IRJAEH
International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 10 October 2025
Page No: 3930-3934
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0573

Enhancing Communication Methods for Individuals with Physical Disabilities

Shravan[1], Karthiga G[2], Shwetha[3], Tarun R N[4], Soumya R[5]
[1,3,4,5]UG – Computer Science and Engineering, AMC Engineering College, Bengaluru, Karnataka, India.
[2]Associate Professor, Computer Science and Engineering, AMC Engineering College, Bengaluru, Karnataka, India.
Emails: *chavanshravan84@gmail.com[1], karthiga.gurunathan@amceducation.in[2], shwethagundad@gmail.com[3], tarun393939@gmail.com[4], soumyaramanagoudra@gmail.com[5]*

Abstract
Communication is a basic human need. However, people with physical disabilities, especially those who are deaf, mute, or both, face significant challenges in sharing and understanding information. Current options like sign language, lip-reading, and text-based communication either lack universal application or need expensive hardware, like sensor-equipped gloves. Additionally, many existing systems only work one way and do not support Indian Sign Language (ISL), which is essential in India. This paper introduces a real-time communication platform powered by AI to address these issues. The system uses Convolutional Neural Networks (CNNs) to recognize gestures. It also uses Natural Language Processing (NLP) for understanding spoken and written text, Automatic Speech Recognition (ASR) for voice input, and text-to-speech technology. It provides five communication methods: sign-to-text/speech, speech-to-ISL animation, speech-to-text, text-to-speech, and sign-to-text for mute users. The platform is built with open-source tools like MediaPipe, OpenCV, and Google TTS. It only requires a standard webcam and microphone, making it affordable and accessible. The system aims to enhance inclusivity in education, healthcare, workplaces, and public services by allowing real-time, two-way, and cost-effective communication.
Keywords: *Sign Language Recognition, Indian Sign Language, CNN, Speech-to-Text, NLP, Accessibility.*

1. Introduction
1.1. Background
Human communication is essential for participation in education, healthcare, work, and everyday interactions. For people who are deaf, mute, or both, communication barriers often result in isolation and fewer opportunities. The World Health Organization (WHO) reports that over 1.5 billion people worldwide face some form of hearing loss. In India alone, more than 63 million people have hearing disabilities. Technological progress has introduced solutions like sign language interpreters, text-messaging tools, and gesture recognition devices. However, these options can be costly, non-portable, or restricted to certain languages, making them unsuitable for common use.

1.2. Problem Statement
Current communication systems for disabled individuals have several issues:
- **Hardware dependency:** Many systems rely on gloves or sensors, which raises costs and reduces convenience.
- **Unidirectional interaction:** Most systems only change gestures to text, lacking support for two-way communication.
- **Language limitations:** A focus on American Sign Language (ASL) often overlooks Indian Sign Language (ISL), creating a gap in the Indian context.
- **Accessibility issues:** Current solutions are not affordable or scalable enough for widespread

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 10 October 2025
Page No: 3930-3934
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0573

use.

1.3. Objectives

The goals of this research are to:

- Create a real-time, multi-directional communication platform for deaf, mute, and hearing individuals.
- Change ISL gestures into text and speech using CNN-based image recognition.
- Transform spoken or written text into ISL animations with NLP and animation models.
- Enable real-time communication without needing special hardware.
- Encourage inclusivity and independence for disabled individuals in education, healthcare, and workplaces.

1.4. Contributions

This paper offers:

- A gesture-to-text and gesture-to-speech system using CNN and MediaPipe.
- A speech-to-ISL animation module for deaf access.
- A speech-to-text and text-to-speech system for mute individuals.
- An integrated software platform that supports five communication methods.

2. Related Work

2.1. Gesture Recognition Approaches

CNNs and vision-based methods have been widely researched for recognizing sign language. Patil et al. (2022) achieved high accuracy in translating ISL gestures to speech using CNNs and Google TTS. Deshpande and Shettar (2023) employed MediaPipe Holistic with CNN for ISL recognition and produced multilingual speech output. Chaudhary et al. (2024) used MediaPipe with CNNs to recognize ASL alphabets, achieving 95% accuracy.

2.2. Multimodal Communication Systems

Researchers have also looked into multimodal solutions. Balamanikandan et al. (2024) proposed a modified VGG16 model for real-time gesture-to-speech translation. Kowsigan et al. (2024) developed a system that combines speech-to-sign and sign-to-text translations. Puri et al. (2023) introduced a multilingual live-captioning system that supported ASL, ISL, BSL, and FSL, with over 96% recognition accuracy.

2.3. Limitations of Existing Systems

Despite these developments, challenges still exist:

- Hardware-based systems, such as glove-based recognition (Manoj, 2020), are cheap but lack scalability.
- Many systems only offer one-way communication.
- Limited support for ISL makes it less useful in the Indian context.
- High costs and specific environmental needs lower real-world applicability.

2.4. Research Gap

There is no integrated, software-only, real-time platform that:

- Works with ISL.
- Supports multiple communication pathways.
- Operates with affordable, widely available devices like webcams and microphones.

3. Proposed Methodology

3.1. System Architecture

The proposed platform has a three-layer architecture (Fig. 1):

3.1.1. Input Layer

- Captures visual input through a standard webcam and audio input through a microphone.
- Ensures it works with low-cost consumer devices, so there is no need for special hardware.

3.1.2. Processing Layer

- **Gesture Recognition Module:** MediaPipe and OpenCV extract hand landmarks; a CNN classifies gestures into ISL symbols.
- **Speech Processing Module:** Automatic Speech Recognition (ASR) turns spoken input into text.
- **NLP Module:** Performs tokenization, aligns grammar, and maps to ISL grammar for precise animations.
- **Text-to-Speech (TTS) Engine:** Turns recognized text into natural-sounding audio using Google TTS.

3.1.3. Output Layer

- **Delivers outputs in three forms:** text on

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 10 October 2025
Page No: 3930-3934
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0573

screen, synthesized speech, and ISL animations.

3.1.4. Enables communication across five pathways

- Sign to Text/Speech (Deaf to Normal)
- Speech to ISL Animation (Normal to Deaf)
- Speech to Text (Normal to Mute)
- Text to Speech (Mute to Normal)
- Sign to Text (Deaf to Mute)

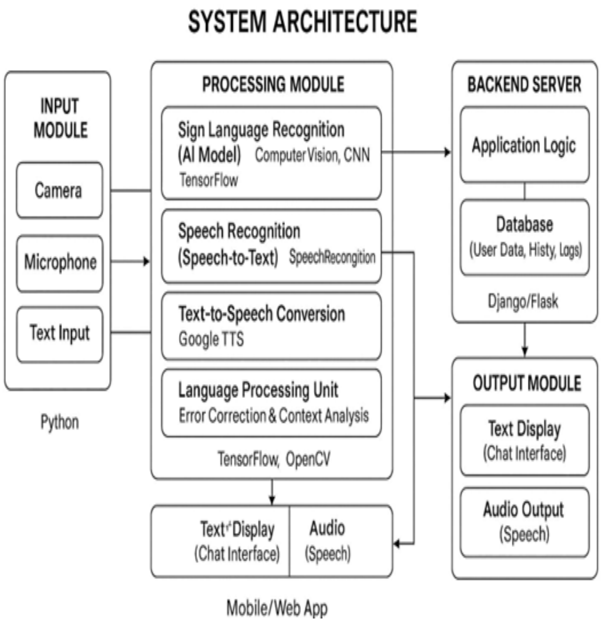

Figure 1 System Architecture

3.2. Gesture Recognition Pipeline (Sign to Text/Speech)

- **Input:** Video frames from webcam.

3.2.1. Preprocessing

- Hand segmentation using OpenCV.
- Keypoint extraction with MediaPipe Holistic.
- **Feature Extraction:** Spatial and temporal features (hand shape and movement) are turned into a feature vector.
- **Classification:** A CNN model trained on ISL datasets connects the gesture to the right symbol or word.
- **Output:** Shown as text on screen and turned into speech.
- This module allows real-time communication from deaf to non-deaf speakers.

3.3. Speech-to-ISL Animation Pipeline (Normal to Deaf)

- **Input:** Spoken language via microphone.
- **ASR Conversion:** Speech is turned into text using pretrained ASR models.

3.3.1. NLP Mapping

- Sentence tokenization.
- ISL grammar adjustment (removing filler words, reordering for ISL structure).
- **Animation Generation:** ISL animations created using a lightweight 3D model or pre-recorded animation library.
- **Output:** Shown as animated ISL signs to deaf users. This module makes sure spoken words are expressed in ISL, promoting inclusive communication.

3.4. Speech-to-Text and Text-to-Speech (Normal to Mute)

- **Speech-to-Text:** Converts audio input into text with ASR, displayed on screen instantly.
- **Text-to-Speech:** Turns typed text into natural audio output, supporting multiple languages. These modules help mute and normal speakers communicate, especially in classrooms and workplaces.

3.5. Integration and User Interface

The final platform is implemented as a desktop or web application with these design principles:

- **Real-Time:** Keeps processing latency under 500 ms for smooth conversation flow.
- **Affordability:** Uses open-source libraries (MediaPipe, OpenCV, TensorFlow, Google TTS).
- **Scalability:** Modular architecture allows for extension to various sign languages (ASL, BSL, etc.).
- **User-Friendly:** Features a simple GUI with microphone and camera controls, a text display panel, and an animation viewer.

3.6. Novelty of the Approach

Unlike previous systems that rely on specific hardware or only work one way, the proposed method:

- Supports communication in all directions among deaf, mute, and normal individuals on

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 10 October 2025
Page No: 3930-3934
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0573

one platform.

- Focuses on ISL, addressing a significant gap in research and application in the Indian context.
- Uses only regular hardware (webcam and microphone), making it affordable.
- Combines CNN, ASR, NLP, and TTS to ensure complete inclusivity and accessibility.

4. Expected Results

The system will provide real-time, two-way communication between deaf, mute, and hearing individuals. It aims for gesture recognition accuracy of 90 to 95% using CNN models and ISL datasets. The platform ensures low latency, affordability, and scalability for use in:

- **Education:** inclusive classrooms with real-time communication.
- **Healthcare:** communication between patients and doctors for individuals with hearing or speech impairments.
- **Public services:** transport, government offices, and emergency services.
- **Corporate environments:** better accessibility in the workplace.

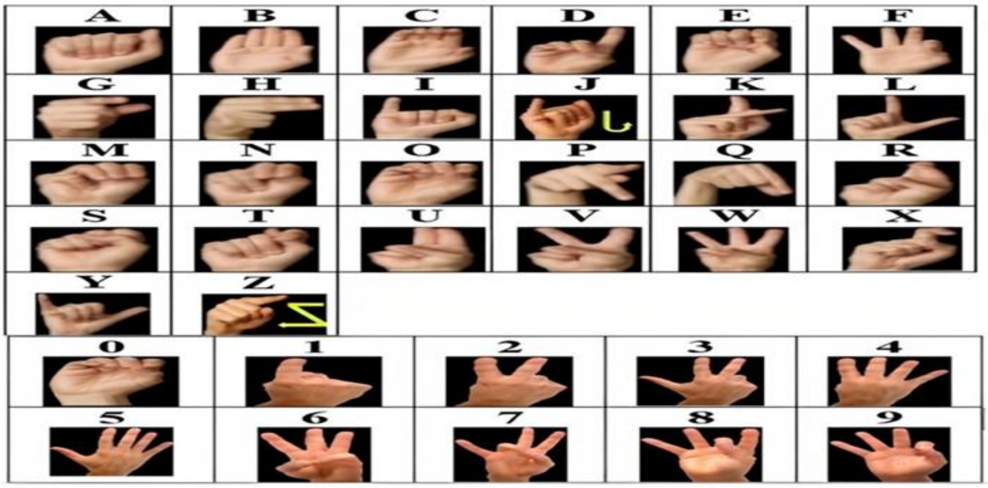

Figure 2 Sign Language Representation

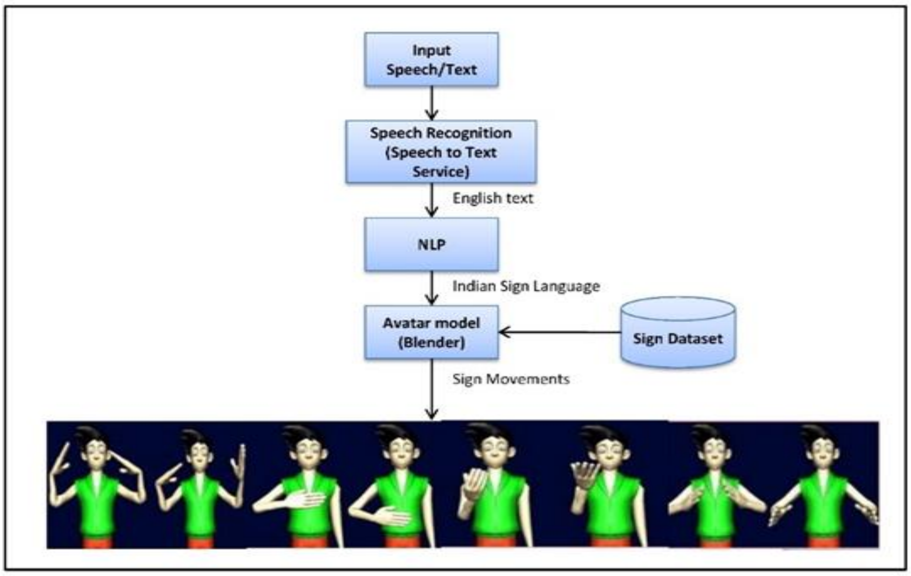

Figure 3 Input – Action Processing for Avatar

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 10 October 2025
Page No: 3930-3934
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0573

Conclusion and Future Work

This work presents an AI-powered communication platform that combines gesture recognition, speech recognition, NLP, and text-to-speech synthesis into a single solution. By removing the need for costly hardware and supporting ISL, the platform promotes inclusivity for millions of people with hearing and speech impairments.

Future improvements will include:

- Extending support for multiple sign languages
- Developing a mobile application for wider adoption.
- Using deep learning models to improve recognition accuracy.
- Deploying in the cloud for remote accessibility.

References

[1]. S. Patil et al., "Conversion of Indian Sign Language to Speech by Using Deep Neural Network," ICCUBEA, 2022.
[2]. J. Ajay et al., "Analysis of Machine Learning Techniques for Sign Language to Text Conversion," ICCCI, 2023.
[3]. S. Sahu et al., "Study of Sign Language Conversion Using Machine Learning," UPCON, 2023.
[4]. M. Puri et al., "Helping the Deaf, Mute & Blind Community through Live Captioning," ASIANCON, 2023.
[5]. B. A. Balamanikandan et al., "Hand Gesture to Speech Empowering Communication for the Speech Impaired," TEECCON, 2024.
[6]. P. Vaidya et al., "Empowering Communication: Converting Sign Language to Text Using CNN Technology," IDICAIEI, 2024.
[7]. M. Kowsigan et al., "An Efficient Speech to Sign Language Conversion and Text Recognition," ICSPCRE, 2024.
[8]. R. S. Ravikumar et al., "Interactive Sign Language Learning System using Deep Learning," ICSSAS, 2024.
[9]. M. Chaudhary et al., "Sign Language to Text Conversion Using Machine Learning," ICAC2N, 2024.
[10]. J. Shajeena et al., "Sign Language Translation and Voice Impairment Support System using Deep Learning," ICDICI, 2024.