International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 10 Issue: 02 | Feb - 2026 SJIF Rating: 8.659 ISSN: 2582-3930

"Helping Hearing Impaired in Problematic Situation Using Deep Learning"

Mr.S.K.Mahajan , Ishan V. Pawar, Jagruti V. Patil , Harshada S. Patil , Akshara S. Patil

Department of Computer Technology
K. K. Wagh Polytechnic, Nashik, India

Abstract - Abstract - The "Hand Sign Detection Using Deep Learning Approach" project aims to develop a robust, real-time hand gesture recognition system capable of detecting and classifying static hand signs with high accuracy. The system leverages advanced computer vision techniques combined with deep learning models to process live video streams or uploaded images. By integrating MediaPipe for fast hand detection, OpenCV for preprocessing, and TensorFlow/Keras for inference, the system can operate reliably under varied conditions. The trained model will be stored in multiple formats - .hdf5, .keras, and .tflite - to ensure compatibility across platforms and deployment environments.The deliverable is a fully hosted, interactive website where users can perform gestures in front of their webcam or upload an image to receive real-time predictions. Backend services will be implemented in Python with Flask or FastAPI , and the frontend will be lightweight, responsive, and device-independent. This approach enables broad accessibility, with potential applications in accessibility solutions, educational tools, and hands-free interfaces.

keywords:
Hand Sign Detection, Open CV, Computer Vision, Mediapipe, Keras, Tester Flow Lite (TF Lite), Image Classification, Live Video Stream.

Introduction

Hand gestures are one of the most natural and intuitive forms of non-verbal communication, playing a vital role in human interaction. They allow individuals to convey emotions, intentions, and instructions without relying on spoken words. With rapid advancements in Artificial Intelligence (AI) and Computer Vision (CV), it has become possible to develop intelligent systems capable of detecting and recognizing hand gestures in real time. Such systems have wide-ranging applications, including sign language interpretation for the hearing-impaired, immersive gaming and virtual reality for enhanced user experiences, robotics for intuitive human–machine interaction, and remote collaboration tools that improve accessibility and communication. This project aims to design a real- time hand gesture recognition system based on a two-stage pipeline. In the first stage, Hand Detection, frameworks such as MediaPipe or OpenCV are employed to accurately identify and localize the hand region. In the second stage, Gesture Classification, a trained deep learning model is used to classify the detected hand into predefined categories. The proposed modular architecture not only ensures accuracy and efficiency but also allows for easy scalability, making it possible to incorporate dynamic gestures, multi-hand tracking, and continuous sign language

translation in future enhancement The current system focuses on real-time recognition of static hand gestures using a single hand as input. It is designed for web-based deployment, allowing users to access and interact with the system through a browser without specialized hardware. This stage establishes a strong foundation by demonstrating the feasibility of combining hand detection and deep learning classification to achieve efficient gesture recognition

1. Literature Survey

1) CNN Architectures: MobileNetV2, EfficientNet, and ResNet have proven effective in low-latency image classification tasks, especially for deployment on resource-limited devices.

2) MediaPipe Hands: Provides 21 3D landmarks for precise hand modeling, improving classification accuracy.

3) TensorFlow.js Deployments: Recent trends highlight running AI models in-browser for better accessibility.

4) Existing Works: Several sign recognition projects exist, but many lack optimized performance for online use or focus solely on offline desktop environments.

5) Datasets: Datasets such as the ASL Alphabet Dataset and Sign Language MNIST have been widely used to train hand sign classifiers.

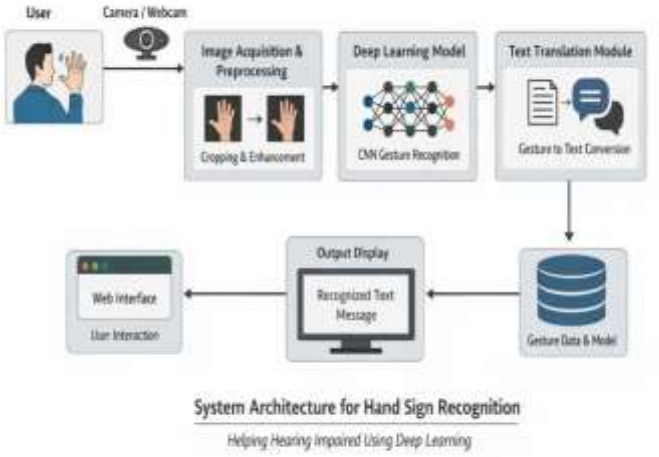
2. Problem Definition

Hearing-impaired individuals face significant communication challenges, especially during problematic or emergency situations. Traditional communication methods such as written notes or interpreters are often slow and impractical, while most people do not understand sign language.

Existing hand gesture recognition systems rely on basic image processing techniques that are sensitive to lighting, background, and hand orientation, leading to inaccurate results.

To overcome these limitations, an automated deep learning–based hand sign detection system is required to accurately recognize gestures in real time and assist hearing-impaired individuals by translating signs into meaningful outputs, thereby improving accessibility and communication.

4. Proposed Working



System Architecture for Hand Sign Recognition

Helping Hearing Impaired Using Deep Learning

The proposed system utilizes deep learning and computer vision techniques to develop an automated hand sign recognition solution that assists hearing-impaired individuals during problematic situations. The system is designed to recognize hand gestures in real time and translate them into meaningful text output, enabling effective communication without the need for human interpreters.

The system consists of a camera module (webcam) that captures live video or static images

of hand gestures, a preprocessing module that enhances and normalizes the input data, and a trained deep learning model that performs gesture classification. The captured hand images are processed using computer vision techniques, and the extracted features are analyzed by a convolutional neural network (CNN) to accurately identify the corresponding hand sign.

The proposed system supports real-time gesture recognition through a web-based interface, allowing users to interact easily with the system. Once a hand sign is recognized, the system displays the interpreted message on the screen, providing a clear and understandable output for communication. The system is designed to be user- friendly, efficient, and reliable, ensuring accurate performance under varying lighting and background conditions.

Overall, the proposed system enhances accessibility, reduces communication barriers, and provides a practical solution to support hearing- impaired individuals in critical or problematic situations.

STEP-BY-STEP PROCESS:

1. Start the system:

The system is initialized by launching the web application, loading the required libraries, and initializing the webcam or camera module.

2. Capture hand gesture:

The user performs a hand sign in front of the webcam, and the live video feed is captured by the system.

3. Hand detection:

The captured video frame is processed using computer vision techniques to detect and isolate the hand region.

4. Preprocessing of input:

The detected hand image is resized, normalized, and enhanced to match the input requirements of the deep learning model.

5. Gesture classification:

The preprocessed image is passed to the trained deep learning (CNN) model, which analyzes the gesture and predicts the corresponding hand sign.

6. Gesture-to-text conversion:

The recognized hand sign is converted into meaningful text or a predefined message.

7. Display output:

The generated text message is displayed on the user interface, enabling others to understand the communication.

8. Continuous operation:

The system continues to process gestures in real time until the user stops the application.

INPUTS:

- Hand gesture input: Live hand signs captured using a webcam or camera device.
- Image/video data: Real-time video frames or uploaded images containing hand gestures.
- Trained dataset: Predefined gesture dataset used by the deep learning model for recognition.

OUTPUTS:

- Recognized text message: Converted text output corresponding to the detected hand sign.
- Visual display: Real-time display of recognized gestures on the web interface.
- System response: Clear and understandable output to assist communication during problematic situations.

5. Result

The Hand Sign Detection Using Deep Learning system successfully demonstrates real-time recognition of static hand gestures using computer vision and deep learning techniques. The project integrates MediaPipe-based hand detection, image preprocessing, and a trained deep learning model to accurately classify hand signs through a web- based interface. The system provides reliable results with minimal latency and a user-friendly interface, making it suitable for accessibility and human–computer interaction applications.

1. Secure Login System

- Separate login access is provided for Admin and Users.
- Authentication is performed using email and password validation.

- Unauthorized users are restricted from accessing system functionalities.

2. User Management (Admin Panel)

- The Admin can add, view, and delete users from the system.
- User credentials and roles are stored securely in the database.
- Enables controlled access to the gesture detection system.

3. Real-Time Hand Gesture Detection

- The system captures live video feed from the webcam.
- MediaPipe detects hand landmarks accurately in real time.
- Supports smooth and continuous gesture recognition.

4. Image Preprocessing and Feature Standardization

- Detected hand regions are resized and normalized.
- Ensures consistent input size for the deep learning model.
- Improves prediction accuracy and system reliability.

5. Deep Learning-Based Gesture Classification

- A trained CNN-based model classifies hand signs effectively.
- The system predicts the gesture label along with a confidence score.
- Supports multiple model formats such as .hdf5, .keras, and .tflite.

6. Result Visualization and User Interaction

- The predicted hand sign is displayed on the screen in real time.
- Confidence percentage is shown for better interpretability.
- Live video feed is annotated with prediction results.

7. Prediction History and Logging

- All predictions are stored in the database with timestamps.
- Enables users and administrators to view past detection results.
- Useful for performance analysis and future model improvements.

8. Database Design & Security

- Relational database tables store user details, gesture data, predictions, and logs.
- Proper constraints ensure data integrity and security.
- Prevents unauthorized data access and duplication.

9. Cost Estimation (COCOMO Model)

- The total project cost was estimated using COCOMO Model.
- The final cost including development, hardware, and travel expenses is ₹13650/-

Fig 1: Sign Up

Fig 2: Login

Fig 3: Dashboard

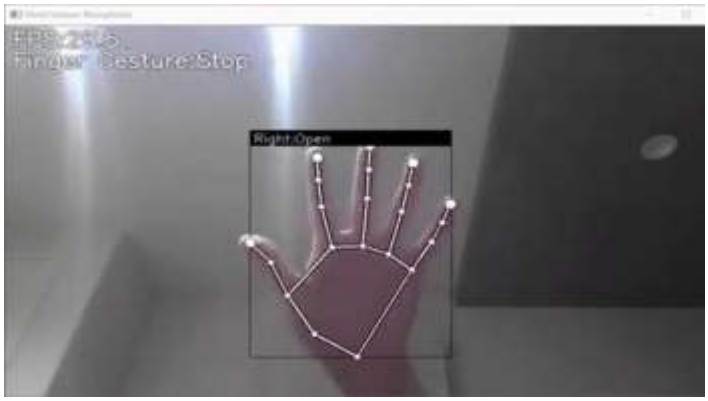


Fig 4: Dashboard

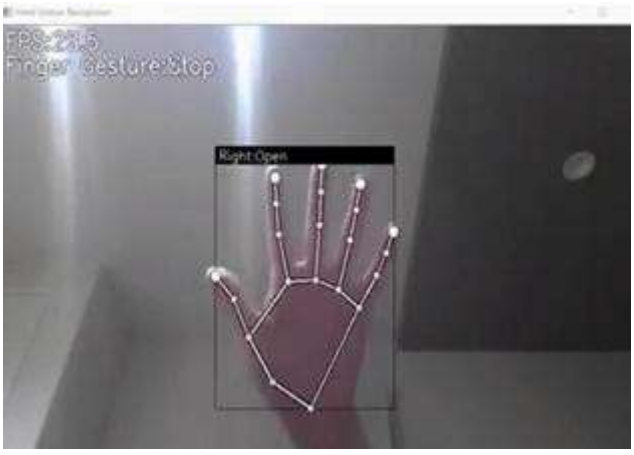


Fig 5: Prediction Output Panel

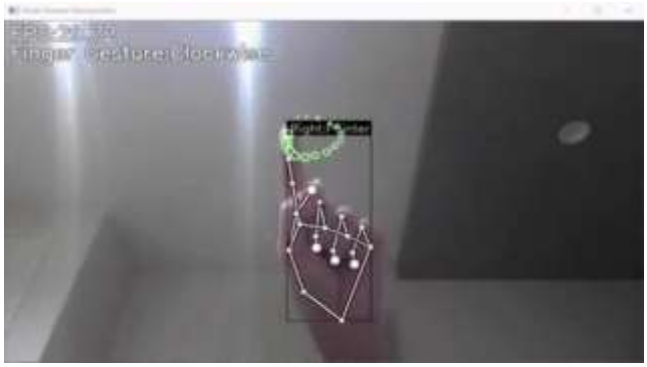


Fig 6: Prediction Output Panel

Fig 7: Predict Gesture Clockwise

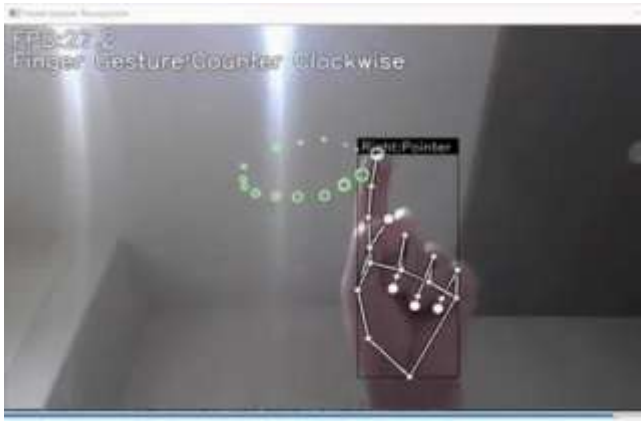


Fig 8: Predict Gesture Counter Clockwise

6. Conclusion

The development of sign language recognition systems using CNN architectures such as MobileNetV2, EfficientNet, and ResNet, combined with MediaPipe Hands for precise hand landmark detection and TensorFlow.js for browser-based deployment, demonstrates a significant step toward

assisting hearing-impaired individuals. These systems provide real-time and accurate gesture recognition, enabling communication in situations where human interpreters are unavailable, such as hospitals, schools, public places, or emergencies. By translating hand gestures into text or speech, they enhance independence, accessibility, and social inclusion for hearing-impaired users. Despite the advantages, challenges like lighting conditions, background clutter, limited gesture vocabulary, hand occlusion, and dataset limitations still affect system performance. However, the use of lightweight CNN models and efficient hand tracking frameworks ensures that these systems remain practical for web and mobile deployment, even on devices with limited computational resources. With further improvements in dataset diversity, model robustness, and real-world testing, such systems have the potential to bridge the communication gap and provide a cost-effective, reliable, and inclusive solution for daily life and critical scenarios, making a meaningful impact on the lives of hearing-impaired individuals.

7. Reference

1. Rheiner, J. “A High-Accuracy CPU-Real-Time American Sign Language Recognition System.” ScienceDirect, 2025.

2. Li, H. H. “Dynamic Hand Gesture Recognition Using MediaPipe and Deep Learning.” MDPI, 2025.

3. “A Deep Learning Framework for Ethiopian Sign Language Recognition.” Nature, 2025.

4. Cheriyan, J. “Optimization of Deep Learning Models for Real-Time Sign Language to Text Conversion.” Medium, 2025.

5. “Sign Language Recognition System Using TensorFlow Object Detection API.” ResearchGate, 2025.

6. Gil-Martín, M. “Hand Gesture Recognition Using MediaPipe Landmarks.” SCITEPRESS, 2025.

7. Najib, F. M. “Sign Language Interpretation Using Machine Learning and Image Processing.” Springer, 2025.

8. Mutalib, M. A., Islam, S., & Islam, M. “Hand Gesture Recognition Systems: A Review of Methods, Applications, and Challenges.” IJCA, 2025.