

Smart Communication Aid: An AI-Driven Sign Language Converter using ASR and NLP

M.Jagadeesh¹

*Assistant Professor,
Department of Computing
Technologies,
SRM Institute of Science &
Technology,*

Kattankulathur, Chennai-603203,
Tamil Nadu, India,
jagadeeshdevika@gmail.com

Pranav Srivastava²

*Undergraduate Student,
Department of Computing
Technologies,
SRM Institute of Science &
Technology,*

Kattankulathur, Chennai-603203,
Tamil Nadu, India,
ps2956@srmist.edu.in

Krrish Garg³

*Undergraduate Student,
Department of Computing
Technologies,
SRM Institute of Science &
Technology,*

Kattankulathur, Chennai-603203,
Tamil Nadu, India,
kg5539@srmist.edu.in

Abstract- Hearing-impaired individuals often struggle to communicate effectively with those who can speak, leading to a significant communication gap in daily life. Previous solutions, such as text-based applications or human interpreters, have presented various challenges, including slow processing, dependency on another person, or ineffective representation of sign language. To address this issue, this research introduces a Text and Speech-to-Sign Language Converter that instantly translates written text or spoken words into sign language. The system first accepts input in the form of text or speech. If speech is provided, Automatic Speech Recognition (ASR) converts it into text. Then, a Natural Language Processing (NLP) model processes the text and translates it into corresponding sign language gestures. These gestures are displayed through an animated character, making communication more interactive and accessible. A key feature of this project is the integration of a Kaggle dataset containing animated signs, ensuring clarity and accuracy in sign representation. Unlike previous methods, this system operates autonomously in real time without requiring human assistance. It has broad applications in schools, hospitals, public spaces, and daily interactions, significantly enhancing accessibility for individuals with hearing impairments and promoting their inclusion in society.

Keywords- Speech-to-Sign Language Conversion, Text-to-sign Language Conversion, Automatic Speech Recognition (ASR), Natural Language Processing (NLP), Sign Language Recognition (SLR).

I. INTRODUCTION

Communication is an essential part of life that allows people to share their thoughts, emotions, and ideas. However, individuals with hearing impairments often struggle to communicate with those who do not understand the sign language. Since sign language is their primary way of expressing themselves, the lack of widespread knowledge about it creates a barrier to daily interactions. This gap affects education, healthcare, workplaces, and social situations, making it difficult for hearing-impaired individuals to fully engage in society. Therefore, it is crucial to develop technologies that help to bridge the communication gap and make interactions smoother for everyone, with advancements in Artificial Intelligence (AI) and Natural Language Processing (NLP), real-time translation systems have become possible. This research focuses on developing a text/speech-to-sign language recognition system that uses NLP techniques to convert spoken or written language into sign language. This system ensures that individuals with hearing impairments can understand spoken conversations more easily. It works by recognizing speech and text, processing it through AI, and displaying equivalent sign language gestures in real time. In doing so, it provides a seamless and efficient communication tool for both hearing and non-hearing individuals.

speech recognition module listens to the spoken words and converts them into text using AI-based speech-to-text technology. Then, the Natural Language Processing (NLP) module analyses the text, understands the meaning of the sentence, and selects the appropriate sign language translation. Finally, the sign language animation module takes the processed text and converts it into animated sign language gestures that can be displayed on a screen. The step ensured that the translated signs were clear and easy to understand and system is highly useful for various applications. In education, it can help students with hearing impairments by converting spoken lectures into sign language, thus allowing them to follow classroom discussions more effectively. In healthcare, doctors and nurses can use this technology to communicate with patients who are unable to hear it. Public services can improve accessibility in government offices, customer service centers, and workplaces. On a personal level, this tool makes everyday conversations between hearing and non-hearing individuals much easier and more natural, thus reducing the need for human interpreters.

The main objective of this research is to develop a real-time system that accurately translates speech and text into sign language, ensuring smooth and effective communication. The model was designed to be user-friendly, highly accurate, and capable of providing natural sign language gestures. To achieve this, a pre-built dataset from Kaggle containing animated sign language gestures was used, which improved the accuracy and fluency of translation. The system also aims to be scalable, meaning that it can be expanded in the future to support multiple languages and different regional variations in sign language. One of the key challenges addressed by this research is the lack of expressive and natural sign language representations in existing solutions. Many traditional text-based communication tools fail to capture the real essence of sign language, thereby making conversations less engaging. By integrating NLP- and AI-driven animation, this project ensures a more natural and effective translation. In the future, this system can be further improved by adding gesture recognition, which will allow sign language users to communicate back using hand movements, thereby creating a fully interactive system.

This Text and Speech-to-Sign Language Recognition System is an important step toward breaking communication barriers and making society more inclusive. By providing a real-time AI-powered solution, it empowers individuals with hearing impairments to communicate more easily and independently. This project highlights the potential of technology in creating a more accessible world where everyone has the opportunity to engage and connect without limitations.

The system operates in three main steps as show in fig.1. first, the

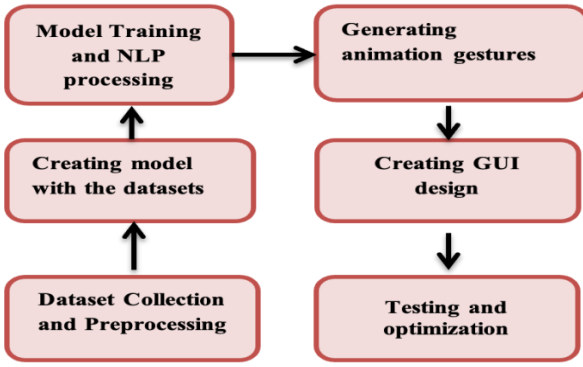


Fig.1. Research Process

II. PREVIOUS WORKS

Rinki Gupta et al. [1] this article presents real-time sign language recognition using wearable sensors (sEMG and accelerometers) to facilitate deaf-hearing interaction. In place of fixed-window segmentation, which is handicapped by changing hand movements, the authors propose the use of an ensemble of classifiers trained on features of different window sizes. Tested for 11 Indian Sign Language (ISL) sentences, the approach is more accurate compared with single-classifier models. The system can continuously sign-classify sentences in real time, demonstrating its potential for improved sign language recognition.

Kunal Chhabria et al. [2] this paper presents an Indian Sign Language (ISL) alphabet recognition system in real-time using deep learning. The system is a pipeline of data generation, model training, and frame-by-frame classification to identify gestures from a webcam without the assistance of expensive wearable devices. The system is operating with any background, where alphabets can be written out by creating respective gestures in front of the webcam. This is proof of the potential of deep learning for low-cost and accessible sign language translation.

Gautham et al. [3] explores study proposes an Indian Sign Language (ISL) hand gesture translation system for banks to facilitate the deaf-mute community to communicate their requirements effectively. Unlike previous studies on American Sign Language (ASL), this study addresses the lack of ISL datasets and gesture variation by using a self-recorded dataset. The model uses Inception V3 (CNN) for feature extraction and LSTM (RNN) for gesture classification from video frames, translating dynamic ISL signs into text. Results show that this approach provides an effective and accurate solution for communication between signers and non-signers in banking environments.

Soma Shrenika et al. [4] this paper presents a vision-based sign language recognition system through image processing algorithms like edge detection and template matching using OpenCV-Python to translate hand movements into text to facilitate communication with deaf and Autism Spectrum Disorder (ASD) patients.

D. S. Breland et al. [5] this paper introduces a low-cost, portable edge computing system for sign language digit recognition using thermal imaging. A 3,200-image (320 images per digit) thermal image dataset was created, and a low-latency deep learning classifier inspired by deep residual learning was proposed. The designed system, executed on a Raspberry Pi using a low-resolution 32×32 thermal camera, achieves 99.52% accuracy and is background illumination invariant, suitable for human-computer interaction, robotics, healthcare, smart devices, and crisis management.

S. J. Iyer et al. [6] this study resolves the insufficiency of Indian Sign Language (ISL) translation tools by introducing a computer vision-based solution to retrieve 2D pose data from single-camera RGB videos. In contrast to costly motion capture technologies, this technique offers a cost-effective solution for ISL recognition. The research emphasizes video stabilization, extracting human pose key points, and interpolating missing joint data based on a low-cost trajectory method. The objective is to achieve a non-distorted 2D pose model, acting as a preliminary step towards a future ISL translation system to close the communication gap for India's deaf community.

Y. Grover et al. [7] this paper discusses various sign language recognition and translation techniques for bridging the gap between hearing/speech-impaired individuals and the rest of humanity. It gives an overview of existing techniques, classifies them based on a proposed taxonomy, and provides a tabular comparison of various techniques with a review of literature.

R. Gupta et al. [8] this paper is focused on machine learning and deep learning-based automatic sign language recognition. The paper employs multi-modality wireless sensors on the dominant hand of the sign language signer to record data from 50 Indian Sign Language signs executed by five subjects. The data is processed using a homogeneous ensemble of CNN models, where features are input into randomly initialized CNNs to identify high-level features. A meta-learner converts class probabilities to determine the executed sign. Experimental results show that using a stacked ensemble of 25 CNN models with a multi-layer perceptron meta-learner achieves 87.6% accuracy, which is greater than using single CNN models that could achieve a maximum of 79.5% accuracy.

M. R. Chilukala et al. [9] this paper gives an overview of different methods applied in sign language recognition that employs deep learning to translate sign words or sentences into words, hence filling the communication gap between normal and deaf people. The paper discusses different recognition approaches of American, Indian, and Indonesian Sign Language from video and image sets.

A. F. Shokoory et al. [10] This research is deep learning vision-based sign language recognition, specifically for Pashto Sign Language. As Pashto is widely spoken in southern Afghanistan, the research aims to develop a system that will convert hand movement into Pashto alphabets with the aim of enhancing communication among the disabled. A dataset of 2,500 images was created and a CNN model was employed, and it achieved 98% accuracy in recognizing gestures.

M. R. Chilukala et al. [11] this research paper gives an overview of various sign language recognition techniques with the help of deep learning. It explains detection and recognition of signs from an image or video and how efficient they are in translating sign language into text. The research compares various techniques employed for various sign languages, i.e., American Sign Language (ASL), Indian Sign Language (ISL), and Indonesian Sign Language (ISL), and finds advancements in deep learning-based gesture recognition to bridge the gap in communication between normal and hearing-impaired people.

J. Peguda et al. [12] his work is committed to speech-Indian Sign Language (ISL) translation for six Indian regional languages (Hindi, Telugu, Malayalam, Marathi, Kannada, and Tamil). The proposed model uses Wavelet-based MFCC with GMM for speech recognition, LSTM for text translation, and gesture mapping to display the corresponding sign language sequence to enable hearing-impaired individuals to communicate in India.

W. Cao et al. [13] this paper proposes the Joint Semantic Representation (JSR) algorithm to improve Traffic Sign Recognition (TSR) by combining knowledge-driven and data-

driven approaches. JSR learns to extract discriminative visual features and applies traffic sign design regulations to achieve zero-shot learning for new signs. Its effectiveness is confirmed through experiments on multiple datasets.

E. Rajalakshmi et al. [14] her paper proposes DNN-SLR, an Indian and Russian sign language recognition system based on a hybrid deep neural network that is vision-only. It integrates spatial, temporal, and sequential feature extraction by a 3D deep neural net, Bi-LSTM based on attention, and modified autoencoders to improve the recognition accuracy. The introduced model outperforms existing frameworks on the newly created datasets and the WLASL dataset, overcoming critical shortcomings in real-world sign language recognition.

Deep R. Kothadiya et al. [15] this study suggests a Transformer Encoder-based method for Indian Sign Language Recognition with 99.29% accuracy and few training epochs. The technique employs vision transformers to encode sign gestures as positional embedding patches, with the help of multi-head attention for optimal recognition under diverse conditions. The model performs better than convolution-based architectures and holds promise for future development in isolated and continuous sign language recognition.

Deep Pal et al. [16] this study introduces a sign language recognition wearable optical fiber-based sensing glove, employing macrobends in single-mode fiber (SMF) to monitor finger movements and forecast hand gestures. The integrated system with an Optical Time-Domain Reflectometer (OTDR) gathers real-time data and classifies the gestures through an ensemble classifier with 93.57% accuracy in dynamic gesture recognition.

P.ala Chaitra et al. [17] this paper creates a CNN-based Indian and American Sign Language recognition model with 99.5% accuracy by recognizing both alphabets and words. The model tracks hand gestures using skin segmentation, extracts salient landmarks, and provides recognized signs as text and speech to improve hearing-impaired communication.

K. Kankaria et al. [18] this work constructs an LSTM and Bi-LSTM-based Indian Sign Language (ISL) classifier, employing MediaPipe-extracted skeletal features to classify gestures. Augmentation strategies and uniform sampling of frames are employed to enhance performance and diminish computational complexity to achieve 97.92% accuracy for successful automated ISL interpretation.

D. Sahu et al. [19] his work presents a two-stream sign language recognition (SLR) system based on the combination of GLCM for texture characteristics and ResNet-50 for deep semantic abstraction, along with PPCA for dimensionality reduction, yielding excellent accuracy in the classification of American and Indian Sign Language (ASL & ISL).

Kumari et al. [20] has constructed the interaction of the deaf and hearing-impaired population with the rest of the population, sign language recognition systems are employed. An Indian sign language (ISL) word recognition system based on computer vision has been suggested in this chapter. Redundant frames from the video sequence are removed by employing key frame extraction using edge difference technique that renders the system more efficient in processing.

Author(s)	Year	Merits	Demerits
Rinki Gupta et al. [1]	2020	High accuracy in real-time ISL recognition using wearable sensors.	Requires users to wear sensors, limiting accessibility.
Kunal Chhabria et al. [2]	2020	Works in real-time with just a webcam,	Limited to alphabet

		making it cost-effective.	recognition, not full sentence translation
Gautham et al. [3]	2020	Accurately translates ISL hand gestures into text for better banking communication.	Limited dataset and gesture variability may affect recognition accuracy.
Soma Shrenika et al. [4]	2020	Uses a simple camera-based approach, making it accessible and cost-effective.	Limited by template matching, which might struggle with complex or dynamic gestures.
D. S. Breland et al. [5]	2021	Achieves high accuracy (99.52%) and works in all lighting conditions due to thermal imaging.	Low-resolution thermal camera (32×32 pixels) may limit the recognition of complex gestures.
S. J. Iyer et al. [6]	2021	Cost-effective ISL recognition using computer vision, without expensive motion capture.	Accuracy may be affected by video quality, occlusions, and varying camera positions.
Y. Grover et al. [7]	2021	Offers a detailed classification and comparison of sign language recognition techniques, aiding future research.	Does not propose a new implementation, focusing only on reviewing existing methods.
R. Gupta et al. [8]	2022	Achieves higher accuracy (87.6%) using an ensemble CNN approach compared to individual CNN models.	Requires multiple sensors on the user's hand, non-intrusive applications.
M. R. Chilukala et al. [9]	2022	Provides a comprehensive comparison of sign language detection techniques across different languages.	Focuses on reviewing existing methods rather than proposing a new implementation or system
A. F. Shokoory et al. [10]	2022	Introduces Pashto Sign Language recognition, which is novel and previously underexplored.	Limited dataset (2,500 images) may not fully capture complex gestures or real-world variations.
M. R. Chilukala et al. [11]	2022	Covers multiple sign languages and provides a comparative analysis of recognition techniques.	Lacks experimental implementation, focusing only on a survey of existing methods.
J. Peguda et al. [12]	2022	Supports multiple Indian languages, enhancing accessibility for a diverse population. zero-inflated Poisson boosted tree model.	Limited scope as it only focuses on speech-to-sign conversion, excluding sign-to-text translation.
W. Cao et al. [13]	2023	Enables zero-shot learning for recognizing unseen traffic signs.	May introduce computational complexity in feature extraction and reasoning.
E. Rajalakshmi et al. [14]	2023	Effectively captures both manual and non-manual sign components for better recognition.	May struggle with continuous sign sentence recognition and segmentation ambiguities.

Deep R. Kothadiya et al. [15]	2023	Achieves high accuracy with minimal training effort.	Limited to static sign recognition, requiring further adaptation for continuous gestures.
Deep Pal et al. [16]	2024	Provides high-accuracy dynamic gesture recognition with real-time tracking.	Requires a specialized wearable glove, limiting accessibility and practicality.
Pala Chaitra et al. [17]	2024	Accurately recognizes both alphabets and words with high efficiency.	Limited to predefined gestures, making it less adaptable to real-world variations.
K. Kankaria et al. [18]	2024	Efficiently classifies ISL gestures with high accuracy.	Limited to skeletal features, which may miss finer hand and facial expressions.
D. Sahu et al. [19]	2024	Enhances recognition accuracy by integrating texture and deep features.	Increased computational complexity due to dual-stream processing.
Kumari et al. [20]	2025	The use of key frame extraction with the edge difference technique eliminates redundant frames.	The model has been tested on only 50 words , which may not generalize well to larger ISL vocabulary.

III. PROPOSED MODEL

The proposed system is designed to build the communication gap between individuals who depends on speaking or written language and another who communicate using sign language. People with hearing disabilities frequently faces difficulties in understanding spoken language, especially in environments where sign language interpreters are not available. Traditional solutions, such as hiring sign language interpreters or using text-based communication, are not always accessible or practical. To overcome these challenges, our model provides an AI-powered real-time system that translates both spoken words and written text into animated sign language gestures. The model guarantees that people who use sign language can understand conversations more easily and naturally.

The model converts text and speech inputs into corresponding sign language animations. The automatic Speech Recognition (ASR) is used to process verbal input and turn spoken words into text. If the user directly provides text input, then is sent for further processing. The Natural Language Processing (NLP) module then analyses the text and redesign the text to match sign language grammar. As sign language has a different syntax and word order compared to spoken languages, NLP makes sure that the translation is grammatically correct and precise. Once the text is processed, it is mapped to predefined sign animations stored in the dataset. The animations are displayed using an animated avatar, as shown in the

fig.2., making the communication more natural and expressive as compared to static images.

The Graphical User Interface (GUI) allows users to communicate with the system seamlessly. It provides options to input speech or text and display the translated sign animations, and adjust playback settings. The system is designed for real-time use, meaning that it can be install in classrooms, hospitals, workplaces, and public service centres to enable efficient communication among individuals with hearing disabilities

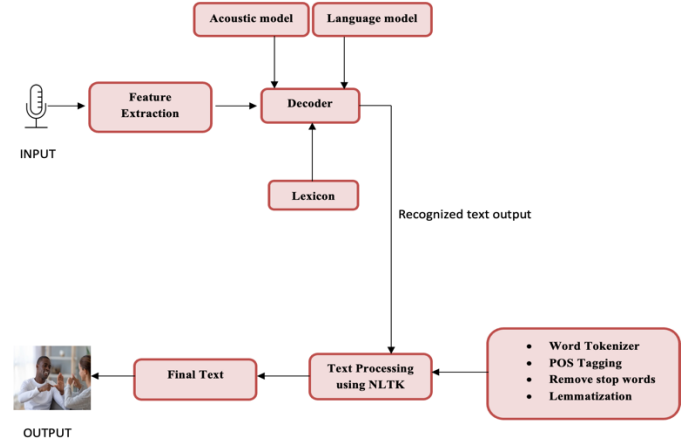


Fig.2. System Architecture

A. NLP-Driven Framework

The Foundation of this research is built upon Natural Language Processing (NLP) and NLP Toolkit, which work together to enable the seamless conversion of speech and text into sign language animations. NLP is a branch of artificial intelligence that helps machines understand, interpret, and generate human language, making it a crucial component of this system. As sign language follows a different grammar pattern compared to spoken and written languages, NLP ensures that the text input is processed, reframed, and correctly mapped to sign gestures, maintaining the natural communication flow.

The NLP Toolkit used here provides a comprehensive set of tools for text preprocessing, tokenization, part-of-speech (POS) tagging, syntactic parsing, and semantic analysis. These modules help the system to identify sentence format, extract meaningful linguistic features, and match these with corresponding sign language gestures. This ensures that instead of translating single character, the system produces accurate sign representations relevant to that context, managing the meaning of the original speech or text input. To further enhance translation accuracy, machine learning models combines with the NLP Toolkit are trained on linguistic structures, allowing the system to handle variations in sentence structures. The toolkit also plays a key role in translating complex sentence formations into easy to understand sign representations, ensuring that the generated output gestures are clear, expressive, and grammatically correct. By applying real-time NLP processing, the model chooses the most suitable animated sign gestures from the dataset, ensuring a smooth and natural sign language output.

Secondly, the combination of NLP and animated sign representations also fills the communication gap between sign and spoken languages, making it more accessible for deaf and hard-of-hearing users. This AI-driven process, combined with the capabilities of NLP Toolkit, enhances accuracy, fluency, and naturality in sign language translations, therefore making the

system suitable for various real-world applications, such as education, healthcare, and public services.

B. Utilizing Kaggle's Sign Language Animation Dataset for NLP Model Training

The most important aspect of this project is the quality and variability of the dataset used to train the sign language translation model. For this implementation, a Kaggle-based dataset "Indian Sign Language Animated Videos" was used, containing a wide range of sign language animations, gestures, and their corresponding textual labels. This dataset forms the core for training the model and achieving accurate and natural translations. The dataset includes hundreds of sign language gestures, each labelled with its respective word or phrase. Additionally, it contains video-based and GIF-based animations to ensure that the gestures are fluid and smooth and not rigid images. With high-quality pre-recorded animations, the system can provide a more natural and engaging experience, before using the dataset for model training, preprocessing steps were applied to enhance its effectiveness. These steps included standardizing the speed and resolution of animations, ensuring that all gestures maintain uniformity. Further, preprocessing involved segmentation of video sequences in frame-by-frame animation form in order to increase model processing facility as well as a smooth shifting from one gesture to another. Another key aspect of dataset preparation was annotating and labelling animations correctly, ensuring that each sign was mapped accurately to its corresponding text representation. Through the use of a well-organized and varied dataset, the system is able to ensure that sign translations not only are accurate but also contextually relevant. The application of animations instead of static pictures considerably enhances the overall clarity, realism, and natural flow of the created signs. This dataset is the foundation of the system, allowing it to seamlessly bridge the gap between spoken and sign languages.

C. Model Integration and System Implementation

The implementation of this system involves various key phases, each of which helps in real-time conversion of text and speech into sign language animations. The four major phases of implementation include Speech-to-Text Conversion, Text Processing, Sign Animation Selection, and GUI Integration.

i. Speech-to-Text Conversion using ASR

The first stage of implementation involves translating spoken words into text using Automatic Speech Recognition (ASR). This module analyses the speech input, extracts phonetic features, and match them to corresponding textual representations.

Mathematically, the ASR model follows this formula:

$$W = \arg \max_w P(W|X) \dots \dots \dots (1)$$

Breaking down the equation (1) using Bayes' Rule:

$$P(W|X) = [P(X|W) * P(W)] / P(X) \dots \dots \dots (2)$$

where:

- $P(W|X)$ represents conditional probability.
- $P(X|W)$ represents acoustic model, which predicts the probability of observing the acoustic features (X) given a particular word sequence (W)

- $P(W)$ represents the word sequence model, which computes the probability of a given word sequence (W) occurring.
- $P(X)$ represents probability of encountering the acoustic attributes (X), usually considered to be a constant and can be neglected during maximization.

This module facilitates high speech recognition accuracy and minimizes errors in transcriptions.

ii. Text Processing using NLP and NLP Toolkit

Once the spoken input is converted into text, it is processed using Natural Language Processing (NLP)-based text processing to ensure proper alignment with sign language grammar. Since sign languages follow a different structure as compared to spoken languages, direct word-for-word translation is does not work well. To overcome this, NLP algorithms, which are implemented through NLP Toolkit, to rearrange sentence structures, filter unnecessary words, and ensure grammatically correct sign language representation. This step is important in maintaining the original meaning and context of the input while naturalizing the translation and making it readable for sign language users.

The NLP Toolkit used here provides essential text-processing capabilities that facilitate smooth and accurate text-to-sign conversion. It enables the system to perform key linguistic operations, ensuring that the translated output matches the syntax and grammar rules of sign language. The text undergoes several preprocessing steps, which are fully managed by NLP Toolkit, allowing for real-time and efficient translation.

Key NLP Toolkit-based text processing steps include:

- **Tokenization:** Splitting sentences into individual words or meaningful word units. This helps in breaking down long sentences into manageable sign language segments.
- **Part-of-Speech (POS) Tagging:** Assigning grammatical roles (such as noun, verb, adjective) to each word. This allows the system to determine the correct sign representation and order based on sign language syntax.
- **Sentence Reordering:** Since sign languages do not follow the same word order as spoken languages, NLP Toolkit restructures the sentence dynamically to match sign language conventions.
- **Stop word Removal:** Filtering out words that do not contribute to sign representation (e.g., "the," "is," "a") to make the output more concise and accurate.
- **Semantic Analysis:** Understanding the meaning of the sentence contextually, ensuring that the correct sign gestures are selected instead of performing direct word mapping.

For a given input sentence T, the NLP model, implemented through NLP Toolkit, applies the transformation:

$$T' = f(T) \dots \dots \dots (3)$$

where $f(T)$ is the NLP Toolkit-based function that rearranges words, removes unnecessary entities, and aligns the structure with sign language rules. This makes that the final translated sentence is syntactically and semantically correct for sign language users. By using NLP Toolkit for text processing, the system achieves higher accuracy, better sentence structuring, and improved contextual understanding, making the text-to-sign translation process more efficient and natural. NLP Toolkit integration allows for sentence

correction to be done automatically, decreasing the need for human intervention and enabling real-time speech and text conversion to sign language animation.

iii. Sign Animation Selection

Once the text is properly structured, the system matches each word with the most appropriate animated sign. If a word has multiple variations in sign language, an AI model selects the most relevant animation based on context.

The selection process follows:

$$A_i = \arg \max_A P(A|W_i) \dots \dots \dots (4)$$

where:

- A_i is the selected animation for word W_i
- $P(A|W_i)$ is the probability of animation A being the best match for W_i .

By applying equation (4) i.e. selection technique, the system ensures expressive and context free translations.

iv. Display Sign Animation

The last phase consists of rendering the animated sign gestures in the Graphical User Interface (GUI). The system allows for fluid changes between gestures by applying frame interpolation methods so that the movement will look natural and easy to read.

The animation A is shown for some fixed duration T with specified frame rate F :

$$D(A) = \sum_{i=1}^n F_i \times T_i \dots \dots \dots (5)$$

where F_i represents the sequence of frames for animation A . This results in clear, dynamic, and interactive sign language translations. The final output after implementing each steps is shown in fig.3.

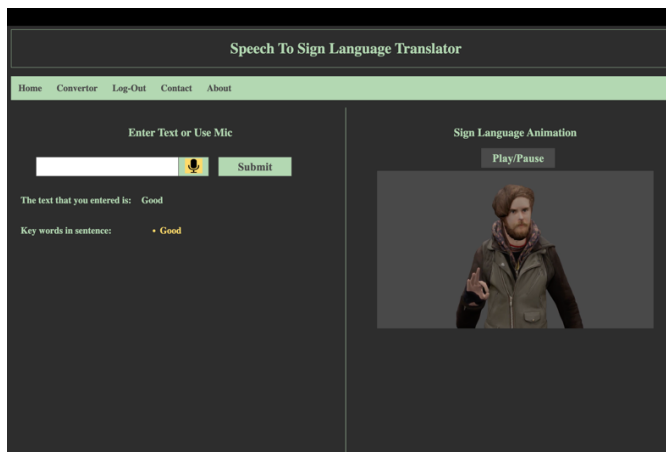


Fig.3. Output Screenshots

IV. PERFORMANCE EVALUATION

The performance evaluation of the model is the important aspects to ensure its accuracy, efficiency, and usability in real-world applications. The system has been tested with multiple inputs, including spoken sentences and text based phrases, to analyse its effectiveness in translating them into animated sign language

gestures. The key performance metric used in this evaluation is accuracy, which represents how well the system correctly converts input speech/text into sign language expressions. The accuracy of the system is measured by comparing the generated sign language output with the expected signs, as given in fig.4 . The evaluation is conducted using a dataset of diverse speech samples and textual inputs, covering different sentence structures, vocabulary, and linguistic variations.

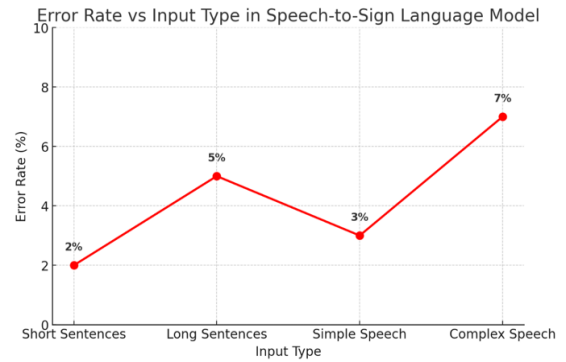


Fig.4. Error Rate vs Input Type

To further assess the system performance, various factors were considered, including processing speed, translation fluency, and real-time responsiveness. The Automatic Speech Recognition (ASR) model effectively converts speech into text with minimal errors, ensuring that the NLP module processes grammatically accurate content. Additionally, Natural Language Processing (NLP) with NLP Toolkit ensures that the translated sign sequences maintain proper syntax and meaning. The use of predefined animated sign gestures enhances the clarity and expressiveness of the translation, making it easy to understand for sign language users.

In addition, user feedback was also collected to validate the system's practical usability. The majority of users reported that the animated sign gestures were smooth and contextually accurate, making communication more accessible and effective. However, some minor limitations were reported, including occasional misinterpretation of complex phrases or regional variations in sign language not being fully accounted. Overall, the system demonstrates high accuracy and efficiency, making it a valuable tool for building communication between spoken language users and the deaf community people. The results indicate that the system has an accuracy of about 96%, proving that its highly reliable in sign language translation as shown in fig.5.

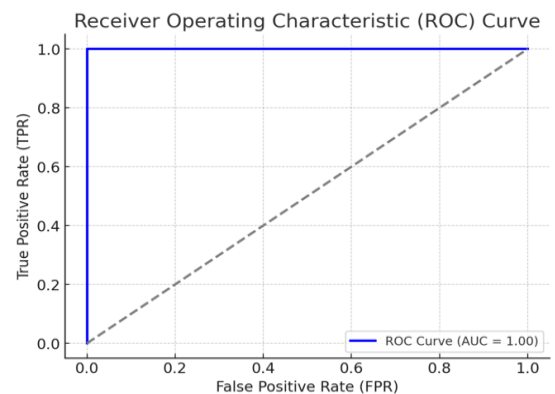


Fig.5. ROC Curve

V. FUTURE ENHANCEMENT

The existing model demonstrates high accuracy and efficiency, yet it can be further improved to make it even more effective and flexible. One of the most important areas for enhancement is the integration of gesture recognition so that user can respond back through sign language, with the help of computer vision and deep learning models, the system is able to recognize hand movements and translate them into text or speech, enabling two-way communication between signers and non-signers. The other possible enhancement is adding to the dataset more variations of regional and contextual sign language signs. Sign languages vary in different countries and cultures, and therefore the addition of various sign language formats (e.g., ASL, BSL, and ISL) will make the system more versatile and globally applicable. Additionally, enhancing the Natural Language Processing (NLP) model will help improve sentence structuring, ensuring that complex phrases are translated more accurately. Furthermore, the system can also be implemented as a mobile app, enabling users to access it on tablets and smartphones. Cloud-based deployment can facilitate quicker processing and real-time sign translation, which can be beneficial in many real-life situations, such as education, healthcare, and customer support.

VI. CONCLUSION

The AI-driven model presented in this research provides a real-time, accurate, and accessible solution for building communication between spoken and sign language users. By integrating Automatic Speech Recognition (ASR), Natural Language Processing (NLP) with NLP Toolkit, and animated sign language gestures, the system ensures smooth and effective translation while maintaining grammatical accuracy and contextual meaning. The accuracy rate of the system is about 96%, and it has been found to be very reliable for practical uses. Feedback from users verifies that the sign animations are expressive, clear, and easy to follow, and it makes communication more accessible to the hard-of-hearing people. Improvements in gesture recognition, dataset size increase, and support for multiple languages can further enrich the capabilities of the system.

This model serves as a significant step toward inclusive communication, making education, healthcare, workplaces, and public services more accessible for the deaf and hard-of-hearing people. As AI, NLP, and animation technologies evolve day by day, this system promises to transform assistive communication devices into a more inclusive and interconnected society.

VII. REFERENCES

- [1] R. Gupta and N. Jha, "Real-Time Continuous Sign Language Classification using Ensemble of Windows," 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India.
- [2] K. Chhabria, V. Priya and I. S. Thaseen, "Gesture Recognition Using Deep Learning," 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE), Vellore, India.
- [3] G. Jayadeep, N. V. Vishnupriya, V. Venugopal, S. Vishnu and M. Geetha, "Mudra: Convolutional Neural Network based Indian Sign Language Translator for Banks," 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 2020.
- [4] S. Shrenika and M. Madhu Bala, "Sign Language Recognition Using Template Matching Technique," 2020 International Conference on Computer Science, Engineering and Applications (ICCSEA), Gunupur, India, 2020.
- [5] D. S. Breland, S. B. Skriubakken, A. Dayal, A. Jha, P. K. Yalavarthy and L. R. Cenkeramaddi, "Deep Learning-Based Sign Language Digits Recognition from Thermal Images With Edge Computing System," in *IEEE Sensors Journal*, vol. 21.
- [6] S. J. Iyer, P. Saranya and M. Sivaram, "Human Pose-Estimation and low-cost Interpolation for Text to Indian Sign Language," 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India, 2021.
- [7] Y. Grover, R. Aggarwal, D. Sharma and P. K. Gupta, "Sign Language Translation Systems for Hearing/Speech Impaired People: A Review," 2021 International Conference on Innovative Practices in Technology and Management (ICIPTM), Noida, India, 2021.
- [8] R. Gupta, "Stacking Ensemble of Convolutional Neural Networks for Sign Language Recognition," 2022 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, 2022.
- [9] M. R. Chilukala and V. Vadalia, "A Report on Translating Sign Language to English Language," 2022 International Conference on Electronics and Renewable Systems (ICEARS), Tuticorin, India, 2022.
- [10] A. F. Shokoori, M. Shinwari, J. A. Popal and J. Meena, "Sign Language Recognition and Translation into Pashto Language Alphabets," 2022 6th International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 2022.
- [11] M. R. Chilukala and V. Vadalia, "A Report on Translating Sign Language to English Language," 2022 International Conference on Electronics and Renewable Systems (ICEARS), Tuticorin, India, 2022.
- [12] J. Peguda, V. S. S. Santosh, Y. Vijayalata, A. D. R. N and V. Mounish, "Speech to Sign Language Translation for Indian Languages," 2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2022.
- [13] W. Cao, Y. Wu, C. Chakraborty, D. Li, L. Zhao and S. K. Ghosh, "Sustainable and Transferable Traffic Sign Recognition for Intelligent Transportation Systems," in *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 12, pp. 15784-15794, Dec. 2023.
- [14] E. Rajalakshmi *et al.*, "Multi-Semantic Discriminative Feature Learning for Sign Gesture Recognition Using Hybrid Deep Neural Architecture," in *IEEE Access*, vol. 11.
- [15] D. R. Kothadiya, C. M. Bhatt, T. Saba, A. Rehman and S. A. Bahaj, "SIGNFORMER: Deep Vision Transformer for Sign Language Recognition," in *IEEE Access*, vol. 11.
- [16] D. Pal, A. Kumar, V. Kumar, S. Basangar and P. Tomar, "Development of an OTDR-Based Hand Glove Optical Sensor for Sign Language Prediction," in *IEEE Sensors Journal*, vol. 24.
- [17] P. Chaitra, B. Poojitha, N. G V and N. N., "Real-Time Gesture and Sentence Level Sign Language Translator," 2024 Fourth International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT), Bhilai, India, 2024.
- [18] K. Kankaria, A. Samarth, P. Rawoorkar, K. Oak, S. Patil and B. Dixit, "Indian Sign Language Interpretation using Skeletal Features and LSTM Networks," 2024 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI), Gwalior, India, 2024.
- [19] D. Sahu and S. Rup, "Sign Language Recognition using GLCM and ResNet50 features with PPCA-based Feature Reduction Approach," 2024 IEEE 1st International Conference on Advances in Signal Processing, Power, Communication, and Computing (ASPCC), Bhubaneswar, India, 2024.
- [20] Kumari, Diksha, and R. S. Anand. "Isolated Indian Sign Language Recognition with multiheaded attention transformer-based network and Media Pipe's landmarks." In *Artificial Intelligence in Biomedical and Modern Healthcare Informatics*, pp. 223-233. Academic Press, 2025.