

Audio to ISL Translator

B S Anil Kumar

Department of CSE
Gokaraju Rangaraju Institute of
Engineering and Technology
Hyderabad, India – 500060

N Raghavi

Department of CSE
Gokaraju Rangaraju Institute of
Engineering and Technology
Hyderabad, India – 500060

P Yochana

Department of CSE
Gokaraju Rangaraju Institute of
Engineering and Technology
Hyderabad, India – 500060

P Sai Ruthvika

Department of CSE
Gokaraju Rangaraju Institute of
Engineering and Technology
Hyderabad, India – 500060

D Anitha

Department of CSE
Gokaraju Rangaraju Institute of
Engineering and Technology
Hyderabad, India – 500060

Abstract-Interaction between hearing-impaired and normally hearing people creates a major impediment in day-to-day communication. To overcome the communication barrier, we introduce a Sign Language Translation System that ensures real-time bidirectional interaction between hearing-impaired and non-hearing-impaired people. The system takes advantage of state-of-the-art speech recognition, computer vision, and a rich database of sign languages to convert verbal language to sign language and sign language to verbal language.

Here, the process of communication starts when a typical individual speaks into a microphone. The voice is recorded and processed by the Sign Language Translation Application using natural language processing (NLP) as well as speech recognition algorithms to transform the spoken words into text. This text is then synchronized with corresponding sign language representations—images or GIFs—deducted from a special database. These sign language images are shown on a screen so that the hearing-impaired individual can interpret the oral message.

On the other hand, when an individual who is hearing-impaired wants to talk, their gesture is recorded real-time by way of a camera that is also linked to the same translation app. Through application of computer vision and gesture recognition technologies, the app recognizes what sign language has been used and translates it to equivalent textual or verbal language. This is transmitted to the non-impaired individual either by displaying text or uttering voice.

The system facilitates smooth, efficient, and real-time communication without the use of an interpreter and greatly increasing the inclusivity of the hearing-impaired community. Not only does it facilitate normal conversations, but it is also likely to have applications in education, customer service, healthcare, and beyond. This framework demonstrates the impressive convergence of audio processing, image recognition, and machine learning technology to build a more accessible society. With ongoing development and training on various sign languages and dialects, this solution can evolve into a globally usable communication tool for the hearing-impaired to ensure accessibility and equal opportunities.

Keywords— Indian Sign Language (ISL), translator, text, speech, two-way communication, Natural Language Processing (NLP), accuracy, Multiple Indian Languages.

I. INTRODUCTION

Indian Sign Language is the native language for millions of hearing-impaired individuals across India. It goes beyond a language for these people, who rather consider ISL as their door to thoughts, emotions, and expressions. But still, much has to be done regarding the importance of this language, which is largely unheard by the mainstream population, paving way for a considerable communication gap between hearing-impaired and non-hearing-impaired sections. This communication barrier indeed causes several expansions in their already wider restricted lives, which limit access to significant healthcare and educational facilities and job opportunities.

The absence of viable translation instruments between natural languages and Indian Sign Language (ISL) compels hearing-impaired users to seek the services of human interpreters, even for simple communication. This reliance presents important challenges, particularly in areas such as healthcare, education, and government services where interpreters may not be readily present. The system proposed in this work solves the problem by way of a real-time, two-way communication channel with the use of speech recognition, NLP, and computer vision. By interpreting verbal language into ISL and the reverse, the system enables deaf individuals to communicate on their own, mitigating social exclusion and facilitating full participation in day-to-day societal life.

None of these limitations stand in the way of inventive innovation in artificial intelligence, machine learning, and computer vision. These new technologies have created the best opportunity to open avenues that could lead to closing this communication gap. Such technological developments would provide new ways for communication to take place by allowing the interpretation and translation of ISL into spoken language and vice versa in real-time. By the development of speech-ISL-speech systems, the accessibility, inclusiveness, and efficiency of communication would be achieved in an environment that hearing-impaired people find in connecting

themselves into more resources of communication with others.

This project envisages the usage of both these technologies to develop a two-way communication system that seamlessly translates from spoken language to ISL and vice-versa. The project involves the development of the system in real-time such that instant translation is possible without the need to have any external human input. The system will thus include a mixture of computer vision and natural language processing (NLP) technologies, as the system will not simply identify the ISL signs but also understand the speech and render it into sign language so that the hearing-impaired can communicate effectively with others unaware of ISL.

By this system, empowerment and inclusivity for the hearing-impaired would be achieved. Being able to communicate without assistance from anyone in various environments—school, hospital, workplace, and public services—will make them much more integrated with society. Creating equal platform communications between the hearing-impaired and normal individuals would allow for more meaningful interactions, thus creating lesser stigma towards disability, greater social inclusion, and eventual.

II. LITERATURE REVIEW

The proposed article discusses the different methods of a certain communication system, where people use two different communication methods for their variable communication device and online learning systems like the keyboard method global system, the convenient nose touch screen i.e., that belong to the variable. communication method. All methods used different types of accelerometer sensor model, which was used to convert text to speech on the screen and keyboard to microcontroller, usually to enter a message between the deaf and non-deaf. Here, the external device was necessary, the external device made this method difficult to solve the problem, authors used e-learning as different methods

[1] Chevella Anil Kumar, V. Sagar Reddy, Nainika Kandarpa, and Pallavi Sharma (2024) recommend a glove-based system in their paper “ISL to Speech Converter: A Wearable Glove-based System for Bridging Communication Gaps” to convert Indian Sign Language (ISL) movement to speech. The system relies on an Arduino UNO microcontroller linked to pushbuttons to detect pre-programmed hand gestures, which are converted into sound output by a Python-based text-to-speech (TTS) module. Transfer of data from the glove to the processing unit is achieved via serial communication. While the system provides an affordable and mobile method of real-time ISL translation, it is limited by the use of raw pushbutton gestures and the hardware constraints in its scalability into more complex and sophisticated gestures. There is no common dataset employed; it relies on real-time gesture input via hardware sensors.

[2] It is an article by Ms. Pradnya Repal (2024) which discusses “Real Time Sign Language Translator Using Machine Learning” a web translation system that will translate sign language to text and sound and vice versa from sound to sign language by utilizing advanced machine

learning methods. This system utilizes technology such as TensorFlow, MediaPipe, Convolutional Neural Networks (CNNs) Recurrent Neural Networks (RNNs), and speech recognition APIs to achieve real-time, two-way communicative solutions. The ASL alphabet and number dataset were utilized in the proposed system for gesture training, recognition, and implementation. While the approach is promising, it recognized the challenges in gaining accuracy in gesture recognition and speech-to-text translation. The article also refers to the extent of improvement in precision and robustness for various conditions.

[3] Gautham Jayadeep, N.V. Vishnupriya, Vyshnavi Venugopal, S. Vishnu, and M. Geetha’s “ISL Hand Gesture Motion Translation Tool for Banks” (2020) is a tool to overcome the hearing gap between the hearing-impaired and bank officials by recognizing Indian Sign Language (ISL) gestures as executable instructions. The system uses state-of-the-art technologies such as Convolutional Neural Networks (CNN), Inception V3, Long Short-Term Memory (LSTM), and Recurrent Neural Networks (RNN) to identify hand gestures and map them to useful actions. Though the system is promising, it has certain limitations also, mainly because of the limited dataset on which it is trained and because of the challenges in processing gesture videos of different lengths. In addition, inconsistency in gesture can also influence overall accuracy, and the self-recorded dataset might not account for the whole range of gestures used in practical applications.

[4] Bandi R. Reddy, Darukumalli S. T. Reddy, Sandeep P. MC, and Susmitha Vekkot (2022) in their article titled “Speech to Sign Language Translation (Telugu) Using Deep Learning” research a speech to sign language translation deep learning model with reference to Telugu. The system utilizes GRU and Bi-LSTM models in processing the speech input as well as translating the input into corresponding sign language gestures in GIF format. Google APIs are utilized by the authors for improving speech recognition and then utilizing a graphical user interface (GUI) to facilitate the use of the system. Despite that, there are some restrictions in the system, i.e., it relies on a known set of GIFs, which is only limited for alphabetic characters of international languages. The system itself is also Telugu input specific, which restricts its use everywhere universally. Its Google API based dependency also introduces probable limitations from the reliability as well as scalability point of view. In spite of the above limitation, the paper presumes a workable speech-to-sign language translation, i.e., to the Telugu-speaking public.

[5] Shobana Devi P, Vidya V, and Balan C’s (2022) “GAN-Based Sign Language Translation” paper suggests an innovative method of translating sign language with the help of Generative Adversarial Networks (GANs) and Neural Machine Translation (NMT). The avatars are created to replicate sign language movements, translating text into sign language in real time. By utilizing GANs and NMT incorporation, the research attempts to demystify communication barriers among sign language users and non-

users. However, the paper also has some weaknesses, i.e., that the system is text- and audio/video transcription-based, thus its capability to encompass the whole array of dynamic sign language variations is constrained. The method also needs huge datasets in order to train optimally, something that may not be present every time. In spite of the above challenges, the paper introduces a new approach to enhancing sign language communication to be readily available, primarily in virtual media, but whose creation requires perfection and larger datasets for universal use.

III. METHODOLOGY

1. Sign-to-Text & Speech Conversion

MediaPipe Hands:

Applied for real-time hand landmark detection (21 keypoints per hand). Provides variable lighting and background robustness.

OpenCV:

Applied for video capture feed, preprocessing frames (resize, grayscale, thresholding, etc.). Assists in ROI extraction for improved gesture classification.

Convolutional Neural Network (CNN):

Specific CNN model trained on skeletal images obtained through MediaPipe. Classifies ISL alphabets from hand landmarks and gestures.

Buffering and Thresholding Logic:

Ensures stability by predicting a gesture only after a Consistent number of frames.

Text-to-Speech (TTS):

Python's pyttsx3 or gTTS used to convert recognized text into speech.

2. Speech/Text-to-Sign Conversion

Speech Recognition:

Uses the speech_recognition library for real-time voice input and transcription.

Sign Representation:

Maps each letter/word of the transcribed text to corresponding ISL sign images or GIFs. Displays the visual representation using frontend templates (HTML/CSS/JS).

Frontend-Backend Integration:

Flask handles backend processing (sign prediction, audio playback). Frontend updates sign images dynamically based on input text or voice.

Translation into Multiple Languages:

Before mapping the recognized text to signs, Google Translate is used to translate it into the user's preferred Indian language (such as Hindi, Telugu, or Tamil) using the deep-translator Python library.

IV. PROPOSED SYSTEM

The goal of this project is to develop a real-time translator that fills the Indian Sign Language (ISL) and spoken or written language gap. It's meant to facilitate effortless two-way communication—translating gestures to audio or text and vice versa. The core part of the system is a smart server that receives incoming inputs, identifies whether it's a gesture or voice command, and forwards it to the respective module for processing. For gesture recognition, the system uses

MediaPipe, OpenCV, and a CNN model to precisely identify hand movement from video input and translate it into readable text in an efficient and reliable manner.

At the end of the speech, the translator employs tools like Google Speech Recognition or Whisper to record what is being said and parse it into clean, translated text. The text may be output as sign language graphics or played back as sound in reverse through text-to-speech tools like gTTS or pyttsx3. To make the process more interactive and natural, the system offers gesture animations like sign GIFs or sentence-level graphics and shows the live translation on a web interface that is designed using Flask or React. Generally, this project aims to create a more universal and accessible way of having people convey speech and sign language in real-time.

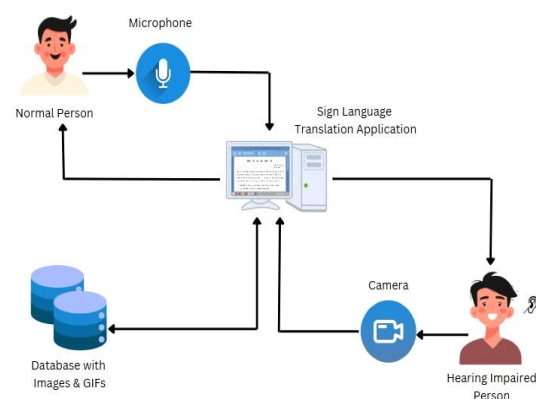


Fig 1. System Architecture

The System Architecture Diagram provides the architectural structure of the SignBuddy system. It presents how different modules are interlinked to enable real-time sign and speech translation. The system is segregated into two fundamental pipelines: Sign-to-Speech and Speech-to-Sign. Input devices such as the camera and microphone receive user gestures and voice commands, which are interpreted by the respective modules with libraries such as MediaPipe, OpenCV, and SpeechRecognition. The backend combines APIs, including Google Speech-to-Text and Text-to-Speech, and logic controllers translate gestures or speech into actions. The PostgreSQL database is used to provide efficient data management, and the final outputs, whether text, speech, or animations, are provided through the user interface to provide smooth, intuitive, and accessible communication.

Server Module:

An IoT-enabled microcontroller, such as the ESP32, collects data from all connected sensors. This device wirelessly transmits the gathered information to the Firebase Cloud via Wi-Fi, serving as the initial communication point within the system.

1. Sign Recognizer Module:

Captures and classifies hand gestures using MediaPipe and CNN to convert them into text.

2. Speech Processor Module:

Converts spoken input into clean, formatted text using speech recognition.

3. Text-to-Speech Module:

Transforms the recognized or typed text into audible speech output.

4. Sign Display Module:

Displays corresponding sign language animations or images for input text or speech.

5. Language Translation Module:

Translates recognized text into the user's preferred Indian language using Deep Translator.

6. Gesture Confirmation & Buffering Module:

Ensures stable predictions by buffering gestures and confirming input with a trigger gesture (like palm).

V. IMPLEMENTATION AND RESULTS

1) Flow chart for Audio to Sign Conversion

The flow chart shows the text to sign conversion as shown in Fig. 1. When the code is initialized, it processes and runs the following code then it shows that user has to give the text. After that the text will be displayed from the given word and it checks whether the given letter is matched with the database or not. If it matches with the database then the output will be displayed in the gesture form. Each letter has to get

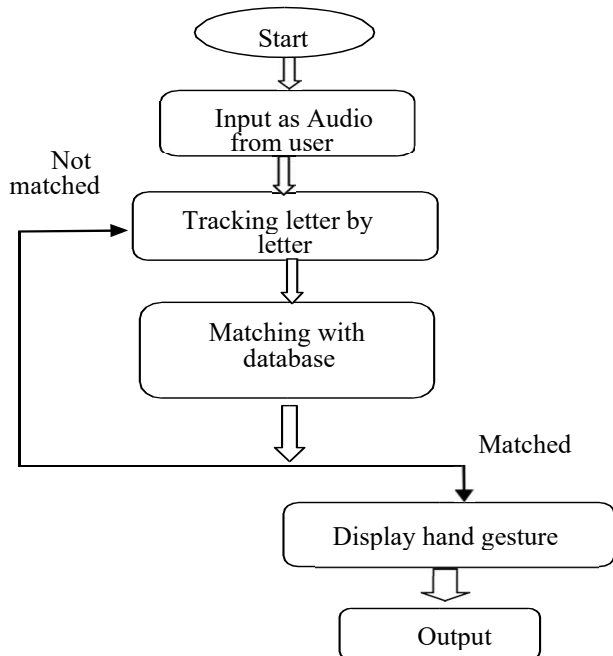


Fig1: Audio to Sign display

separated from the word and it displays each letter gesture. If it not matched with the database then it backs to the home page. This situation is not same for the following code. However, the code will able to give all the gestures according to the words given by the user. This flow chart shows the text to sign conversion. It depends on the user input and gif will

be displayed as the output. It can be sees when the code is executed. It can able to take only English alphabets rather than Telegu and Hindi.

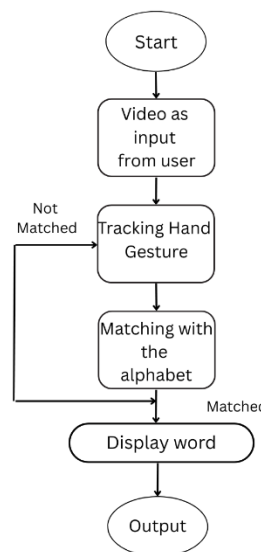


Fig 2. Sign to Audio conversion

In Fig. 2, the flow chart shows the sign to text conversion. Here the code is written in the text form. When the user places the gesture in the front of camera then the code is able to track the gesture of the user and check the gesture is matched with the database or not if it is not matched with the given database, it shows the error. If it is matched with the database, the output will be displayed in the form of text. Here some words are included like call me, thumbs up, thumbs down, fist and other words.



Fig 1. Home Page

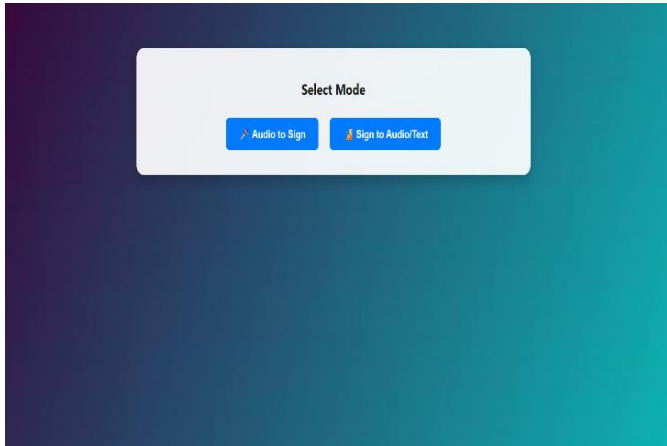


Fig 2. Main Page

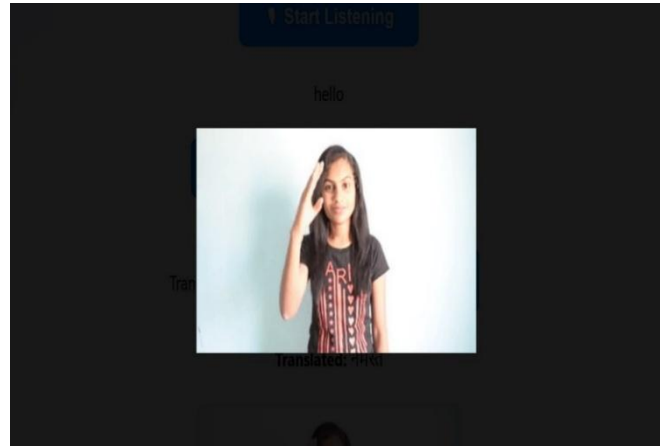


Fig 5. Sign GIF

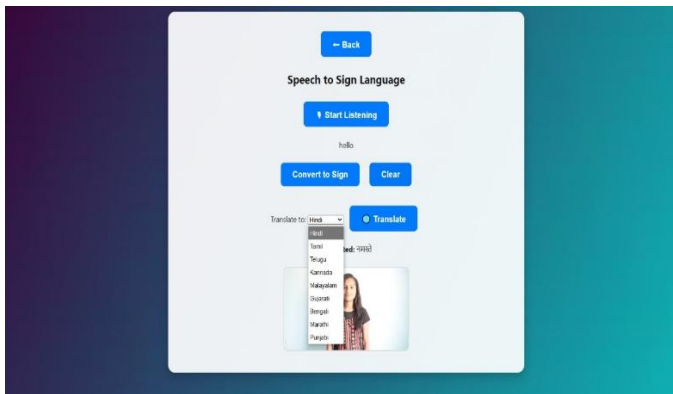


Fig 3. Audio to Sign

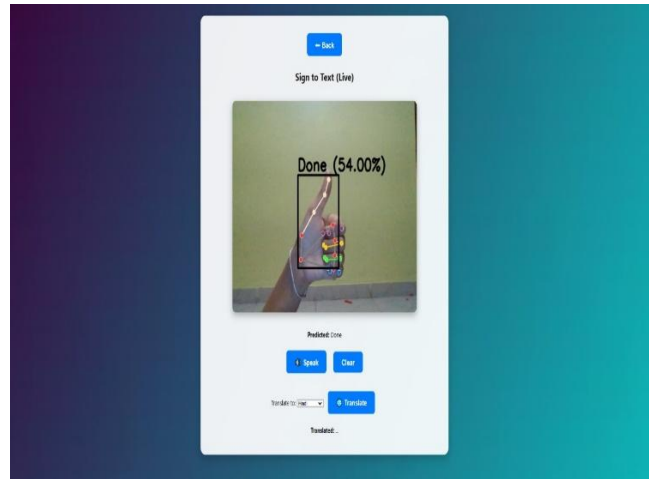


Fig 6. Sign To Audio

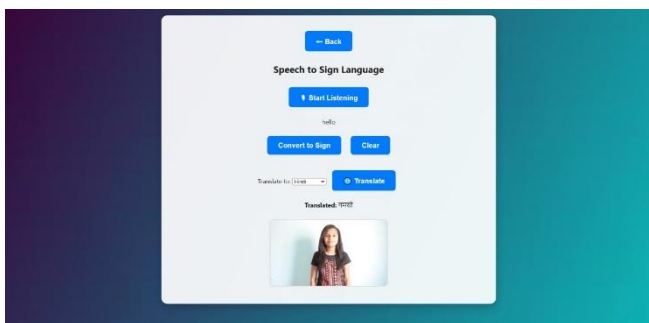


Fig 4. Sign GIF with the Language Translator

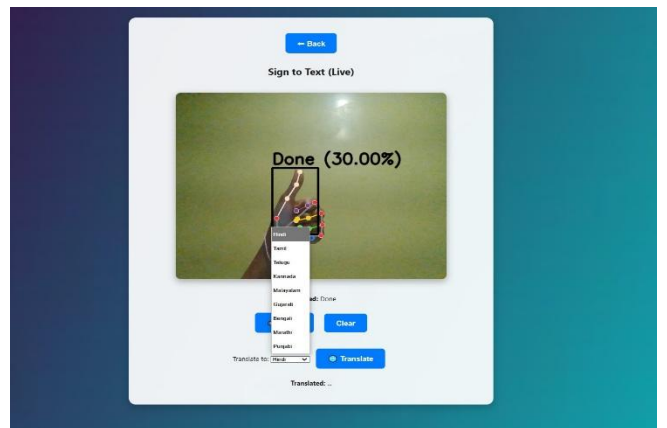


Fig 7. Audio Generated

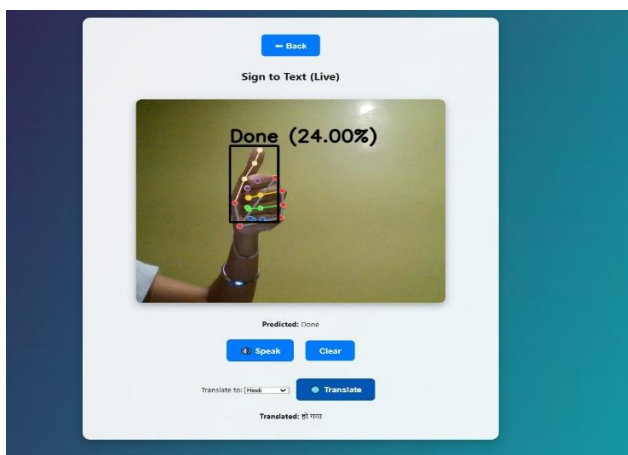


Fig 8. Sign to Text Language Translated

VI. SYSTEM TRAINING AND VALIDATION

The system includes training a CNN/LSTM model with MediaPipe-processed gesture data for sign recognition, verified through accuracy and confusion matrix over a held-out set. Speech input is processed by Google Speech Recognition API without any training. Identified text is mapped to ISL images/GIFs from a curated database. Validation constitutes cross-validation for gestures, WER for speech, and real-world testing to determine translation accuracy, response time, and user satisfaction. End-to-end testing verifies proper communication between hearing and hearing-impaired users in both directions through camera and microphone inputs.

TABLE I. TESTING

Table. 1. Audio input to Sign output

| Test Case | Input (Audio/Text) | Expected Output (ISL Signs Displayed) | Actual output | Final Verdict |
|-----------|---------------------|--------------------------------------------------------------|--------------------------------------------------------------|---------------|
| 1 | "Hello" | Signs for H → E → L → L → O | Signs for H → E → L → L → O | Pass |
| 2 | "I am" | Signs for I, A → M | Signs for I, A → M | Pass |
| 3 | "" (No input) | Error: "No input detected" | Error: "No input detected" | Fail |
| 4 | Background noise | Error: "Could not recognize speech" | Error: "Could not recognize speech" | Fail |
| 5 | "123" | Error: "Digits not supported" or ignored | Error: "Digits not supported" or ignored | Fail |
| 6 | "Thank you" | Signs for T → H → A → N → K, Y → O → U | Signs for T → H → A → N → K, Y → O → U | Pass |
| 7 | "Do you have money" | Signs for D → O, Y → O → U, H → A → V → E, M → O → N → E → Y | Signs for D → O, Y → O → U, H → A → V → E, M → O → N → E → Y | Pass |
| 8 | "Are You Hungry" | Signs for A → R → E, Y → O → U, H → U → N → G → R → Y | Signs for A → R → E, Y → O → U, H → U → N → G → R → Y | Pass |

Table. 2. Sign input to Text output

| Test Case | Input (Audio/Text) | Expected Output (Text Displayed) | Actual output | Final Verdict |
|-----------|-----------------------------|----------------------------------|---------------------------|---------------|
| 1 | Hand sign for "H" | "H" | "H" | Pass |
| 2 | Hand sign for "Hello" | "HELLO" | "HELLO" | Pass |
| 3 | Continuous signs for "I AM" | "I AM" | "I AM" | Pass |
| 4 | Sign for "My name is..." | "MY NAME IS" | "MY NAME IS" | Pass |
| 5 | Ambiguous hand movement | "Unrecognized" | "Unrecognized" | Fail |
| 6 | Hand sign + palm detected | Auto-capture next token | Auto-capture next token | Fail |
| 7 | Fast repeated gestures | Stabilized letter or word | Stabilized letter or word | Fail |
| 8 | Gesture for "PLEASE" | "Please" | "Please" | Pass |

VII. DISCUSSION

The Sign Language Translation System spans the communication gap between hearing-impaired and normal users through the combination of speech recognition, gesture recognition, and visual translation modules. For the recognition of gestures, a CNN or an LSTM model is trained on MediaPipe to identify hand keypoints from video input. It learns from a labeled ISL gesture dataset and is tested via accuracy measurement and confusion matrix using a different validation set or cross-validation. Normal user speech is translated to text via the Google Speech Recognition API, which is pre-trained and doesn't involve local model training. Recognized text is then matched against equivalent ISL gestures via a carefully curated database of images or GIFs. System validation entails checking the precision of gesture recognition, assessing speech input via Word Error Rate (WER), and conducting real-time, end-to-end testing to validate accurate bidirectional translation. Real-world testing concentrates on user satisfaction, responsiveness of the system, and translation clarity. Such an integrated approach makes sure that the system is technically sound but also practical and effective for real-time communication among differently-abled persons.

VIII. LIMITATIONS

Even though effective, the system is subject to various limitations. Accuracy in recognizing gestures can deteriorate under conditions of low illumination or cluttered background, with MediaPipe struggling to accurately locate hand landmarks. The performance of the model is also likely to decrease when exposed to a minimal or imbalanced dataset, weakening generalization towards real-world use cases. Recognition of speech using Google API will be compromised when exposed to loud noise or extensive accents, rendering the system less reliable. Further, the image/GIF database is based on mappings that have to be predefined, which restricts flexibility for dynamic or complicated sentences. Real-time processing can cause latency based on hardware or network speed. Lastly, the system could be challenged with regional differences in Indian Sign Language or lack contextual knowledge, resulting in translation errors. These issues need to be overcome in order for the system to scale and be used in a mass population.

IX. CONCLUSION

System proposed will have a simple and generalized way to fill a gap between hearing-impaired and non-hearing-impaired people to talk. It enables bi-directional real-time translation from Indian Sign Language (ISL) to spoken language and vice versa. The solution makes a huge difference in improving accessibility. The hearing-impaired people can, by using technology-mediated interaction with the system, be included in every other aspect such as education, public services, and health care.

With the help of the Translator class, which combines computer vision technology, speech recognition, and language translation, hand gestures are captured by the program and either converted into ISL representation or verbal output, while also supporting multiple Indian languages. Hence, the need for human interpretation is avoided in day-to-day communication and provides independence to the ISL-users. The first implementation had a very positive response in terms of really accurate results coming from both speech processing and gesture recognition, paving the way for future improvement.

X. REFERENCES

- [1] N. Dhruva, S. R. Rupanagudi, S. K. Sachin, B. Sthuthi, and R. Pavithra, "Novel segmentation algorithm for hand gesture recognition," In *2013 International Multi-Conference on Automation, Computing, Communication, Control and Compressed Sensing (iMac4s)*, pp. 383-388, 2013.
- [2] A. Y. Dawod, J. Abdullah, and Md. J. Alam, "Adaptive skin color model for hand segmentation," in *2010 International Conference on Computer Applications and Industrial Electronics*, pp. 1-4, 2011.
- [3] V. Bhame, R. Sreemathy, and H. Dhumal, "Vision based hand gesture recognition using eccentric approach for human computer interaction," in *2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pp. 949953, 2014.
- [4] A. S. Ghotkar, and G. K. Kharate, "Study of vision-based hand gesture recognition using Indian sign language," *International journal on smart sensing and intelligent systems*, vol. 7, no. 1, pp.96-115, 2014.
- [5] Y. Zhu, Z. Yang, and B. Yuan, "Vision based hand gesture recognition," in *2013 international conference on service sciences (ICSS)*, pp. 260-265, 2013.
- [6] S. N. Sawant, "Sign language recognition system to aid deaf-dumb people using PCA," *Int. J. Comput. Sci. Eng. Technol. (IJCSSET)*, vol. 5, no. 05, 2014.
- [7] M. J. Taylor, and T. Morris, "Adaptive skin segmentation via feature-based face detection," *Real-Time Image and Video Processing*, vol. 9139, pp. 167-178, 2014.
- [8] Y. Madhuri, G. Anitha, and M. Anburajan, "Vision-based sign language translation device," in *2013 International Conference on Information Communication and Embedded Systems (ICICES)*, pp. 565-568, 2013.
- [9] M. Elmahgiubi, M. Ennajar, N. Drawil, and M. S. Elbuni, "Sign language translator and gesture recognition," in *2015 Global Summit on Computer & Information Technology (GSCIT)*, pp. 16, 2015.
- [10] I. Dagher, N. E. Sallak, and H. Hazim, "Face recognition using the most representative sift images," *International Journal of Signal Processing, Image Processing and Pattern Recognition*, vol. 7, no. 1, pp.225-2363, 2014.
- [11] N. C. Camgoz, S. Hadfield, O. Koller, H. No and R. Bowden, "Neural translation of sign language," in *Proceedings of the IEEE Conference on Computer Science Vision and Pattern Recognition (CVPR)*, pp. 7784– 7793, 2018.
- [12] J. Huang, W. Zhou, Q. Zhang, H. Li, and W. Li, "Visual-based sign language recognition without temporal segmentation," in *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI-18)*, vol. 32, no. 1, 2018.
- [13] S. K. Ko, C. J. Kim, H. Jung, and C. Cho, "Neural sign language translation based on human key point estimation," *Applied sciences*, vol. 9, no. 13, p. 2683, 2019.
- [14] D. Guo, W. Zhou, H. Li, and M. Wang, "Hierarchical LSTM for sign language translation," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 32, no. 1, 2018.
- [15] O. Koller, N. C. Camgoz, H. Ney, and R. Bowden, "Weakly supervised learning with multi-stream CNN-LSTM-HMMs to discover sequential parallelism in sign language videos," *IEEE transactions on pattern analysis and machine intelligence*, vol. 42, no. 9, pp. 2306-2320, 2019.
- [16] B. Sridhar, "A Blind Image Watermarking Technique using most Frequent Wavelet Coefficients," *International Journal on Smart Sensing & Intelligent Systems*, vol. 10, no. 4, 2017.
- [17] P. Yakaiah, and K. Naveen, "An Approach for Ultrasound Image Enhancement Using Deep Convolutional Neural Network," in *Advanced Techniques for IoT Applications: Proceedings of EAIT 2020*, pp. 86-92, 2022.
- [18] M. Alsulaiman, M. Faisal, M. Mekhtiche, M. Bencherif, T. Alrayes, G. Muhammad, H. Mathkour, W. Abdul, Y. Alohali, M. Alqahtani, and H. Al-Habib, "Facilitating the communication with deaf people: Building a largest Saudi sign language dataset," *Journal of King Saud University Computer and Information Sciences*, vol. 35, no. 8, p.101642, 2023.
- [19] B. Hdioud, and M. E. H. Tirari, "A Deep Learning based Approach for Recognition of Arabic Sign Language Letters." *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 4, 2023.
- [20] A. Dixit, S. Sharma, P. D. Rao, V. Reddy, M. Janaki, R. Thirumalaivasan, and M. M. Subashini, "Audio to indian and american sign language converter using machine translation and nlp technique," in *2022 Third International Conference on Intelligent Computing Instrumentation and Control Technologies (ICICT)*, pp. 874-879, 2022.
- [21] A. Chaikaew, "An applied holistic landmark with deep learning for Thai sign language recognition," in *2022 37th International Technical Conference on Circuits/Systems, Computers and Communications (ITCCSCC)*, pp. 1046-1049, 2022.