Journal of Artificial Intelligence and Computational Technology

Journal homepage: https://ojs.omgfzc.com/index.php/JECT

Hand Gesture Recognition for Sign Language Translation

Jithu Varghese Jacob[1], Anwin K Biju[2], Dr. Rajesh Kanna R[3]

Anwer Mustafa Hilal[4], Kawthar Ishag Ali Fadlallah[5]

[1]Student, Master of Computer Applications, CHRIST University, Bangalore
[2]Student, Master of Computer Applications, CHRIST University, Bangalore
[3]Assistan Professor, Department of Computer Science, CHRIST University, Bengaluru, Karnataka, India, 560029
[4]Department of Information System, Omdurman Islamic University, Omdurman, Sudan
[5]Department of Computer Science, Omdurman Islamic University, Omdurman, Sudan

**Correspondence: E-mail: alwin.jaison@mca.christuniversity.in

Article Info

Article History:
Submitted/Received 5 February 2025
Revised in revised format 25 February 2025
Accepted 17 Mar 2025
Available 17 Mar 2025
Publication Date 01 Apr 2025

Keyword:
Indian Sign Language,
Deep Learning,
Gesture Recognition,

Cite this article:

ABSTRACT

This project aims to develop software that translates Indian Sign Language (ISL) hand gestures into real-time text and speech. A custom symbol module enables users to add new gestures, enhancing user experience. Voice call integration allows real-time gesture-to-speech translation during calls. Optimized machine learning models ensure efficiency and accessibility, creating a solution for inclusive communication. The domain for hand gesture recognition in sign language translation, where the development of automated systems for the interpretation of sign languages using gestures would make fluid communication possible with deaf or hard-of-hearing patients. The recognition of hand gestures relies on the combination of algorithms in image processing and machine learning to capture, understand, and translate signs into text or spoken language. Over time, these systems have evolved from the traditional rule-based and statistical methods to more advanced models that focus on achieving greater accuracy and reliable recognition through machine learning, especially deep learning.

1. Introduction

The main mode of communication among the Deaf and Hard of Hearing is sign language. The translation of sign language into text or speech remains quite limited in many applications, such as real-time voice calls. The proposed project will create a mobile application to translate Indian Sign Language gestures into text and speech in real time, thereby bridging the gap between sign language users and non-sign language users. This project allows users to personalize the application using new symbols for words that are not included in the standard dataset. Recording short videos of new gestures makes the application automatically ex- tract key frames; this makes the multiple image capture process automatic, thus making it smooth and friendlier to use. Optimized mobile device models applied to machine learning will have a flexible framework that may enable the predefined user vocabulary. Other than gesture translation, the application will offer voice calls and convert sign language into synthesized speech in future update. This will make it easier for a person with sign language to communicate real-time in two ways with someone with spoken language. This allows interactions to be more accessible and inclusive. They are using TensorFlow Lite for gesture recognition, signalizing the state of communication and WebRTC to make voice calls feel natural in sign language conversations. The technology stack, system architecture, and implementation strategy for an innovative tool that would completely eliminate the existing limitations of translation of sign language to speech, improve user experience, and reduce computational costs.

2. Related Works

Ebey Abraham et al [1]. It designed the system in order to bridge the gaps of communication between deaf and others through making ISL into speech translation. For this system, the hand gestures were detected through a sensor glove containing flex sensors and IMU. Classified data have been collected which were classified through LSTM networks, which are efficient learners of long-term dependencies. It may classify 26 gestures with an accuracy of 98%. Thus, this proves that LSTM based neural networks have real-time capability for the translation of ISL.

Harini R et al [2]. proposed computer vision and machine learning work that translates sign language into text. It captures a sign gesture using a webcam. Subsequently, it preprocesses the sign through background subtraction, while classification is done through a CNN. The model does pretty well with high accuracy at 99.91%. No equipment such as gloves and sensors is needed, but it relies on computer vision to understand the gestures.

Mohammed Elmahgiubi et al [3]. proposed a wearable device in the form of a smart glove, equipped with sensors, that translates sign language gestures to readable text. The glove, equipped with flex sensors, contact sensors, and accelerometer, captures hand gestures to send data for interpretation to the microcontroller. The textual output might be displayed at an LCD screen or forwarded to a smartphone. It recognizes 20 out of 26 letters of American Sign Language with 96% accuracy. It is simple and compact and cost-effective in design.

Zeyu Liang et al [4]. summarizes a vast array of current sign language translation approaches along with their techniques. Ch. outlines challenges in- volved together with advances in converting it to spoken or written out- put via machine learning. This work splits SLT into four significant tasks: Sign2Gloss2Text, Sign2Text, Sign2(Gloss+Text), and Gloss2Text. More so, the paper outlines transformer-based architectures applied in SLT and discusses some of the major challenges such brings: scarcity of data and complexity in grammar of sign languages. Lastly, the survey is summed up by proposing directions of future work to make improvements on the SLT model (Sign Language Translation).

Necati Cihan Camgoz et al [5]. proposed Neural SLT as a new distinct task from SLR. Therefore, the objective of SLT is the direct translation of videos from sign language into spoken language text directly,

DOI: https://doi.org/10.70274/jaict.2025.2.1.
p- ISSN 34562-3478 e- ISSN 5443-1243

capturing the difference in the grammar and word orders of both languages. Authors propose an encoder-decoder model using convolutional and recurrent neural networks with attention mechanisms. They contribute a continuous SLT data-set called RWTH PHOENIX-Weather 2014T for the German Sign Language, and state baseline results for several architectures of SLT.

Mary Jane C. Samonte et al [6]. looks at deep learning models of translation from sign language into text, aimed at increasing accessibility in communication in cases where the concerned person may either be speech and hearing impaired. Here, the importance of CNNs, CTC, and DBN concerning text production when interpreting hand gestures is brought to attention. A deep learning based proposed model was developed with respect to careful review of all relevant studies for the precise recognition of signs and language processing. Such enhanced systems of translation would minimize communication barriers as well as barriers of accessibility by the deaf and hard-of-hearing community.

Adria´n Nu´ n˜ ez-Marcos et al [7]. offered a holistic view of methods, challenges, and advances in sign language translation. Traditional approaches include rule based methods and statistical machine translation. Recently, there also appeared a great advance, using neural models based on deep learning techniques, referred to as NSLT; a good quality of such robust datasets, like the mentioned RWTH-PHOENIX-2014T, is also one more requirement in training its suitable machine learning model using correct results. In this survey, it can be seen how gesture segmentation, gloss translation, and regional sign language variation will be further advanced to support smooth translation with precision across languages.

3. Discussion

3.1. Project Methodology and Requirements

A. Tools and Technologies

MediaPipe: MediaPipe is a highly optimized library for real-time hand tracking and gesture detection. It gives you the ability to accurately track hand and finger movements by providing precise hand landmark detection. It would be a great benefit for any ISL (Indian Sign Language) applications, since the scenerio of detecting gestures on both Mobile and Desktop is persistent.

TensorFlow Lite: Tensor Flow Lite is used for deployment of custom trained model to recognize gestures on mobile. TensorFlow Lite framework provides an optimized runtime for machine learning models that is extremely useful in keeping performance and reducing latency on resource-constrained mobile devices.

OpenCV: It is used to add new gestures in capturing and processing video frames.5 OpenCV can automatically take frames from a video that saves the user considerable work when adding custom gestures as he need not capture several still images.

Google TTS API & pyttsx3: This API offers the facility for on-cloud speech synthesis. It converts ISL gestures recognized to live audible speech. Good performance and high-quality output that would be perfect for developing an ISL accessible application is maintained with multilingual support, making the output surprisingly very natural.

B. Model Training

Utilize TensorFlow and Keras for model development.

Employ transfer learning and data augmentation to improve robustness.

DOI: https://doi.org/10.70274/jaict.2025.2.1.
p- ISSN 34562-3478 e- ISSN 5443-1243

Optimize for mobile environments using TensorFlow Lite.

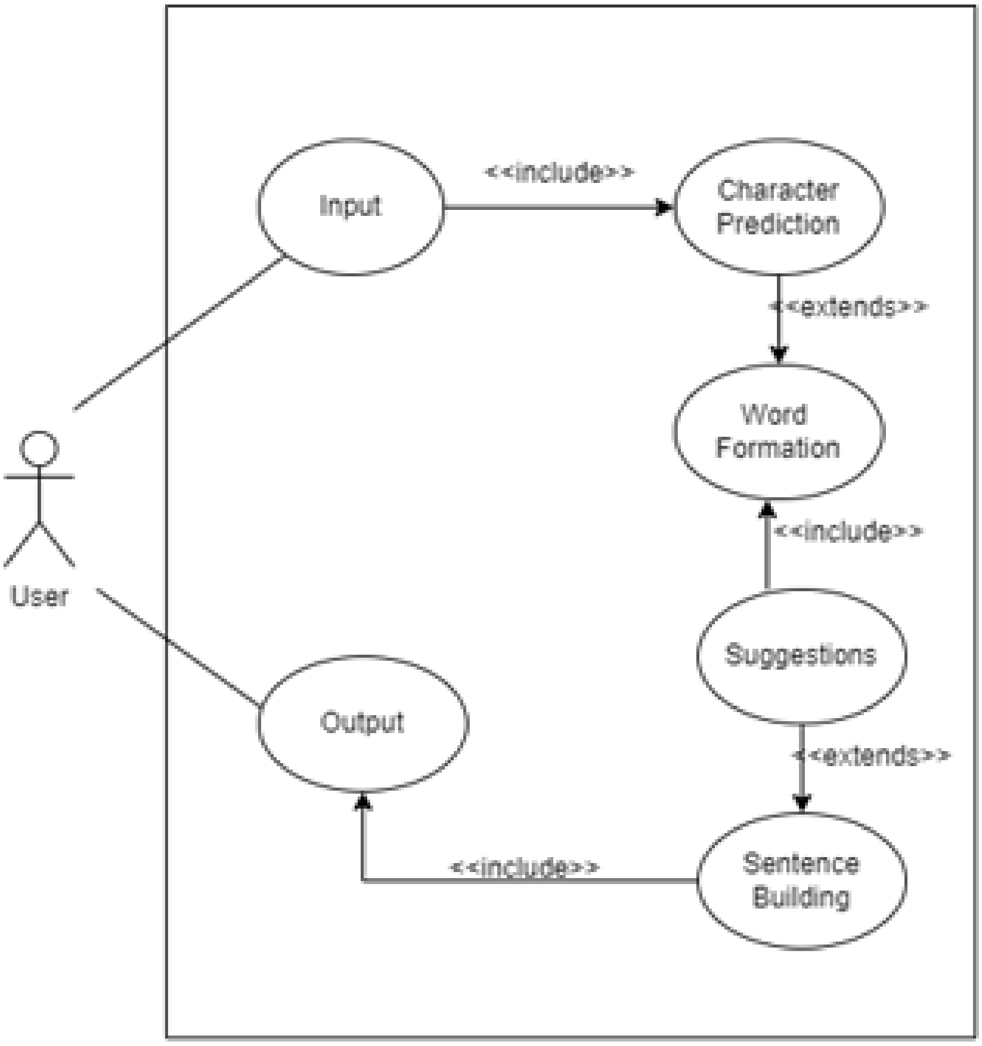

Fig. 1 USE-CASE DIAGRAM

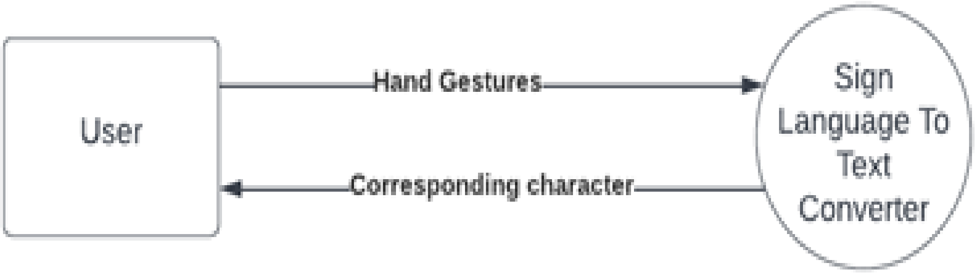

Fig. 2 Data Flow Diagram- Level-0

DOI: https://doi.org/10.70274/jaict.2025.2.1.
p- ISSN 34562-3478 e- ISSN 5443-1243

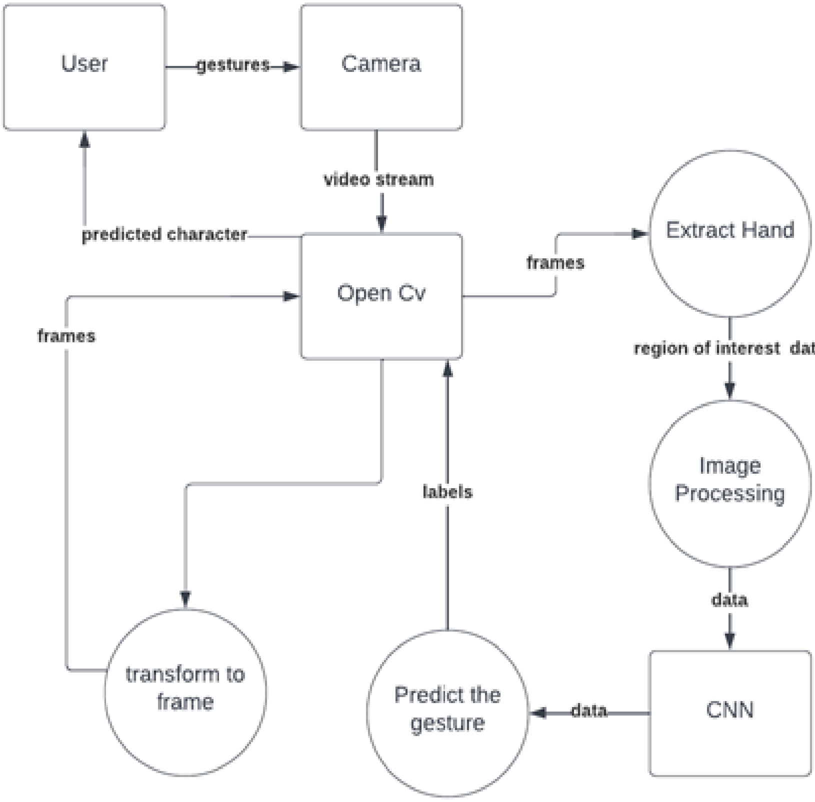

Fig 3 Data Flow Diagram - Level 1

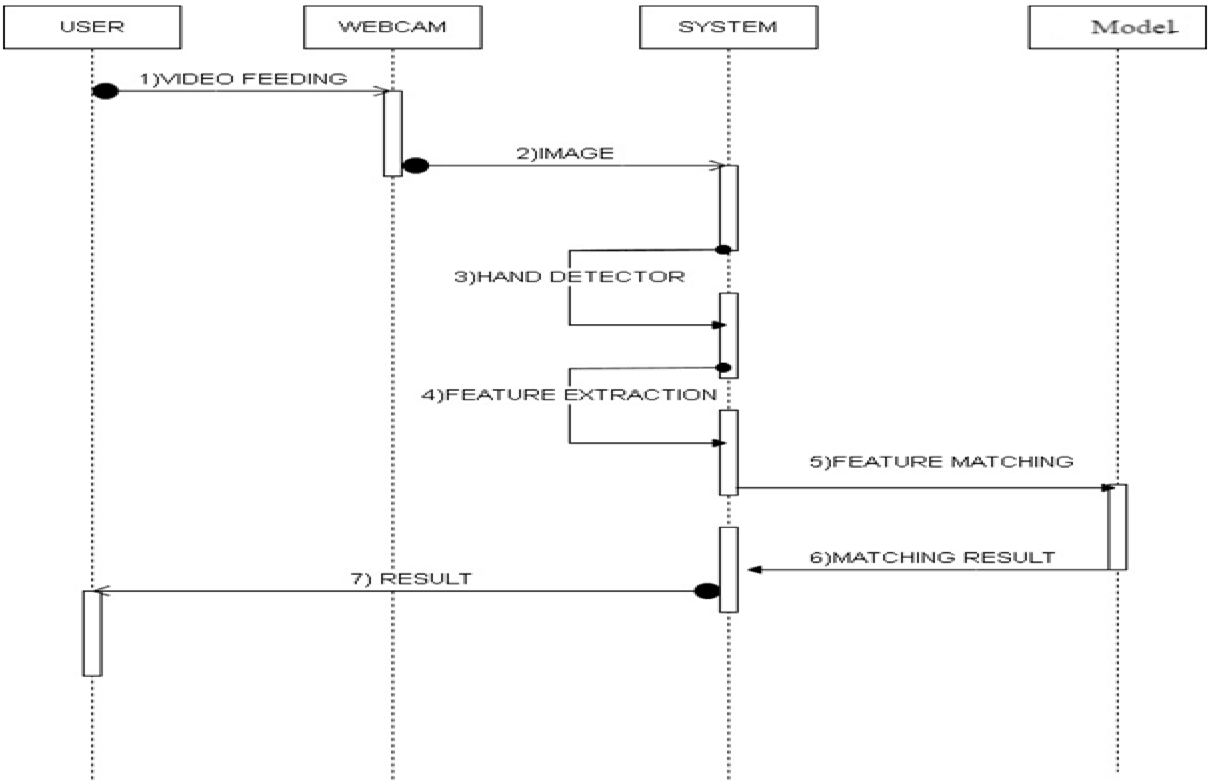

Fig 4 Sequence Diagram

DOI: https://doi.org/10.70274/jaict.2025.2.1.
p- ISSN 34562-3478 e- ISSN 5443-1243

Table 1: Comparison of Existing Systems

System	Functionality	Limitations
Gesture Recognition Apps	Recognizes gestures for specific applications, such as gaming or virtual controls	Limited to predefined gestures; not tailored for sign language translation
Translation Tools	Translates text or speech between spoken languages	No support for sign language or non-verbal communication
Educational Tools	Provides resources for learning sign language	Static content; lacks real-time interaction and translation features

3.2. Technical Feasibility

Recent developments of gesture recognition, mobile computing, and text-to-speech synthesis are mature enough, and a system of real-time translation into the Indian Sign Language may be made technologically feasible. Mobile platforms allow hand tracking using great feasibility with tools such as MediaPipe light, a part of TensorFlow Lite because it allows applications of relatively smooth hand-tracking implementation that fit into devices based on as minimal resources that their constraints require. Further on-device optimization is achieved for the learning model added to TensorFlow Lite. Therefore, with these, several models for gesture recognition can be run fast and accurately on a mobile. These tools enable any system to easily detect, using almost no computational resources, no processing, gestures meant specifically for ISL. Other ways to use the TTS engines for generating text-to speech syntheses in real time apply through the Google TTS and pyttsx3 TTS engine. Google TTS can synthesize very good speech with a very natural-sounding voice; recognized gestures are translated to speech as fast as possible within this cloud-based application; therefore, it is best used in stable Internet environments. As an offline version, it is possible to consider Pyttsx3 which supports TTS functionality even while being offline, although likely at compromised quality. The TTS methods support the low-latency requirements of a real-time translation system that would make it possible to provide spoken output for each detected gesture in terms of mobility application compliance. It is very easy with incremental learning to include some custom gestures into the system. Users can capture video clips of their new gestures, extract frames that can be labeled or added into the model. Using transfer learning it just adapts the use of new gestures by allowing an efficiency added with a whole-new vocabulary and does not break all this. Thus, one balancing approach between the flexibility as usability is achieved through which very flexibility has allowed an experience of the user as simple and will leave an untouched demand, that was computational as such. The further optimization of the application increases its technical feasibility. Lighter architectures in MediaPipe and TensorFlow Lite allow the operation to be battery efficient, which is a must for mobile devices. The techniques of frame skipping and selective processing may reduce consumption in the application to significant levels, thereby

DOI: https://doi.org/10.70274/jaict.2025.2.1.
p- ISSN 34562-3478 e- ISSN 5443-1243

increasing battery life and making the system more viable for longer periods of usage. Problems related to sensitivity to variation in lighting and background can be compensated by multiple training datasets and preprocessing. The project is technically valid, equipped with tools and techniques which may provide realtime ISL for mobile devices. By the effective gesture recognition, responsive TTS output, and user-friendly customization, the system has full potential to help users of ISL benefit in making access to communication-in-the-large, especially that during conversational interaction-in-real-time possible.

5. Discussion

Real-Time ISL to Text and Audio Translation: The project will enable immediate translation of ISL gestures into text and spoken language. This outcome will enhance communication for ISL users, bridging the communication gap with non-signers in various environments.
Custom Gesture Recognition: Add custom gesture for user's vocabulary and particular use. This personal interface would help users to configure the system without demanding advanced technical know-how.
Integrated TTS with Cloud and Offline Options: This will be an application featuring integration of TTS function both with Cloud and offline Options. It will thus use the offline TTS version, which is base upon pyttsx3 while supporting translation functionalities even on cases of connectivity.
Scalable Gesture Recognition Model: This project will deliver an adaptive model, which, as the vocabulary of ISL continues to grow, will take into consideration new gestures. That flexibility will continuously allow it to be evolved to be strong and scalable.
Greater Inclusivity for ISL Users: Providing a mobile application that will translate ISL is aimed at opening access to education, public services, social interaction, and voice communication. Through this project, more inclusive conditions for these activities can be attained.
Broader Impact on Accessibility: These technologies con- tribute significantly to accessibility as they promote the use of assistive technology for the ISL users and have opened up access to different social structures.

5. Conclusion

The proposed project will introduce the entire approach of developing a real-time Indian Sign Language translation for PC application. The application is designed to offer accessible choices of communication for the users of the ISL. It will successfully translate sign language into text and speech using gesture recognition, machine learning, and TTS technologies. This is the feature that allows the model to add custom gestures, thereby catering to the needs of individual vocabularies without requiring much retraining to remain adaptive and responsive. It further brings out the user-centered design aspect, thus easy to use and efficient in PC-based applications. Through accessible tools such as TensorFlow, MediaPipe, and TTS libraries, this application brings about the possibility of having a cost-effective yet high-performance ISL translation tool. This is why this project is critical in filling the communication gap for ISL users and promoting inclusivity in digital interactions and paving the way for further developments in assistive technology for sign language translation.

6. Acknowledgment

This work was supported by the Deanship of Scientific Research at CHRIST University, Bangalore, Karnataka, India

Funding:

"This research received no external funding"

References

[1] Ebey Abraham et al.," Real-Time Translation of Indian Sign Language Using LSTM," 2019.
[2] Harini R et al.," Sign Language Translation," 2020.
[3] Mohammed Elmahgiubi et al.," Sign Language Translator and Gesture Recognition," 2015.
[4] Zeyu Liang et al.," Sign Language Translation: A Survey," 2023.
[5] Necati Cihan Camgoz et al.," Neural Sign Language Translation," 2018.
[6] Mary Jane C. Samonte et al.," Deep Learning in Sign Language Translation," 2022.
[7] F. Ronchetti, F. M. Quiroga, C. Estrebou, L. Lanzarini, and A. Rosete," Sign Language Recognition without Frame-Sequencing Constraints: A Proof of Concept on the Argentinian Sign Language," arXiv preprint arXiv:2310.17437, Oct. 2023.
[8] O. Koller, N. C. Camg¨oz, and H. Ney," Weakly Supervised Learning of Sign Language Interpretations via Parallel Corpus Mining," IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019.
[9] S. Goyal, P. Pandey, and N. Bhatia," Real-Time Sign Language Recognition Using Mediapipe and Tensor- Flow," International Journal of Advanced Computer Science and Applications, 2021.
[10] C. Thomas and T. Ahmad," Exploring Sign Language Translation Challenges Using Transfer Learning," Computer Vision and Image Understanding, 2020.
[11] 11) Y. Shi, L. Yang, and X. Feng," Attention-Based Encoder- Decoder Models for Sign Language Translation," Neural Computing and Applications, 2023.
[12] 12) D. Sun and Z. Li," A Review of Gesture Recognition Techniques for Human-Computer Interaction," IEEE Access, 2022.
[13] J. Chung and D. Lee," Hand Pose Estimation for Robust Gesture Recognition in Complex Environments," Journal of Machine Learning Research, 2021.
[14] K. Shukla and R. Jain," Hybrid Models Combining CNN and LSTM for Sign Language Gesture Recognition," Springer Advances in Neural Networks and AI, 2022.
[15] P. Wang and F. Chen," Open-Source Solutions for Real- Time Ssign Language

DOI: https://doi.org/10.70274/jaict.2025.2.1.
p- ISSN 34562-3478 e- ISSN 5443-1243