International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 12 | Dec – 2025 DOI: 10.55041/ISJEM05230
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

Multilingual Conversion of Sign Language to Text

Sindhu M K[1], Shraddha S J[2], Vibha Gunaga[3], Y G Sneha[4]

[1]*Student, Dept. of Information Science Engineering, AMC Engineering College, Karnataka, India*
[2]*Student, Dept. of Information Science Engineering, AMC Engineering College, Karnataka, India*
[3]*Student, Dept. of Information Science Engineering, AMC Engineering College, Karnataka, India*
[4]*Student, Dept. of Information Science Engineering, AMC Engineering College, Karnataka, India*

Abstract - *This paper, Multilingual Conversion of Sign Language to Text aims to bridge the communication gap between the hearing-impaired community and non-sign language users through intelligent automation. Sign language, being a visual mode of communication, varies across regions, making translation across multiple languages a major challenge. This project utilizes computer vision and natural language processing (NLP) techniques to recognize and interpret hand gestures from sign language videos or live input. The captured gestures are pre-processed using background subtraction, image normalization, and frame extraction methods to enhance accuracy. A deep learning model, such as a Convolutional Neural Network (CNN) combined with a Long Short-Term Memory (LSTM) network, is employed to classify gestures and translate them into corresponding text. The translated text is then converted into multiple natural languages using NLP-based translation models. The system is designed to support the Indian Sign Language (ISL), ensuring inclusivity and adaptability. This multilingual feature makes the system applicable in diverse regions and enhances accessibility in education, workplaces, and public communication. Overall, the proposed system offers an efficient, real-time, and scalable solution to foster inclusivity and promote equal communication opportunities for the hearing-impaired community.*

***Key Words*: *Hearing- impaired, Indian Sign Language, Hand gestures, NLP, LSTM, Multiple languages*.**

1. INTRODUCTION

Effective communication is essential for social inclusion, yet individuals with hearing and speech impairments often face difficulties because most people do not understand sign language. In India, this challenge is more pronounced due to the limited awareness of Indian Sign Language (ISL), the primary and culturally rooted language of the Deaf community. To address this communication gap, AI- and ML based systems have emerged as powerful tools for automatic gesture interpretation.

Advancements in computer vision and deep learning now enable accurate recognition of complex hand gestures. This study proposes a multilingual sign language–to–text conversion system focused on ISL. The approach involves acquiring and pre-processing ISL gesture data using techniques such as background subtraction, image normalization, and frame resizing. Key features—including hand shape, orientation, and movement—are extracted using tools like OpenCV and Media Pipe and classified using a hybrid CNN–LSTM model capable of capturing both spatial and temporal characteristics. Once recognized, the gesture is converted into text and translated into multiple regional languages using NLP-based translation. This multilingual framework promotes accessibility, strengthens inclusivity, and contributes to ongoing efforts to standardize ISL recognition technologies for real-time communication support.

1.1 Problem Statement

Individuals with hearing and speech impairments rely on Indian Sign Language (ISL) for communication, yet most of the population in India is not familiar with ISL. This lack of understanding creates a significant communication barrier, limiting accessibility, social inclusion, and independence for the Deaf community. Existing sign language recognition systems often focus on single languages, offer limited accuracy, or fail to support the multilingual output required in India's linguistically diverse environment. Moreover, many current models struggle with real-time performance, dynamic gesture recognition, and scalability when exposed to varying lighting conditions or hand orientations.

Therefore, there is a need for an intelligent, robust, and real time system that can accurately recognize ISL gestures and convert them into readable text across multiple regional languages. Such a solution would help bridge communication gaps in public services, education, healthcare, and day-to-day interactions. This project aims to address this gap by developing a multilingual ISL-to-text translation system using advanced computer vision and deep learning techniques, enabling seamless, inclusive, and accessible communication for the Deaf community in a diverse linguistic ecosystem.

2. Methodology

The proposed system for Multilingual Conversion of Indian Sign Language (ISL) to text is developed through a structured pipeline integrating computer vision, deep learning, and

© 2025, ISJEM (All Rights Reserved)

natural language processing techniques. The methodology comprises the following stages:

1. Data Acquisition:

- ISL gesture data is collected from video recordings and image sequences representing both static and dynamic gestures commonly used in fingerspelling and basic vocabulary. The dataset includes variations in background, illumination, and hand positioning to ensure model generalization and robustness.

2. Pre-processing:

- Background Subtraction: Removes irrelevant visual information and isolates the hand region.
- Image Normalization: Standardizes brightness and contrast to reduce variability.
- Frame Resizing: Ensures uniform input dimensions compatible with deep learning models.
- Hand Segmentation: Extracts the region of interest using skin detection or Media Pipe-based hand landmark detection.

3. Feature Extraction:

- Feature extraction focuses on capturing both spatial and temporal characteristics of ISL gestures
- Spatial Features: Hand shape, finger configuration, and orientation are extracted using OpenCV-based contour analysis and Media Pipe hand landmarks.
- Temporal Features: Movement trajectories and dynamic transitions between frames are captured to support recognition of sequential gestures.

4. Gesture Recognition:

- The combined CNN–LSTM model enables accurate recognition of both static and continuous gesture sequences.
- Convolutional Neural Network (CNN): Learns spatial features from individual frames such as hand shape and orientation.
- Long Short-Term Memory (LSTM) Network: Captures temporal dependencies and sequential information in dynamic gestures.

5. Text Conversion and Multilingual Translation:

- Once a gesture is classified, the recognized label is converted into text in the source language. This text is then processed through NLP-based translation modules or API-driven translation models to generate output in multiple regional languages. The system supports multilingual display to improve accessibility for users across different linguistic backgrounds.

6. User Interface Display:

- The final translated text is presented through a user friendly interface designed for real-time interaction. The interface displays the recognized gesture, the intermediate text, and the translated output in the selected language.

3. IMPLEMENTATION

The proposed Sign Language Recognition System was implemented using a combination of Media Pipe for hand landmark detection and machine learning models for gesture classification. The system begins by capturing live video input from a webcam, which is processed in real time using the Media Pipe Hands module. This framework extracts 21 three-dimensional hand landmarks from each frame, providing highly accurate geometric information about the user's hand posture. These coordinates are normalized and converted into structured feature vectors, which serve as inputs to the classification model. Several algorithms—including CNN, SVM, and Random Forest—were evaluated during development; however, lightweight machine-learning models were ultimately preferred because they offered high accuracy while maintaining real-time performance with minimal latency. Once a gesture is predicted, the corresponding text is displayed immediately on screen, and a text-to-speech module generates spoken output to further enhance accessibility.

In addition to the core recognition system, a dedicated webpage was developed to present the project in a more interactive and user-friendly manner. This webpage includes sections explaining the project workflow, model architecture, dataset details, and live demonstrations. The web interface makes the system accessible to a broader audience and allows users to understand, test, and interact with the gesture recognition pipeline without requiring technical setup. Together, the backend recognition engine and the webpage form a complete system capable of demonstrating the potential of AI driven sign language interpretation.

4. RESULT AND DISCUSSION

The implemented Sign Language Recognition System successfully recognizes ISL fingerspelling gestures in real time with high accuracy and stable performance. Media Pipe consistently detects hand landmarks at fast frame rates, enabling smooth and reliable gesture tracking across different lighting and background conditions. The trained classifier

performs well for alphabet gestures, producing accurate text predictions that appear instantly on the screen. The addition of text-to-speech further enhances usability by converting recognized gestures into spoken output, improving communication between deaf-mute and hearing individuals.

The decision to use Media Pipe proved effective, as it offered lower latency and better real-time performance compared to heavier models like YOLO or Faster R-CNN. The accompanying webpage helped showcase system functionality and allowed users to interact with the model easily, improving accessibility and understanding of the project. Minor challenges such as reduced accuracy during rapid movements or extreme distances from the camera were observed, but these can be addressed with larger datasets and sequence-based models in future work. Overall, the results confirm that the system is efficient, practical, and well-suited for assistive communication applications.

5. CONCLUSION

This work presents an intelligent and multilingual system designed to bridge the communication gap between individuals with hearing impairments and the hearing community by converting Indian Sign Language (ISL) gestures into text. By integrating computer vision techniques with a hybrid CNN–LSTM deep learning architecture, the system effectively recognizes both static and dynamic ISL gestures with improved accuracy. The incorporation of NLP based translation further enables the generated text to be converted into multiple regional languages, making the system highly relevant in India's diverse linguistic landscape.

The proposed methodology demonstrates the potential of AI-driven technologies in promoting accessibility, inclusivity, and independence for the Deaf community. Through real time gesture recognition, multilingual output, and a user friendly interface, the system contributes to ongoing efforts toward standardizing ISL-based communication tools. Future enhancements could include expanding the dataset, incorporating full-sentence recognition, and extending support to additional sign languages, enabling broader applicability and impact.

REFERENCES

[1] Deepika Patil, Amol More, Suraj Shinde, "Sign Language to Text Conversion", vol. 12, 2024.

[2] Carl Jose, Ron Andrew, Mark Joseph, John Ray D. Tamayo, "Using Deep Learning in Sign Language Translation to Text", 2022.

[3] Medhini Prabhakar, Prasad Hundekar, Sai Deepthi B P, Shivam Tiwari, "Sign Language Conversion to Text and Speech", vol. 9, 2022.

[4] Anuja v. Nair, Bindu V, "A Review on Indian Sign Language Recognition," International Journal of computer Applications, July 2013.

[5] Ramesh M. Kagalkar, Nagaraj H.N, "Methodology for Translation of Static Sign Symbol to Words in Kannada Language", 2020.

[6] https://www.ibm.com/in-en/topics/convolutional-neural-networks

[7] https://learnopencv.com/introduction-to-mediapipe/