Our project will definitely help to remove the communication gap between the society of people and as well as the people from the deaf and dumb society, the tool that we did is a small step from that process, the small step is conversion of live feed audio into text which is currently used for online class for the people of deaf and dukb community.On a long run we have planned to integrate the project with a database so the inoformtion that have been searched will be saved thereby the person can access it and see whenever they want to infact that is the main reason why the option of setting up one account was created. The next idea is to add more languages detecting and conversion , because not just english is spoken along the world there are a lot of languages which have been spoken so the conversion just being able to convert english is not a big milestone.Now we are able to convert english audio file to sign language, we might add another way in which the sign language also is converted to audio format which is like the reverse of the process that we did. The last one is to be able to capture the direct audio that is like now we are able to take a live audio and then record it in another device but we are planning to make this webpage to detect the audio from the same device and then show it in sign language like the work done by google caption in a google meet.

References

[1] Pawan Kumar et al. International Journal of Recent Research Aspects ISSN: 2349-7688, Special Issue: Conscientious and Unimpeachable Technologies 2016, pp. 30-33 © 2016 IJRRA All Rights Reserved page - 30- Generating Indian Sign Language Text Using English/Hindi Text Pawan Kumar, Savita Khatri

[2] International Journal of Advanced Research in Computer and Communication Engineering Vol. 4, Issue 2, February 2015 Copyright to IJARCCE DOI 10.17148/IJARCCE.2015.4258 264 Study of Sign Language Translation using Gesture Recognition Neha Poddar1 , Shrushti Rao2 , Shruti Sawant3 , Vrushali Somavanshi4 , Prof. Sumita Chandak5

[3] ISSN: 2277-3754 ISO 9001:2008 Certified International Journal of Engineering and Innovative Technology (IJEIT) Volume 9, Issue 10, April 2020 DOI:10.17605/OSF.IO/EZ6AG Page 30 Audio to Sign Language Translation for Deaf People Ankita Harkude#1, Sarika Namade#2, Shefali Patil#3, Anita Morey #4 1,2,3,4#Department of Information Technology, Usha Mittal Institute of Technology, SNDT Women's University,