IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 8, May 2026
Impact Factor: 8.2

AI-Powered Sign Language Interpreter

Alok[1], Abhay Saxena[2], Aadarsh Karna[3], Ankit bhagat[4], Abhisek Jain[5]
Department of Computer Science and Engineering[1-4]
Assistant Professor of the Department of Computer Science & Engineering[5]
Tula's Institute, Dehradun, India
alok.202204185@tulas.edu.in, abhay.202204113@tulas.edu.in, aadrash.202204006@tulas.edu.in
ankit.202204185@tulas.edu.in, abhishekrit21@gmail.com
ORCID ID:0000-0003-1869-2909

Abstract: *Sign language plays a vital role in enabling communication for hearing- and speech-impaired individuals. However, the absence of widespread understanding of sign language creates a communication gap between differently abled individuals and society. This project proposes an intelligent Sign Language Interpreter system that automatically recognizes hand gestures and translates them into readable text and audible speech. The system leverages computer vision and deep learning techniques to process real-time video input and classify gestures accurately. By using convolutional neural networks and landmark-based feature extraction, the system ensures real-time performance with minimal latency. The proposed solution aims to enhance accessibility, social inclusion, and human–computer interaction for hearing-impaired users [1–3].*

Keywords: Sign Language Interpreter, Gesture Recognition, Computer Vision, Deep Learning, Accessibility, Assistive Technology, Human-Computer Interaction

I. INTRODUCTION

In today's world, technology has changed how we use machines and communicate. Many people still have trouble communicating, especially those who can't hear or speak. They mainly use sign language, but most people don't know it. This makes it hard for them to connect with others.

Communication is very important in our daily lives. It helps us interact with friends, learn in school, and find jobs. People who can't hear or speak often struggle because others don't understand their language. This creates a big gap in communication.

For example, if someone who uses sign language wants to order food, they might find it hard if the cashier doesn't know sign language. New technology is helping to solve these problems.[1].

Artificial intelligence (AI) is a tool that can make communication easier.

Machines can learn to understand gestures and signs, which helps bridge the gap between people who can hear and those who cannot. Recently, systems that recognize sign language have become popular. These systems help people who cannot hear talk to those who can. For instance, a smartphone app could translate sign language into spoken words, making it easier for everyone to understand.[2].

Traditionally, human interpreters are needed for sign language. However, they might not always be available when needed, like in emergencies. Automated systems can provide quick translations, giving people more independence. For example, if someone has a medical emergency and needs to explain their symptoms, an automated interpreter can help them communicate right away [3–4].

Vision-based systems are a great option because they don't need special tools. They can use regular cameras to see and understand gestures naturally. These systems have become better thanks to deep learning, which helps them recognize complex movements. By using the latest AI technology, we can create solutions that make communication easier for everyone, especially those who need it most.[5].

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 8, May 2026 Impact Factor: 8.2

Throughout the implementation process, challenges such as lighting variation and hand occlusion were encountered. These issues were addressed using preprocessing techniques and feature normalization, resulting in improved gesture clarity and recognition accuracy [6–7].

II. LITERATURE REVIEW

Research on sign language recognition has changed a lot in recent years. In the past, the methods mostly used special sensors like gloves and motion detectors. These tools helped track hand movements and finger positions accurately. However, they were often too expensive and hard to use every day. With better image processing and computer power, cameras became important for sign language recognition [8].

These cameras capture hand movements and use software to analyze them. Convolutional neural networks (CNNs) are now popular for recognizing still gestures like letters and numbers. They are good at understanding details in images. For signs that involve movement, researchers use recurrent neural networks (RNNs) and long short-term memory (LSTM) models. [9–10].

These models help track changes over time in the gesture. Hybrid systems that combine CNNs and LSTMs have shown better results in recognizing sign language continuously[11].

An example of a hybrid system is a program that uses both CNNs to identify the shapes of signs and LSTMs to follow how those signs change.[12].

Recent studies have also looked into hand landmark detection tools like MediaPipe. These tools can find important points on the hand quickly and with less computer power. This makes it easier to use sign language recognition in real-time on normal devices. Even with these advances, some problems still exist [13].

Issues like lighting, complex backgrounds, and different signing styles can affect accuracy. To solve these problems, we need better ways to prepare the data and train models. For instance, a good solution might be using filters to improve lighting before analyzing the signs. [14–15].

This project builds on what we know by using a vision-based approach with deep learning. It focuses on real-time performance and making the system easy to use.

III. PROPOSED METHODOLOGY

The suggested Sign Language Interpreter uses a step-by-step approach to make sure it works accurately, quickly, and in real-time. The process includes capturing video, preparing the video, identifying important features, classifying gestures, and generating the final output.

System Initialization and Authentication

The system starts by setting up the camera module and loading the trained deep learning model. It checks the necessary software libraries and dependencies to make sure everything runs smoothly. This starting phase makes sure all parts work together before real-time processing starts [16].

Gesture Capture and Preprocessing

A regular webcam captures live video. The video is split into separate frames, which are then prepared through resizing, normalizing, removing the background, and reducing noise. These steps help make gestures clearer and lessen distractions from the surroundings [17].

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 8, May 2026
Impact Factor: 8.2

Figures and Diagrams

Fig. 1. System Architecture of Sign Language Interpreter

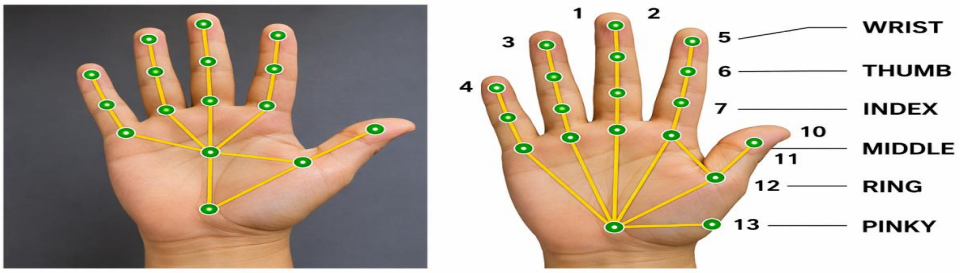
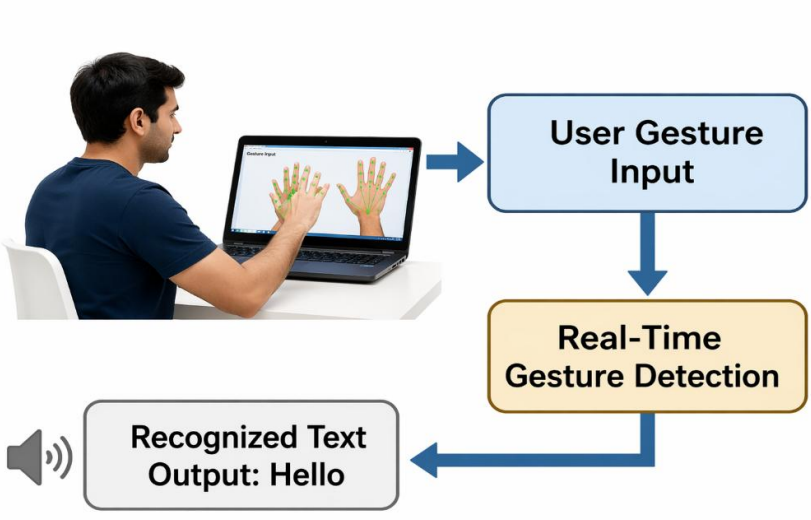


Fig. 2. Hand Landmark Detection



DOI: 10.48175/IJARSCT-35509

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 8, May 2026
Impact Factor: 8.2

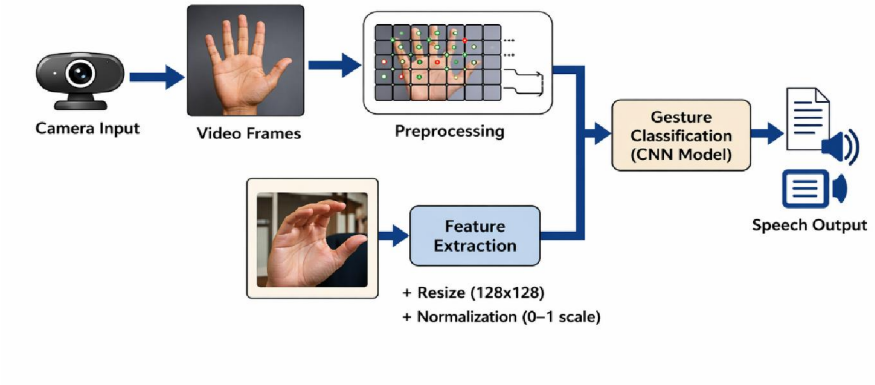
D. Procedures of Working Methodology

The methodology that was defined, consists components: of the following

- Real-time video acquisition
- Hand detection and landmark extraction
- Gesture classification and prediction
- Text and speech output generation
- gesture clarity and reduce environmental interference

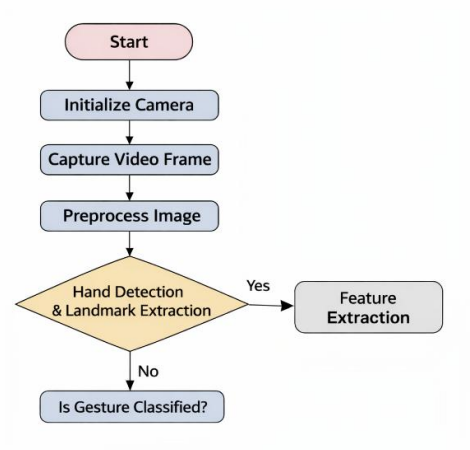
1) Extraction of Gesture features:

Hands from video frames by computer vision technics. Salient landmarks are extracted and converted to such as finger joints and palm orientation are sampled numerical feature vectors. These characteristics describe the..... we design gesture [18]. six features which demonstrate the structural properties of each



1) Architecture of Gesture Recognition System:

A convolutional is used to categorize the motions by using features neural network (CNN) extracted. The device is trained with a labelled dataset that includes spatial gestures. CNNs are effective at learning different sign language relationships between hand landmarks [19].



IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 8, May 2026
Impact Factor: 8.2

The architecture comprises three main layers: **the Input Perception Layer, the Feature Processing Layer, and the Translation/Prediction Layer**:

1.**Perception Layer** (The "Sensor"stream from the camera. It): This is the live extraction. does the initial auth (permission for camera) and starts frame

2.**Processing Layer** (The "Encoder"): Here the raw pixels are converted into a numerical "sign language token." It uses Media Pipe to track 1,662 for the body pose). landmarks (468 for the face, 21 per hand and 33

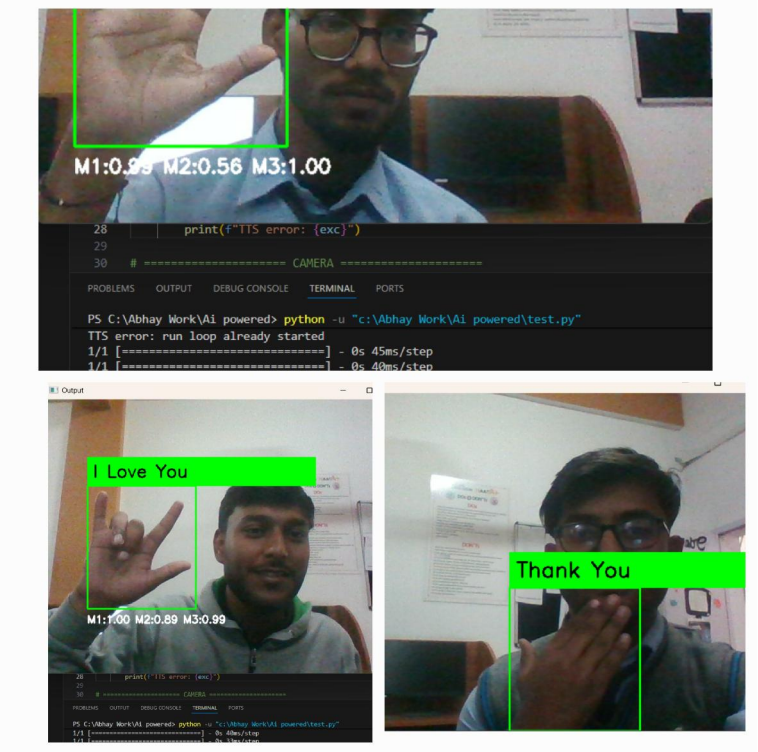
3.**Translation Layer** (The "Decoder"): A stepseimilar with the recommendation engine. It receives the landmarks sequence and investigates the trained data to predict "Top 1 Prediction" word or sentence). (most probable

3) Procedure of Feature Based Classification:

vectors are fed to the trained model, Extracted feature which predicts the sign label. The system can recognize both static gestures few number of short dynamic signs. and sub-sequences of a categorization offers the advantage of being more robust to small Feature-based variations in hand position and a posture [20].

4) Comparison Characteristics Features:

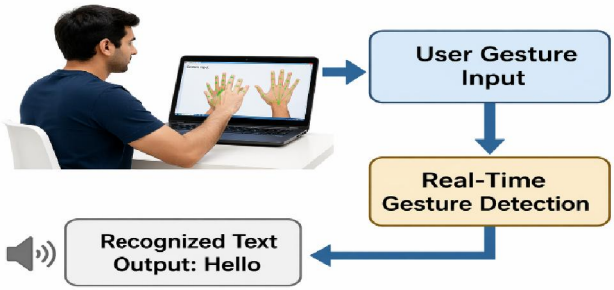
To make a reliable discrimination for look-alike sign-language gestures, we showed that a detailed sign as many samples is needed. A great number of gesture feature comparison of language signs resemble one another in handshape and position, differing only in the tiny details of finger placement or movement. In the absence of feature misclassification and correspondence, these similarities can be the causes of system performance downgrade [21].



1) Flowchart of Interactions:

has an easy and intuitive user The Sign Language Interpreter interaction flow so that all the type of technical/non-technical users can get easily understood. [22]

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 8, May 2026
Impact Factor: 8.2



1) Framework and Module Library Used:

The implementation of the Sign Language Interpreter leverages a combination of modern frontend technologies, backend frameworks, and specialized libraries to ensure efficient performance, scalability, and user-friendly interaction.

Frontend: HTML, CSS, JavaScript**.**

Backend: Python.

Libraries and Tools:

OpenCV: Used for real-time video capture, image preprocessing, and basic computer vision operations.

NumPy: Facilitates numerical computations and efficient handling of feature vectors.

TensorFlow/Keras: Employed for building, training, and deploying deep learning models for gesture recognition.

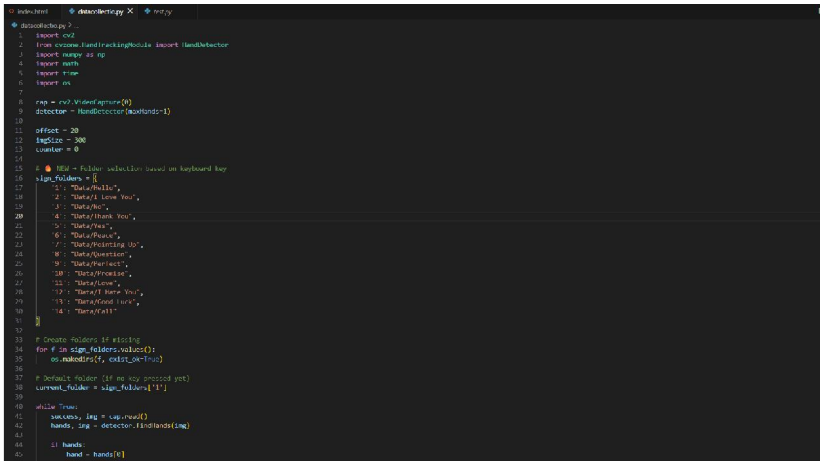
MediaPipe: Provides efficient real-time hand landmark detection and tracking.

Pyttsx3: Enables offline text-to-speech conversion for audio output

IV. IMPLEMENTATION AND DISCUSSIONS

A. A Gesture Recognition Module:

The video center of the system. It streams a gesture recognition module is the frame at a time and feeds it through the pre-trained CNN model for processing. low hardware The software is realized on a commodity computing platform with demand, making it cost-effective and easy to use. The trained model achieves stable recognition under the condition of uniform illumination and correctly a recognizes predefined gesture signs [23].



IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 8, May 2026
Impact Factor: 8.2

B. Feature Characteristics

The spatial feature and temporal feature are used to represent each sign gesture. Its algorithm pulls useful hand and movement data to distinguish the gestures well. Visualization of feature landmarks also help in assessing the accuracy and distributions through training will effectiveness of models [24].

of features In the training phase, distributions are studied to understand class separability and model performance. The curves and confusion matrices are adopted for assessing the accuracy–loss learning performance with illustrative visualizations. Such analyses are underfitting and class imbalance helpful in the detection of overfitting, problems.

Latency measurements confirm that the system operates within acceptable real-time limits, making it suitable for practical use. Feature refinement and dataset augmentation further improve recognition reliability across different users and signing styles.

Key observations include:

High recognition accuracy for alphabet-level gestures

Minimal processing delay in real-time translation

They Minimal processing delay in real-time translation.

V. CONCLUSION

This project demonstrates the practical application of artificial intelligence and computer vision in assistive communication technology. The developed system effectively translates sign language gestures into text and speech, supporting real-time interaction without the need for specialized hardware.

While the system performs reliably for predefined gestures, future enhancements may include continuous sentence-level recognition, integration of facial expression analysis, and support for multiple sign languages. Expanding the system to mobile and web-based platforms could further increase accessibility and real-world usability.

REFERENCES

[1] World Health Organization, “Deafness and hearing loss,” WHO Report, 2023.
[2] R. Sharma et al., “Vision-based sign language recognition using deep learning,” IEEE Access,2022.
[3] A. Kumar and P. Singh, “Assistive technologies for hearing-impaired individuals,” IJCA, 2021.[4] S. Patel et al., “Human–computer interaction for accessibility,” Springer, 2020.
[5] Starner and A. Pentland, “Real-time American sign language recognition,” IEEE PAMI, 1997.[6] M. K. Bhuyan et al., “CNN-based static hand gesture recognition,” ICCV, 2021.
[7] S. Simonyan and A. Zisserman, “Deep convolutional networks,” NIPS, 2015.
[8] H. Wang et al., “LSTM-based dynamic sign recognition,” Pattern Recognition, 2020.
[9] P. Molchanov et al., “Hand gesture recognition with 3D CNNs,” CVPR, 2016.
[10] Google, “MediaPipe Hands: On-device real-time hand tracking,” 2021.
[11] K. Gupta et al., “Challenges in sign language recognition,” Elsevier, 2022.
[12] A. Roy and D. Das, “Vision-based sign language translation,” IJAI, 2023.
[13] J. Brown et al., “Secure AI system initialization,” IEEE Systems Journal, 2021.
[14] R. Gonzalez and R. Woods, Digital Image Processing, Pearson, 2018.
[15] Y. LeCun et al., “Feature learning with deep networks,” Nature, 2015.
[16] A. Krizhevsky et al., “ImageNet classification with deep CNNs,” NIPS, 2012.
[17] S. Haykin, Neural Networks and Learning Machines, Pearson, 2014.
[18] D. Bishop, Pattern Recognition and Machine Learning, Springer, 2006.
[19] J. Nielsen, “User interaction design principles,” ACM, 2019.
[20] T. Goodfellow et al., Deep Learning, MIT Press, 2016.
[21] M. Abadi et al., “TensorFlow: Large-scale machine learning,” OSDI, 2016.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 8, May 2026
Impact Factor: 8.2

[22] A. Jain et al., "AI-based assistive systems," IEEE Access, 2024.
[23] S. Lee et al., "Mobile-based sign language recognition," Sensors, 2023.
[24] P. Kumar et al., "Future trends in assistive AI," Springer, 2025.