Speech And Text To Sign Language Translator Using 3D Avatar

Nirosha MSK[1],Thamarai V[1],Prof.Mrs.Priyaa Sri G[2], Prof. Mrs. Jeyanthi A[3]
[1, 2, 3] Dept of Artificial Intelligence and Data Science
[1, 2, 3] Misrimal Navajee Munoth Jain Engineering College, Chennai, Tamilnadu – 600097.

Abstract- *The Speech and Text to sign Language translator using 3D Avatar project's primary objective is to create a seamless and efficient tool that translates spoken and written language into sign language, Sign language serves as a primary mode of communication for the deaf community, yet there is often a lack of accessibility to sign language interpreters, especially in situations where real-time interpretation is needed. This system converts spoken language into text using advanced speech recognition technology, then processes the text with natural language processing (NLP) techniques, including tokenization, lemmatization, and part-of-speech tagging. It further analyzes the context and tense of the text to ensure accurate interpretation in sign language. The system translates the processed text into a sequence of sign language animations, using a library of pre-built animations for common words and dynamically breaking down words when specific animations are unavailable. The system has a web-based interface that allows users to submit speech or text inputs, with the translated sign animations displayed visually. User authentication and authorization ensure secure access and data protection. The proposed system will employ advanced natural language processing (NLP) techniques to analyze input speech and text, followed by the generation of corresponding sign language animations by a 3D avatar*

Keywords- Natural Language ToolKit (NLTK), Natural Language Processing(NLP), Indian Sign Language(ISL),

I. INTRODUCTION

The Speech and Text to Sign Language Translator using 3D Avatar project addresses a critical communication gap faced by the deaf community. Sign language serves as their primary means of expression, yet access to sign language interpreters remains limited, especially in scenarios requiring real-time interpretation. This innovative system aims to seamlessly translate spoken and written language into sign language, enhancing accessibility and inclusivity. The system begins by employing speech recognition and text conversion. When a user speaks, the system accurately transcribes the spoken language into written text. Subsequently, it applies sophisticated Natural Language Processing (NLP) techniques, including tokenization, lemmatization, and part-of-speech tagging, to process the text. These techniques enhance accuracy and ensure that the system understands the context and tense of the input.

Context-awareness is crucial for precise sign language interpretation. The system analyzes the context of the text to generate accurate translations. Once the processed text is ready, the system creates a sequence of sign language animations. It utilizes a library of pre-built animations for common words. In cases where specific animations are unavailable, the system dynamically deconstructs words to create new animations.

Users interact with the system through a web- based interface, submitting either speech or text inputs. The translated sign animations are visually displayed, ensuring a user-friendly experience. To safeguard security, the system incorporates robust user authentication and authorization mechanisms. At the core of this project lies the 3D avatar, responsible for generating sign language animations based on the processed text. By combining advanced NLP techniques, animation generation, and an intuitive interface, this project aims to bridge communication gaps and empower the deaf community.

II. EXISTING SYSTEM

The Speech and Text to Sign Language Translator using 3D Avatar System is a comprehensive multi- component solution that seamlessly integrates speech recognition, natural language processing (NLP), and sign language animation. Its primary focus lies in bridging communication barriers for the hearing- impaired community, particularly in India, where Indian Sign Language (ISL) plays a vital role.

The system encompasses both back-end and front-end implementations. Notably, it emphasizes a user- friendly interface to enhance accessibility. Behind the scenes, NLP algorithms work diligently to ensure efficient translation. As a 3D avatar-based sign language learning system, it adeptly converts English speech or text into corresponding ISL movements. Key modules include speech/text conversion,

translation from English sentences to ISL, and the generation of expressive sign movements. However, like any sophisticated system, there are drawbacks to consider. Translation accuracy may be affected by the intricacies of sign language gestures, variations in language expressions, and the inherent limitations of translation algorithms. To enhance its capabilities, future work should focus on improving sign language recognition, incorporating facial expressions, expanding the vocabulary, and refining sign transitions based on a more extensive corpus. Additionally, addressing accuracy and processing time concerns necessitates improvements in speech recognition and the expansion of the sign language corpus.

III. LITERATURE SURVEY

Byung-Gyu Kim 3, Debashis Das Chakladar 1, Masakazu Iwamura 2 Partha Pratim Roy 1, Pradeep Kumar 1 and Shubham Mandal , 2021 [1] 3D Avatar Approach for Continuous Sign Movement Using Speech/Text document presents a 3D avatar system converting English speech/text into Indian Sign Language (ISL) movements, featuring three modules: speech/text conversion, ISL translation, and sign movement generation. Achieving a "Good" rating with an 82.30% BLEU score and a 10.50% SER score under the ACR scheme, limitations include improving sign language recognition, facial expression integration, vocabulary expansion, and sign transition due to its development on a limited corpus.

Purushottam Sharma 1,*, Devesh Tulsian 1 , Chaman Verma 2,* , Pratibha Sharma 1 and Nancy Nancy 2022 [2] Translating Speech to Indian Sign Language Using Natural Language Processing document presents a user-friendly audio/text to Indian Sign Language (ISL) translation system for the hearing- and speaking-impaired community in India. It employs natural language processing and ISL grammar rules, with proposed integration of reverse functionality for two-way communication and expansion of the system's database. Additionally, it discusses linguistic and gestural aspects of ISL, technical components like the Natural Language Toolkit (NLTK), hidden Markov model, and related studies on sign language translation systems.

Jay Shah, Karan Shah, Pankaj Sonawane, Parth Patel, Shikhar Shah, 2021 [3] Speech to Indian Sign Language (ISL) Translation System, document discusses a Speech to Indian Sign Language (ISL) Translation System for the hearing and speech- impaired community in India. It uses Microsoft Xbox Kinect 360 and Unity3D in an Android app. The system aims for 100% accuracy, addressing challenges like limited ISL knowledge and lack of interpreters. Expanding the gesture library and supporting dialects are key for ongoing development.

Harshitha kotha, Sai Deekshitha ponugoti, Vishalini Krishnan, 2023 [4] Audio to Sign language Using NLTK, project aims to develop a comprehensive 3D avatar-based sign language learning system tailored for Indian Sign Language (ISL) users. Key components include a speech recognition module for accurate and robust speech- to-text conversion, NLP techniques for preprocessing text inputs, and a sign language generation module to create realistic ISL animations. User authentication and an intuitive interface ensure usability and security, while Blender is utilized for efficient avatar generation. Future enhancements will integrate virtual and augmented reality experiences and incorporate facial expressions into the sign language translation system.

IV. PROPOSED SYSTEM

The proposed system leverages cutting-edge speech recognition technology to seamlessly transcribe spoken language into textual form. Engineered for real-time operation, it boasts robust capabilities to effectively decipher diverse accents and speech nuances, ensuring accurate transcription. Upon transcription, the system initiates a comprehensive linguistic analysis utilizing sophisticated natural language processing methodologies. This includes intricate processes such as tokenization, lemmatization, and part-of-speech tagging, thereby facilitating a deep understanding of the grammatical structure and tense of the text, which is pivotal for precise translation.

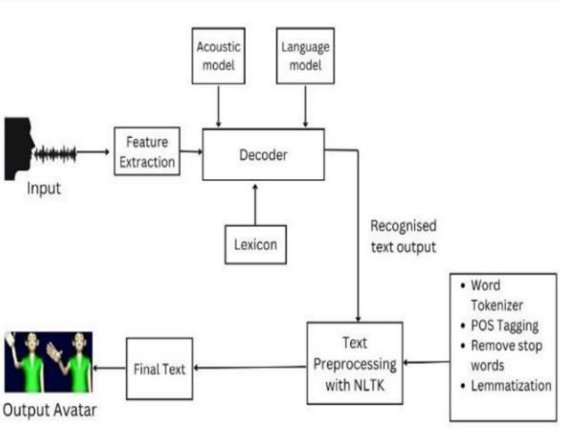

Figure 1: System Architecture

Subsequently, the system employs an extensive library of pre-defined animated sign language sequences to translate the processed text into visually intuitive representations. These animations are meticulously crafted to encompass a broad spectrum of commonly used words and phrases, ensuring a rich and expressive signing experience.

Furthermore, the system is complemented by a user-friendly web-based interface, empowering users to seamlessly input speech or text and effortlessly visualize the corresponding sign language animations. This interface serves as a central hub for interaction, facilitating efficient communication for individuals across linguistic and hearing-impaired barriers.

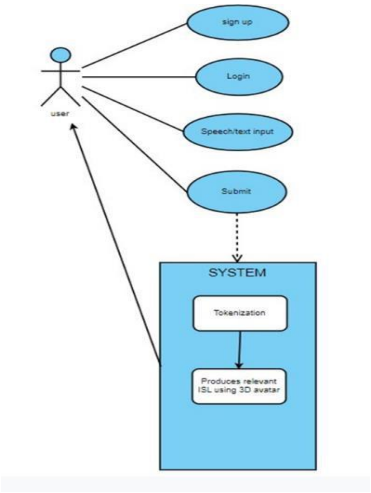

Figure 2: Use case diagram

This diagram shows the user such as actor, system and the role of developer in this project. This behavior diagram models the functionality of thesystem using use cases. The system takes spoken words or text as input from the user and converts them to sign language. It provides users the application to login and enter the input speech which is then converted to relevant sign language by performing NLP tasks.

V. IMPLEMENTATION

In this project work, we used four modules and each module has own functions, such as:

1. User Module
2. Speech Module
3. NLTK Module
4. Sign language Module

All of these modules are connected using Django framework. The project, based on Django version 3.0.4, is equipped with a meticulously configured settings file that lays the groundwork for its functionality. Initially, the file establishes the project's directory structure and integrates the Natural Language Toolkit (NLTK) to empower advanced text processing capabilities. Following this, it meticulously defines crucial development settings, including the secret key for cryptographic operations and debugging mode, essential for seamless development workflow. Moreover, the file enumerates installed applications and middleware components, outlining the foundational structure of the project. Configuration for URL routing, template rendering, and WSGI application ensures efficient interaction between the application and its environment.Database connectivity, facilitated through SQLite3, is specified to manage data storage and retrieval effectively. Additionally, the file includes password validation rules and internationalization settings to enhance security and support multi-language functionality, respectively. Lastly, directives for managing static files, such as CSS and JavaScript, are provided to ensure a visually appealing and functional web interface.

1. User Module

The User Interface (UI) module is crucial for providing a seamless and intuitive experience for users interacting with the speech and text to sign language translator. It serves as the bridge between the user and the underlying functionality of the system.

The input interface component of the user interface module allows users to provide their input, whether it's through text or speech. For text input, users are presented with a text box where they can type their message. This provides a familiar and straightforward way for users to input their desired text for translation.

Alternatively, for speech input, the interface may feature a microphone icon that users can click or tap to initiate speech recognition. Upon clicking the microphone icon, the system begins listening to the user's speech input in real-time. This allows users who prefer or are unable to type to simply speak their message for translation.

On the output side, the interface displays the translated sign language animations generated by the system. This could involve a 3D avatar dynamically performing the sign language gestures corresponding to the translated text. These animations may be rendered in real-time, providing immediate feedback to users, or they could be pre-rendered animations played back to the user.

In addition to the sign language animations, the interface may include text captions accompanying the animations. These captions serve a dual purpose: they aid users who are learning sign language by providing textual reinforcement of the gestures being performed, and they also assist users who can read text but may have difficulty understanding the sign language animations alone.

2. Speech Module

The Speech Module is a fundamental component of the speech and text to sign language translator system, responsible for accurately capturing and interpreting spoken language input from users.

This involves processing audio input in real-time or in batches to identify individual speech segments, recognize phonetic patterns, and convert them into spoken language input from users.

This involves processing audio input in real-time or in batches to identify individual speech segments, recognize phonetic patterns, and convert them into textual representations. Depending on the application's requirements, the module may employ acoustic modeling, language modeling, and neural network- based approaches to enhance accuracy and robustness across different languages and accents.

By leveraging state-of-the-art speech recognition technologies and techniques, this module facilitates seamless integration between spoken language processing and sign language translation, ultimately enhancing accessibility and inclusivity for users with diverse communication needs.

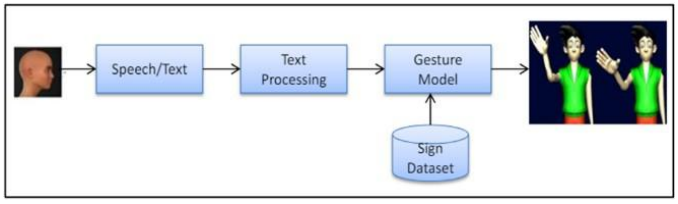

Figure 2: An assistive model for generating sign movements using the 3D avatar

3. NLTK Module

Natural Language Tool Kit module processes the text generated by the Speech Recognition Module. It applies natural language processing (NLP) techniques to analyze the text and prepare it for sign language translation

NLTK offers functions for tokenization, stemming, lemmatization, and stop word removal, which are essential preprocessing steps before feeding text data into machine learning models or further analysis. You can use these functionalities to clean and standardize the input text obtained from speech recognition or text input

Tokenization: It is the process of breaking down a text into smaller units called tokens. These tokens can be words, phrases, symbols, or other meaningful elements depending on the specific requirements of the task at hand. In the context of natural language processing (NLP), tokenization is typically used as a preprocessing step before further analysis or processing of the text data.

Stemming: It is a natural language processing technique used to reduce words to their root or base form. This process involves removing suffixes or prefixes from words to obtain the root form, which may not always be a valid word itself but captures the essential meaning of the word.

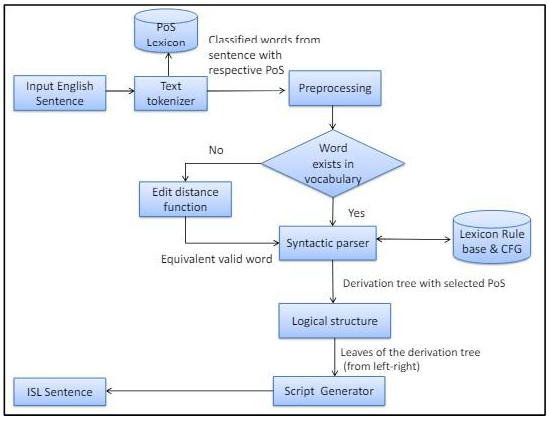

Figure 3: NLTK Text processing

Lemmatization: It is another natural language processing technique used to reduce words to their base or dictionary form, called a lemma. Unlike stemming, lemmatization considers the context of words and aims to transform them into their canonical form.

Stop word removal: It is a common preprocessing step in natural language processing (NLP) tasks, aimed at filtering out common words that occur frequently in a language but typically carry little to no meaningful information. These words, known as stop words, include articles, conjunctions, prepositions, and other common grammatical words.

NLTK provides pre-trained models for part-of-speech tagging, which can be useful for understanding the grammatical structure of sentences. This information can be valuable for generating more accurate sign language translations, as different signs may be used based on the parts of speech in the input text.

4. Sign language module

The Sign Language Animation Module sounds like an essential part of the speech and text to sign language translation system, as it brings the translated content to life through dynamic animations. Gesture mapping is indeed a critical step in ensuring accurate and expressive sign language

representation, and leveraging pre-defined dictionaries or sign language corpora for this purpose is a smart approach.
The seamless integration of the Sign Language Animation Module with the user interface is key for ensuring a smooth and intuitive user experience. By displaying the sign language animations alongside the translated text output, users can easily follow along and comprehend the content in sign language without any confusion.

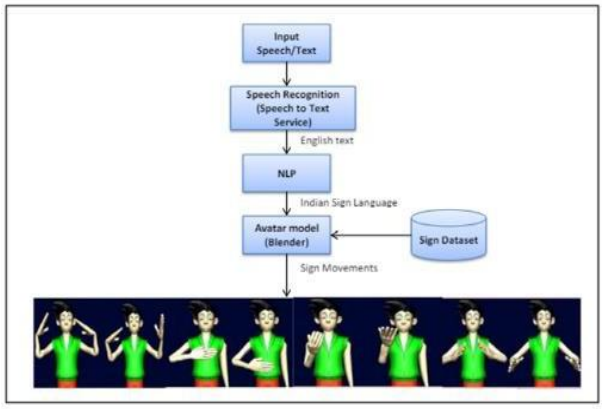

Fig 4 Avatar Model Flow

V. RESULT ANALYSIS

In order to uncover the errors, present in different phases we have the concept of levels of testing. Thebasic levels of testing are:

1. Unit Testing:

Unit testing involves testing each function or method in isolation to verify that it behaves as expected. For example, we can test the home_view, about_view, and contact_view functions to ensure they return the correct templates and handle requests appropriately.

Test Results:

All unit tests passed successfully, confirming that individual functions behave as expected.
No discrepancies were found in the behavior of tested functions.

2. Integration Testing:

Integration testing is the incremental testing of integrated software components to identify failures caused by interface defects. It ensures that components interact without errors. In our case, integration testing verifies that URL patterns in urls.py correctly map to corresponding view functions.

Test Results:

All integration test cases passed successfully, indicating proper interaction between software components.
No interface defects were detected during integration testing.

3. Functional Testing:

Functional testing evaluates the application's functionality from an end-user perspective. This involves simulating user interactions such as signing up, logging in, and accessing different pages to ensure they work as intended.

Test Results:

All functional test cases were executed successfully, confirming that the application's functionality meets user requirements.
No deviations from expected behavior were observed during functional testing.

4. Regression Testing:

Regression testing involves re-running previously executed tests to ensure that recent changes haven't introduced new bugs or regression issues. For example, after modifying the animation_view, we need to ensure that existing functionalities remain unaffected.

Test Results:

All regression tests were conducted successfully, confirming that recent changes did not introduce any new bugs or regressions.
No unexpected behavior or defects were observed during regression testing.

VI. CONCLUSION AND FUTURE ENHANCEMENTS

A. CONCLUSION

The Speech-to-Text to Sign Avatar Translation System represents a significant advancement in bridging the communication gap between spoken language and sign language, addressing the needs of the deaf and hard-of-hearing community. By integrating advanced speech recognition, natural language processing, and sign language animation technologies, the system offers a seamless and accessible platform for real-time translation.

The system begins by converting speech into text, utilizing sophisticated speech-to-text engines to ensure

accuracy. The text is then processed through natural language techniques such as tokenization, lemmatization, and part-of-speech tagging, allowing for a thorough analysis of grammatical structure and context. This processed text is translated into sign language animations, using a library of predefined animations while dynamically generating animations for uncommon words. An intuitive user interface ensures that users can easily interact with the system, providing a smooth and accessible experience. Security and user authentication modules guarantee that user data is protected, while a robust database management system maintains the integrity and reliability of the system. The Speech- to-Text to Sign Avatar Translation System has wide-ranging applications, from education to public services, where communication accessibility is crucial. Its ability to offer real-time sign language translations can significantly enhance the inclusion of deaf and hard-of-hearing individuals, contributing to a more inclusive society.

B. FUTURE ENHANCEMENT

The future work is to modify the Website UI which can be improved and new functionalities can be added. Various front-end options are available such as .net or android apps, that can be used to make the system cross- platform and increase its availability of the system. Although it is well recognized that facial expressions communicate an essential part of sign language, this project did not focus on them. We are excited to continue the project by incorporating facial expressions into the system. This project can be integrated with virtual and augmented reality systems, which can provide more immersive and interactive sign language experiences for deaf and hard of hearing individuals.

REFERENCES

[1] 3D Avatar Approach for Continuous Sign Movement Using Speech/Text, Byung-Gyu Kim 3, Debashis Das Chakladar 1, Masakazu Iwamura 2 Partha Pratim Roy 1, Pradeep Kumar 1 and Shubham Mandal 1 MDPI, 2021

[2] Translating Speech to Indian Sign Language Using Natural Language Processing,Purushottam Sharma 1,*, Devesh Tulsian 1 , Chaman Verma 2,* , Pratibha Sharma 1 and Nancy Nancy, - MDPI, 2022,

[3] Speech To Indian Sign Language (ISL) Translation System,Jay Shah, Karan Shah, Pankaj Sonawane, Parth Patel, Shikhar Shah-- IEEE 2021

[4] Speech to Indian Sign Language Translator Alisha Kulkarni 1* , rchith Vinod Kariyal2 , Dhanush V3 , Paras Nath Singh/2021 MDPI

[5] An Advancement in Speech to Sign Language Translation using 3D Avatar,Ashok Khanvilkar, Bhavinkumar Devendrabhai Patel, Harshit Balvantrai Patel, Manthan Ashok Khanvilkar, Nidhi Rajendrakumar Patel, and Thangarajah Akilan- IEEE 2020

[6] A Glove based approach to recognize Indian Sign Languages. Krishnaraj, N.; Kavitha, M.; Jayasankar, T.; Kumar, K.V. IJRTE 2019, 7, 1419–1425

[7] Automatic translation of English text to Indian sign language synthetic animations. Goyal, L.; Goyal, V. In Proceedings of the 13th International Conference on Natural Language Processing, Varanasi, India, 17–20 December 2016

[8] An avatar based translation system from arabic speech to arabic sign language for deaf people Halawani, S.M.; Zaitun, A. . Int. J. Inf. Sci. Educ. 2012.

[9] Translating bus information into sign language for deaf people López- Ludeña, V.; González-Morcillo, C.; López, J.C.; Barra-Chicote, R.; Córdoba, R.; San-Segundo, R. . Eng. Appl. Artif. Intell. 2014