



# International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)





# Real-Time Sign Language Recognition Using Deep Learning and Computer Vision

**Shreedatt Pandya, Shiva Singh, Prasad Sujalkumar**

UG Student, Computer Science and Engineering, Parul University, Vadodara, Gujarat, India

**Arpita Vaidya, Rashmi Pandey, Fesal Shaikh**

Project Guide/ Assistant Professor, Computer Science and Engineering, Parul University, Vadodara, Gujarat, India

**ABSTRACT:** Sign language recognition (SLR) has gained significant attention in recent years due to the increasing need for inclusive communication technologies for individuals with hearing and speech impairments. Traditional sign language recognition systems rely on sensor-based hardware such as gloves and motion trackers, which can be costly and inconvenient. Recent advancements in artificial intelligence (AI) and deep learning have enabled real-time, vision-based sign language recognition systems using convolutional neural networks (CNNs) and recurrent neural networks (RNNs). This research proposes a real-time sign language recognition system leveraging deep learning and computer vision techniques to translate hand gestures into meaningful text or speech

**KEYWORDS:** Sign Language Recognition, Deep Learning, Computer Vision, Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Real-Time Gesture Recognition, Artificial Intelligence (AI), Human-Computer Interaction (HCI), Accessibility Technology, Speech and Hearing Impairment Solutions.

## I. INTRODUCTION

Information and Communication Technologies (ICTs) and Artificial Intelligence (AI) play a crucial role in bridging communication gaps between visually impaired, deaf, and hearing populations. Automated sign language analysis is essential to facilitate effective interaction and ensure inclusivity in next-generation Human-Computer Interfaces (HCI). According to the 2011 Indian Census, approximately 1.3 million individuals have hearing impairments. However, data from India's National Association of the Deaf suggests that nearly 18 million people—around 1% of India's population—are deaf. These figures highlight the urgent need for assistive technologies to enable seamless communication. Individuals with speech and hearing impairments rely on sign language as their primary mode of communication, yet most hearing individuals do not understand sign language, creating a significant communication barrier. AI-driven solutions, including computer vision, natural language processing (NLP), and deep learning, offer innovative ways to recognize, translate, and interpret sign language in real time. Various sign languages exist worldwide, such as American Sign Language (ASL), British Sign Language (BSL), French Sign Language (FSL), Indian Sign Language (ISL), and Japanese Sign Language (JSL). Extensive research has been conducted to develop recognition systems for different sign languages, enhancing accessibility and inclusivity.

By leveraging AI, machine learning, and gesture recognition, automated sign language translation systems can enable real-time communication between the deaf and hearing populations. These systems enhance accessibility, promote social inclusion, and ensure that individuals with hearing impairments can participate equally in society. However, challenges such as high implementation costs, diverse sign language variations, and the need for extensive datasets must be addressed. Governments, researchers, and tech companies must collaborate to develop cost-effective, scalable solutions to bridge this communication gap and improve human-computer interaction for the hearing-impaired community. Achieving high accuracy in real-time sign language recognition requires robust computer vision techniques capable of handling variations in lighting, hand orientation, and background noise. Additionally, integrating facial expressions and body posture into AI models is crucial, as they play a significant role in sign language communication. Overcoming these technical challenges will enable the development of seamless, real-time translation systems that bridge the communication gap between the deaf and hearing communities, fostering greater inclusivity and accessibility in everyday interactions.



## International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

### II. OBJECTIVES OF STUDY

The primary objective of this project is to develop an efficient, real-time, and user-friendly system that converts sign language into speech, enabling seamless communication between individuals with speech impairments and the hearing population. This system will use computer vision, artificial intelligence (AI), and machine learning to recognize hand gestures and translate them into text and speech output. By leveraging deep learning models, the system will improve accuracy over time and provide an effective means of interaction for those who rely on sign language.

Sign language is a crucial mode of communication for individuals with hearing and speech impairments. However, most people in society are not familiar with sign language, creating a communication barrier. This project aims to bridge this gap by developing an AI-powered system capable of recognizing and interpreting various sign language gestures. Through the use of gesture tracking and real-time processing, the system will facilitate better communication and promote inclusivity. The proposed system will employ camera-based gesture tracking to capture hand and finger movements. This data will be processed using AI algorithms to convert gestures into meaningful text output, which will then be transformed into speech using text-to-speech (TTS) technology. The entire process will be designed to operate in real-time, ensuring instantaneous and accurate communication. This solution will be particularly beneficial for deaf and mute individuals, as it provides an independent way for them to communicate with those who do not understand sign language. Additionally, the system will incorporate self-learning capabilities, allowing it to improve accuracy over time and adapt to different sign language variations. The objective is to make communication more accessible, efficient, and widely available through technological advancements.

### III. SCOPE

#### A. Target Users

The primary users of this system are individuals with speech and hearing impairments who rely on sign language for communication. This system will allow them to communicate more effectively with the hearing population without requiring an interpreter. It will also assist caregivers, educators, and professionals working with the deaf community in understanding sign language more efficiently.

#### B. Machine Learning and Self-Learning Capabilities

The system will integrate machine learning algorithms that enable it to learn from user interactions and improve over time. By continuously analyzing and adapting to different hand gestures and finger movements, the system will enhance its accuracy and effectiveness. This self-learning capability ensures that the system remains relevant and adaptable to different sign language variations.

#### C. Real-Time Assistance in Daily Life

The system is designed to assist users in their daily interactions by providing real-time sign language translation. Whether in educational institutions, workplaces, or public spaces, users will be able to communicate more effectively. This will help them gain more independence, reduce communication barriers, and enhance their overall quality of life.

#### D. Mobile Compatibility and Accessibility

The system will be developed to function seamlessly on Android smartphones, making it accessible to a larger audience. Mobile compatibility ensures that users can carry the tool with them wherever they go, providing instant access to sign language translation. Additionally, efforts will be made to ensure the system remains cost-effective so that it is affordable for a wide range of users.

### IV. LITERATURE REVIEW

Sign language recognition has been a key area of research in the field of human-computer interaction (HCI) and assistive technologies. Over the years, numerous methods and technologies have been proposed to bridge the communication gap between the hearing-impaired and the general population. Traditional approaches relied on wearable sensor-based systems, whereas recent advancements in artificial intelligence (AI), machine learning (ML), and computer vision have enabled more efficient and real-time recognition systems. Several studies have explored





## International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

different methodologies, including image processing, feature extraction, and deep learning algorithms, to improve the accuracy and efficiency of sign language recognition. While many systems have demonstrated high recognition accuracy, challenges such as real-time implementation, computational complexity, and variation in sign languages across different regions remain significant hurdles. This literature review critically examines existing research on sign language recognition, highlighting key contributions, methodologies, and challenges in the field.

### A. Sensor-Based and Wearable Systems for Sign Language Recognition

Glove-based systems use embedded sensors, accelerometers, and microcontrollers to detect hand gestures. Wearable devices map hand and finger movements to text or speech output. Examples include *"Deaf Mute Communication Interpreter"*, which explores wearable communication devices like gloves, keypads, and touchscreens. Similarly, *"Smart Glove with Gesture Recognition Ability"* uses bend sensors, accelerometers, and Hall Effect sensors for gesture recognition.

### B. Image Processing and Computer Vision Techniques

Image processing techniques such as skin color segmentation, histogram matching, and contour where  $m$  is the number of hops in the route,  $TE = TE_{node}$  is the transmission energy between the nodes. The route having minimum total transmission energy i.e.  $\min(TTE_R)$  will be selected as energy efficient route detection play a crucial role in sign language recognition. Feature extraction methods like Scale Invariant Feature Transform (SIFT) and Principal Component Analysis (PCA) help improve recognition accuracy. For example, *"Hand Gesture Recognition System for Dumb People"* utilizes SIFT for feature extraction in static hand gesture recognition, while *"Vision-Based Hand Gesture Recognition Using Dynamic Time Warping"* employs motion-based recognition methods to enhance accuracy.

### C. Machine Learning and Deep Learning Approaches

Machine learning models such as Support Vector Machines (SVM), Artificial Neural Networks (ANN), and Convolutional Neural Networks (CNN) have been widely used to improve the accuracy and adaptability of sign language recognition. AI-driven models enable real-time recognition, making communication more efficient. Studies like *"Hand Gesture Recognition for Sign Language Recognition"* explore different machine learning classification techniques, whereas *"Hand Gesture Recognition of English Alphabets Using Artificial Neural Network"* achieves a recognition accuracy of 92.50%, showcasing the potential of deep learning in gesture recognition.

### D. Mobile and Real-Time Sign Language Recognition

With the rise of mobile applications, sign language translation has become more portable and accessible. AI-powered mobile applications allow real-time gesture recognition and conversion into text or speech, making communication easier for individuals with hearing and speech impairments. For instance, *"SignPro-An Application Suite for Deaf and Dumb"* focuses on developing a mobile application for real-time gesture-to-text conversion. Similarly, *"An Automated System for Indian Sign Language Recognition"* utilizes neural networks and Otsu's thresholding for improved real-time recognition accuracy.

### E. Challenges and Future Directions

Despite significant progress in sign language recognition, several challenges remain, including high computational costs, the need for extensive training datasets, and variations in sign language across different regions. Many studies emphasize the necessity for large-scale multilingual sign language datasets to improve model generalization. Future research should focus on improving accuracy under real-world conditions by integrating multi-modal approaches that combine vision-based recognition with wearable sensor technology. Additionally, AI ethics and data privacy concerns must be addressed to ensure responsible deployment of sign language recognition systems. By critically evaluating existing research, it becomes evident that integrating AI, computer vision, and real-time processing can significantly enhance communication accessibility for the deaf and mute community. Future advancements in AI, cloud computing, and mobile technologies will further drive the development of highly efficient and user-friendly sign language translation systems.



## International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

### V. METHODOLOGY

The development of a vision-based sign language recognition system involves multiple phases, including feature extraction, artificial neural networks, and deep learning models like Convolutional Neural Networks (CNNs). The project integrates various technologies such as TensorFlow, Keras, and OpenCV to process hand gestures efficiently. The workflow follows a structured approach, beginning with project conceptualization, technology selection, front-end and back-end development, and culminating in system testing. This methodology ensures that the proposed system effectively translates hand gestures into meaningful text or speech, improving accessibility for deaf and mute individuals.

### VI. FEATURE EXTRACTION AND REPRESENTATION

Feature extraction involves representing an image as a three-dimensional matrix where pixel values encode depth information. CNN-based models utilize these pixel values to extract significant features that help in recognizing hand gestures accurately.

Artificial Neural Networks:

Artificial Neural Network is a connection of neurons, replicating the structure of human brain. Each connection of neuron transfers information to another neuron

.Inputs are fed into first layer of neurons which processes it and transfers to another layer of neurons called as hidden layers. After processing of information through multiple layers of hidden layers, information is passed to final output layer.

Convolution Neural Network Unlike regular Neural Networks, in the layers of CNN, the neurons are arranged in 3 dimensions: width, height, depth. The neurons in a layer will only be connected to a small region of the layer (window size) before it, instead of all of the neurons in a fully-connected manner. Moreover, the final output layer would have dimensions (number of classes), because by the end of the CNN architecture we will reduce the full image into a single vector of class scores.

Convolution layer In convolution layer we take a small window size [typically of length 5\*5] that extends to the depth of the input matrix. The layer consists of learnable filters of window size. During every iteration we slid the window by stride size [typically 1], and compute the dot product of filter entries and input values at a given position. As we continue this process well create a 2- Dimensional activation matrix that gives the response of that matrix at every spatial position. That is, the network will learn filters that activate when they see some type of visual feature such as an edge of some orientation.

### VII. IMPLEMENTATION DETAILS HARDWARE AND SOFTWARE SPECIFICATIONS

To ensure optimal performance for real-time sign language recognition, the system was implemented using a combination of high-performance hardware and efficient software frameworks.

Pooling Layer We use pooling layer to decrease the size of activation matrix and ultimately reduce the learnable parameters. There are two types of pooling: a. Max Pooling In max pooling we take a window size [for example window of size 2\*2], and only take the maximum of 4 values. Well lid this window and continue this process, so well finally get an activation matrix half of its original Size. b. Average Pooling In average pooling we take average of all values in a window.

### MODEL TRAINING AND OPTIMIZATION TECHNIQUES

To achieve high accuracy in sign language recognition, deep learning models were trained on a diverse dataset of hand gestures. The following techniques were employed to enhance performance



## International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

### VIII. CHALLENGES FACED DURING IMPLEMENTATION

#### A. Variability in Hand Gestures

- Differences in hand sizes, orientations, and skin tones impacted model accuracy.
- Solution: Data augmentation and adaptive thresholding in OpenCV.

#### B. Real-Time Processing Constraints

- Processing video frames in real-time required high computational power.
- Solution: GPU acceleration and model quantization for mobile deployment.

#### C. Lighting and Background Interference

- Changing lighting conditions affected hand detection.
- Solution: Adaptive histogram equalization and background subtraction techniques.

#### D. Complexity in Continuous Gesture Recognition

- Recognizing fluid, dynamic gestures was more challenging than static signs.
- Solution: Implementing LSTM for sequential hand movement analysis.

#### E. Limited Dataset for Rare Signs

- Some sign languages had limited labeled datasets available.
- Solution: Data augmentation and synthetic data generation using GANs (Generative Adversarial Networks).

### REFERENCES

1. J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Miami, FL, USA, 2009, pp. 248-255.
2. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.
3. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436-444, May 2015.
4. A. Graves, A. Mohamed, and G. Hinton, "Speech recognition with deep recurrent neural networks," *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Vancouver, BC, Canada, 2013, pp. 6645- 6649.
5. M. R. Islam, M. A. Hossain, and S. Rahman, "Hand gesture recognition using deep learning for sign language translation," *IEEE Access*, vol. 8, pp. 129797-129809, 2020.
6. P. Molchanov, S. Gupta, K. Kim, and J. Kautz, "Hand gesture recognition with 3D convolutional neural networks," *Proc. IEEE Conf. Comput. Vis. Workshops (CVPRW)*, Las Vegas, NV, USA, 2015, pp. 1-7.
7. J. Wang, Y. Liu, Y. Wu, and J. Yuan, "Learning actionlet ensemble for 3D human action recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 5, pp. 914-927, May 2014.
8. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Adv. Neural Inf. Process. Syst. (NIPS)*, vol. 25, pp. 1097-1105.





INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

 9940 572 462  6381 907 438  [ijircce@gmail.com](mailto:ijircce@gmail.com)



[www.ijircce.com](http://www.ijircce.com)

Scan to save the contact details