



International Journal of Research Publication and Reviews

Journal Homepage: www.ijrpr.com ISSN 2582-7421

Intelligent Hand Gesture Recognition System Using Machine Learning

Ankitha Marvalli Ramesh¹, Ganga Basvaraj Kakamari², Gousiya Noorain³, Madeeha Tazeen⁴, Dr. Chethan Chandra S Basavaraddi⁵

¹⁻⁴Department of Computer Science and Engineering, G M Institute of Technology Davanagere

⁵Guide, Associate Professor, Department of Computer Science and Engineering, GM University Davanagere

Email- ankithamarvalli518@gmail.com, gangakakamari2005@gmail.com, gousiyanoorain1@gmail.com, 1madeehatazeen@gmail.com, chethanchandrasb.cse@gmua.ac.in

ABSTRACT

Communication can become difficult for bedridden, elderly, or physically disabled patients who may have limited speech or mobility. This project presents an intelligent hand gesture recognition system that uses computer vision and deep learning to convert predefined hand gestures into meaningful communication messages. The system uses MediaPipe Hands to detect hand landmarks from video frames and represents each frame using normalized landmark coordinates. Temporal sequences of these features are processed using a Long Short-Term Memory (LSTM) network to learn movement patterns and classify dynamic gestures. An initial development dataset containing six emergency-related gesture classes—Call, Doctor, Help, Hot, Lose, and Pain—was processed to generate 126 features per frame for a two-hand representation. The sequences were resampled to a fixed length of 70 frames and used for LSTM training and evaluation. The current model achieved 96.67% accuracy on the held-out test set of 60 sequences. A real-time webcam prototype was also implemented to display the recognized gesture, confidence, patient message, hand detection status, and system information. The current results demonstrate the feasibility of the proposed approach. The present dataset is an initial development dataset rather than the final dataset, and future work will focus on a larger and more diverse dataset, additional gesture classes, and richer features.

Keywords: Hand Gesture Recognition, Indian Sign Language, MediaPipe, LSTM, Deep Learning, Computer Vision, Assistive Communication.

1. Introduction

Hand gesture recognition is a computer vision and machine learning technology that enables users to communicate or interact with digital systems through hand movements. It has applications in human-computer interaction, accessibility, healthcare assistance, and touchless communication. For patients who cannot communicate comfortably through speech or physical controls, a vision-based gesture interface can provide a simple and non-contact communication mechanism.

Dynamic gestures are different from static hand poses because their meaning depends not only on the shape of the hand but also on how the hand changes across time. Therefore, recognizing a gesture from a single image is insufficient for many real-world gestures. The proposed system addresses this by extracting hand landmarks frame by frame and using an LSTM model to learn temporal relationships across a sequence of frames.

The current implementation focuses on six emergency-oriented gestures: Call, Doctor, Help, Hot, Lose, and Pain. The recognized gesture is mapped to a simple patient-oriented message such as “I need help” or “I need a doctor.” The project is currently at the prototype and initial model-validation stage; the dataset and feature set will be expanded in the next development stage.

2. Literature Review

2.1 Hand Gesture Recognition Based on Deep Learning – Zhao et al. (2023)

Authors and Year: J. Zhao, X. H. Li, J. C. D. Cruz, M. S. Verdadero, J. C. Centeno and J. M. Novelero — 2023

Proposed a vision-based contactless gesture-recognition framework using Google MediaPipe to extract 21 three-dimensional hand landmarks. The study compared RNN, CNN, hybrid CNN–RNN, and Transformer Encoder architectures for gesture classification. Its use of landmark-based representation is relevant to the present project because landmark features reduce dependence on complete image backgrounds and provide a compact description of hand structure. The work also demonstrates the suitability of deep learning for real-time gesture recognition and HCI applications.

2.2 Using Machine Learning and Deep Learning for Static and Dynamic Hand Gesture Recognition – Sebbah and Chelali (2024)

Authors and Year: R. Sebbah and F. Z. Chelali A. Ellie — 2024

Sebbah and Chelali presented a comparative study covering both static and dynamic gestures. The work used ResNet18 for deep feature extraction and LGBPHS for machine-learning-based feature representation, with SVM and KNN used for classification. The reported comparison indicated that deep-learning-based approaches can provide more robust recognition than traditional machine-learning approaches for the studied datasets. This supports the use of a learned representation for the present dynamic gesture problem.

2.3 Exploring Models Approach for Real-Time Hand Gesture Recognition – Asih et al. (2025)

Authors and Year: N. Asih, B. S. Hantono and I. Hidayah — 2025

Asih et al. conducted a systematic literature review of real-time hand gesture recognition for HCI. The review examined models and algorithms for static and dynamic gestures and identified challenges including complex backgrounds, illumination changes, and variations in hand shape and size. It also highlighted research directions toward more accurate, robust, and efficient systems, including applications in assistive technologies. These observations are directly relevant to the proposed system because practical assistive communication requires reliable operation beyond a single controlled recording condition.

2.4 Analysis of Machine Learning for Recognizing Hand Gestures – Sharma et al. (2023)

Authors and Year: A. Kumar Sharma, A. S. Kumar, A. Nagpal, R. Singh, D. Reddy and A. Chadha — 2023

Sharma et al. reviewed artificial-intelligence and machine-learning techniques used for hand-gesture recognition in human-machine interaction. The study considered image-based and model-based approaches and discussed issues such as computational complexity, training time, and recognition performance. The work emphasizes the continuing need for methods that balance recognition quality with computational efficiency, which is important for real-time webcam applications.

2.5 Research Gap and Motivation

The reviewed studies show that MediaPipe landmark representations, deep learning, and temporal models are promising for hand gesture recognition. At the same time, the literature identifies recurring practical issues: environmental variation, differences between users, dynamic sequence handling, computational efficiency, and the need for real-time operation. The present project addresses these concerns by combining compact landmark features with normalization, temporal sequence resampling, and LSTM-based learning in a webcam-oriented assistive communication prototype. The project also treats dataset expansion and robustness testing as explicit future stages rather than assuming that a small development dataset is sufficient for general deployment.

3. Research Methodology

➤ Data Collection

Collect hand gesture video data representing emergency and communication gestures such as **Call, Doctor, Help, Hot, Lose, and Pain**.

Extract gesture sequences from the collected videos for training and testing the recognition model.

➤ Data Preprocessing

Process the videos frame by frame using Python and OpenCV. Detect hands using MediaPipe Hands and extract 21 hand landmarks with their 3D coordinates. Normalize the landmark coordinates to improve consistency.

➤ Feature Extraction and Sequence Preparation

Convert the extracted landmarks into feature vectors containing **126 features per frame** for two-hand input.

Resample the sequences to **70 frames** to provide a fixed-size input for the LSTM model.

➤ LSTM-Based Gesture Recognition

Implement an **LSTM neural network** to learn temporal patterns from hand landmark sequences.

Train the model to classify the input sequence into the corresponding gesture category.

➤ Real-Time Recognition

Integrate the trained model with a webcam-based application for real-time gesture recognition.

Continuously detect hand landmarks, process the sequence, and predict the corresponding gesture.

➤ **Assistive Communication Output**

Convert the recognized gesture into a meaningful **communication message** for the patient.

Display the detected gesture, confidence, and corresponding message through the system interface.

➤ **Evaluation and Analysis**

Evaluate the system using **accuracy, precision, recall, F1-score, loss, and confusion matrix**.

Analyze the recognition results and classification errors to improve the system's performance.

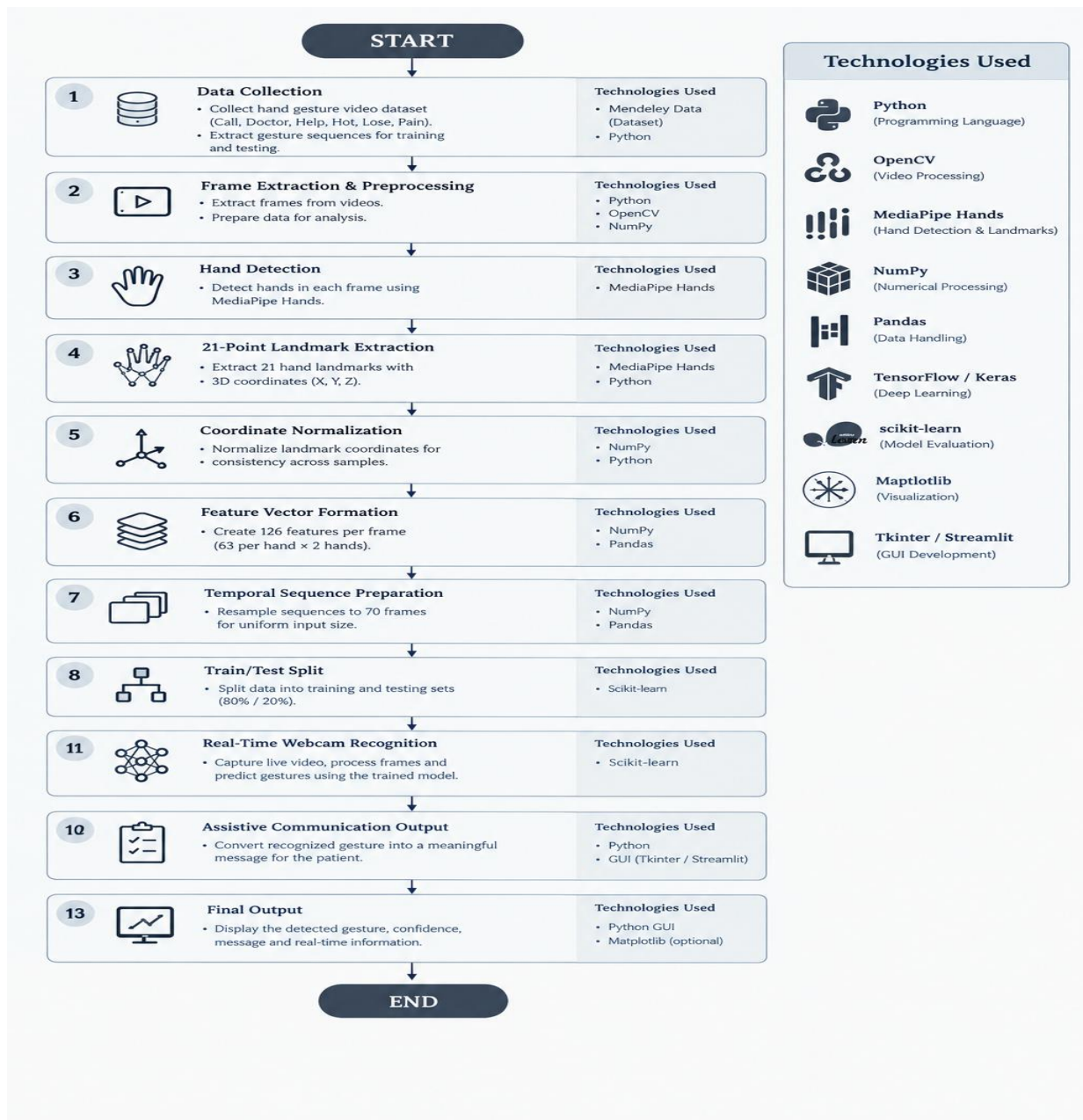


Fig 3.1 : Flowchart

4. Objectives of the Study

1. **To develop** a real-time hand gesture recognition system using MediaPipe and LSTM.
2. **To extract and normalize** hand landmarks from gesture video sequences for effective feature representation.

3. **To recognize dynamic hand gestures** by learning temporal movement patterns using an LSTM model.
4. **To provide assistive communication** by converting recognized gestures into meaningful messages.
5. **To evaluate and improve** the system based on accuracy, precision, recall, F1-score, and real-time performance.

5. Results and Discussions

The current development experiment uses 298 usable samples. After preprocessing and temporal resampling, 238 samples are used for training and 60 samples for testing. Each sample is represented as a 70×126 sequence.

Evaluation Item	Current Result
Training samples	238
Testing samples	60
Input sequence	70×126
Test loss	0.1587
Test accuracy	96.67%
Correct predictions	58 / 60
Main observed confusion	Doctor $\rightarrow$ Pain (2 samples)

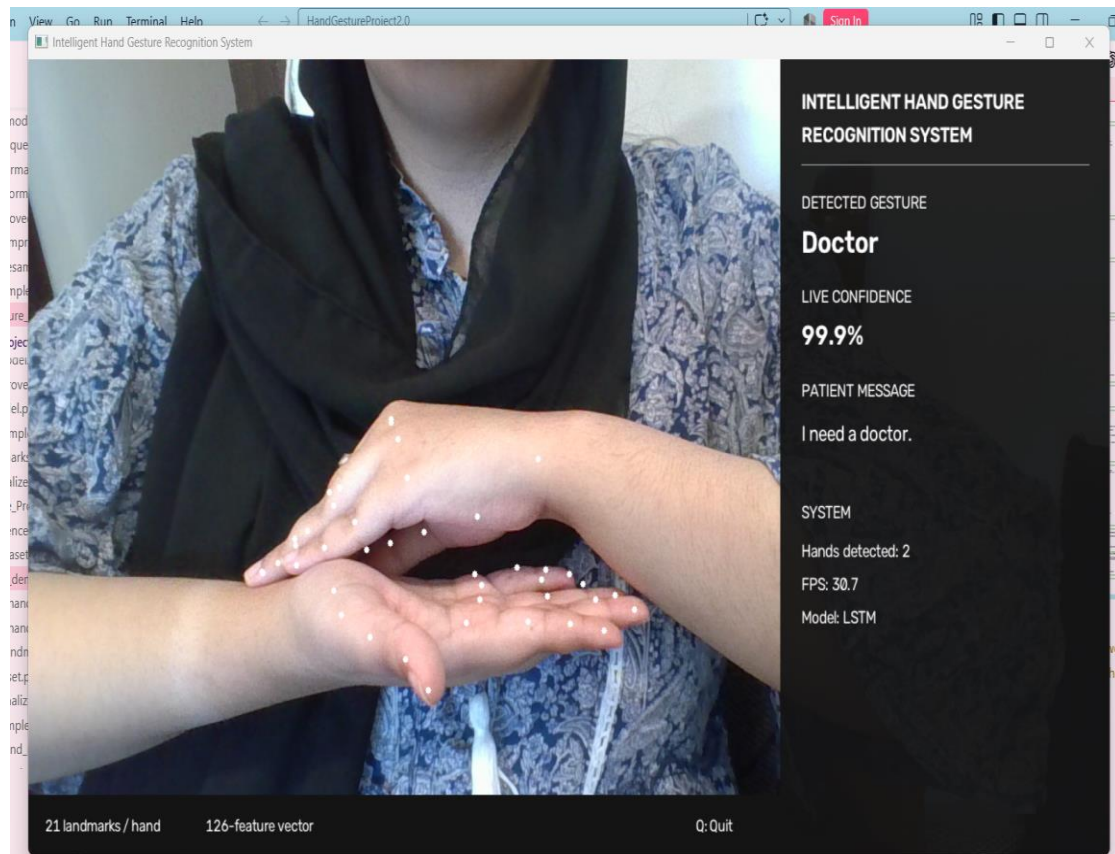


Fig 5.2: Real-Time System Output for “Doctor” Gesture

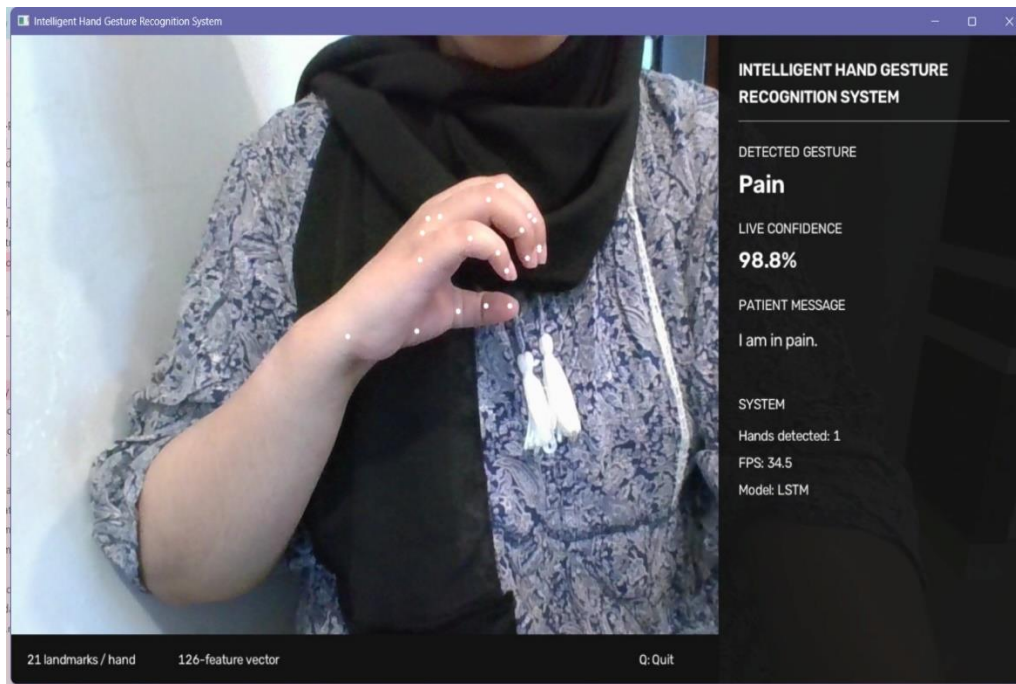


Fig 5.2: Real-Time System Output for “Pain” Gesture

The LSTM model achieved a **96.67% test accuracy**, showing that it was able to effectively learn the temporal patterns of different dynamic gestures. The real-time system successfully detected both **single-hand and two-hand gestures** using MediaPipe landmarks. It also provided high-confidence predictions and converted the recognized gestures into meaningful **patient communication messages**. The prediction stability mechanism helped maintain consistent outputs during continuous movement. Overall, the results demonstrate that our approach is effective for real-time gesture-based communication, while a larger and more diverse dataset can further improve its reliability and generalization.

6. Conclusion

This paper presents an Intelligent Hand Gesture Recognition System that combines computer vision, MediaPipe hand landmarks, normalized feature representation, temporal sequence processing, and LSTM-based classification for dynamic gesture recognition. The current prototype demonstrates a complete pipeline from video acquisition to gesture prediction and communication output. Using six emergency-related gesture classes and a 126-feature two-hand representation, the initial development model achieved 96.67% accuracy on a 60-sample test split. The result demonstrates the feasibility of the proposed approach for the current development setting, while the observed Doctor–Pain confusion and the limited dataset highlight the need for broader evaluation. The project therefore provides a foundation for a more scalable assistive communication system in which additional data, gesture classes, real-world testing, and voice output can be incorporated in future stages.

References

1. R. Sebbah and F. Z. Chelali, “Using Machine Learning and Deep Learning for Static and Dynamic Hand Gesture Recognition,” 2024 International Conference on Advances in Electrical and Communication Technologies (ICAECOT), 2024, pp. 1–6, doi: 10.1109/ICAECOT62402.2024.10828747.
2. J. Zhao, X. H. Li, J. C. D. Cruz, M. S. Verdadero, J. C. Centeno and J. M. Novellero, “Hand Gesture Recognition Based on Deep Learning,” 2023 International Conference on Digital Applications, Transformation & Economy (ICDATE), 2023, pp. 250–254, doi: 10.1109/ICDATE58146.2023.10248500.
3. A. Kumar Sharma, A. S. Kumar, A. Nagpal, R. Singh, D. Reddy and A. Chadha, “Analysis of Machine Learning for Recognizing Hand Gestures,” 2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), 2023, pp. 1844–1847, doi: 10.1109/ICACITE57410.2023.10182676.
4. M. Deepika, S. Choudhary, S. Kumar and K. Srinivas, “Machine Learning-Based Approach for Hand Gesture Recognition,” 2023 International Conference on Disruptive Technologies (ICDT), 2023, pp. 264–268, doi: 10.1109/ICDT57929.2023.10150843.
5. N. Asih, B. S. Hantono and I. Hidayah, “Exploring Models Approach for Real-Time Hand Gesture Recognition,” 2025 International Conference on Advancement in Data Science, E-learning and Information System (ICADEIS), 2025, pp. 1–7, doi: 10.1109/ICADEIS65852.2025.10933403.

-
6. A. Mohanarathinam, K. G. Dharani, R. Sangeetha, G. Aravindh and P. Sasikala, "Study on Hand Gesture Recognition by using Machine Learning," 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA), 2020, pp. 1498–1501, doi: 10.1109/ICECA49313.2020.9297513.
 7. F. Zhang et al., "MediaPipe Hands: On-device Real-time Hand Tracking," arXiv:2006.10214, 2020.
 8. S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.
 9. "A Video Dataset of the Hand Gestures of Indian Sign Language Words used in Emergency Situations," Mendeley Data, DOI: 10.17632/2vfdm42337.1.