iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: VIII Month of publication: August 2025

DOI: https://doi.org/10.22214/ijraset.2025.73557

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue VIII Aug 2025- Available at www.ijraset.com

American Sign Language Detection Using Machine Learning

Supriya Singh[1], Dr. V. Uma Rani[2]
[1]Post Graduate Student, M. Tech (SE) Department of Information Technology, Jawaharlal Nehru Technological University Hyderabad
[2]Professor & Head, Department of Information Technology, Jawaharlal Nehru Technological University Hyderabad

Abstract: This project introduces a system that helps deaf and mute individuals interact more easily by converting sign language into spoken or written words. It uses image processing and a deep learning model called Convolutional Neural Network (CNN) to identify hand gestures. The system analyzes each gesture through multiple layers to extract features and classify them accurately. This approach can improve communication in everyday situations and medical environments, and it can also be expanded into a mobile application for quick and wireless communication support. The human hand, being highly expressive, is frequently used not only for physical interaction but also for communication. For deaf and mute individuals, hand gestures form the foundation of sign language, which is essential for daily communication. Enabling computers to understand and interpret these gestures would mark a significant advancement in human-computer interaction. The development of such systems requires effective manipulation and processing of visual data

Keywords: Gesture Recognition, Convolutional Neural Network, LBP (Local Binary Pattern), Wireless Communication.

I. INTRODUCTION

This project focuses on developing a gesture recognition system through effective neural network training. The goal is to accurately identify gestures in real time, thereby offering a communication aid for individuals with speech or hearing impairments. With rapid advancements in technology, the human race continues to enjoy increasing levels of convenience. However, a significant segment of the population—those who are differently abled—often remain excluded from many of these benefits. . According to the WHO, more then 300people have this issue.This project presents one such effort: the development of a unified system to assist individuals who are blind, deaf, or mute in interacting seamlessly with others, regardless of their abilities. The long-term vision is to create a compact, mobile-compatible communication tool that enables real-time interaction among individuals with varying disabilities Our approach focuses on converting non-verbal communication—specifically sign language—into text or speech. Sign language is a structured visual language where gestures involving the hands, head, and other body parts convey meaningful information. In the proposed system, sign gestures captured as images are processed by a Convolutional Neural Network (CNN) model to interpret and translate the signs into spoken or written language.The system utilizes hidden layers within the neural network to perform both feature extraction and classification. These layers help enhance the relevant features identified from the input image. When a gesture image is provided, the system compares it against the trained dataset stored in the neural network's memory. If a match is found, the corresponding gesture is recognized, and an appropriate output is generated.

II. LITERATURE REVIEW

1) BMC Kumara, HS Nagendrasamy and RL Chinmayi new technique has been introduced to extract spatial features that effectively track the hand movements of the signer. Spatial features are then derived by utilizing both the local centroids of these components and the overall global centroid. To handle the variations in the same sign made by different signers at various times, the concept of interval-valued symbolic data is employed. To assess the performance comprehensive experiments were carried out on a substantial dataset of sign videos compiled during the research process.
2) S Shivashankara and S Srinath both worked together on American Sign Language Recognition (ASLR) predict the action and gives the output.This paper reviews and analyzes multiple techniques developed to identify both static and dynamic ASL gestures. It presents a comparative study of various recognition methods, outlines key challenges in gesture recognition, and summarizes different approaches explored by researchers. Additionally, it includes visual comparisons and highlights the strengths and limitations of each method.

3) Rajesh George Rajan, M Judith Leo found that Sign language has become an essential mode of communication for individuals with hearing impairments, with sign alphabets serving as its foundation. This study employs a hybrid approach combining handcrafted features and deep learning to enhance sign classification accuracy. Skin color segmentation using the YCbCr model and Local Binary Patterns (LBP) are used to capture shape and texture details. Additionally, a pre-trained VGG-19 network is fine-tuned to extract deep features, which are then fused with the handcrafted features using serial fusion. The final feature set is classified using a Support Vector Machine (SVM) model.
4) Sanil Jain and K.V. Sameer Raja explored Indian Sign Language recognition by utilizing color images as input data. They applied various feature extraction techniques including Bag of Visual Words, Gaussian Random Features, and Histogram of Oriented Gradients (HoG). The system was trained using data from three individuals and evaluated on a different user entirely, achieving a recognition accuracy of 54.63%, indicating moderate success in cross-user generalization
5) Haitham Hasan and S. Abdul-Kareem introduced a hand gesture recognition method that leverages shape analysis. Their system focused on classifying six distinct static gestures — such as open, close, cut, paste, maximize, and minimize — using a neural network model. They employed a multilayer perceptron (MLP) architecture trained with the backpropagation algorithm. Their proposed approach achieved recognition accuracy of 86.38%, demonstrating strong performance in gesture classification.

III.EXISTING SYSTEM

Current communication methods for the hearing and speech impaired primarily rely on sign language interpreted by human translators or basic software tools with limited vocabulary. Many existing systems focus on static image-based gesture recognition, which often lacks accuracy in real-time applications and fails to detect dynamic hand movements effectively. These approaches also depend heavily on controlled environments, fixed backgrounds, or wearable sensors, which restrict their usability in real-world conditions. Additionally, most systems are not scalable and struggle to generalize across different users or lighting conditions.

A. *Disadvantages of the Existing System*

1) Hardware Dependency: Many systems require external devices like gloves or depth sensors, increasing cost and reducing portability.
2) Limited Gesture Recognition: Most existing models can only recognize a limited set of static gestures, lacking support for dynamic or sentence-level gestures.
3) Sensitivity to Conditions: Performance often drops with changes in lighting, background, hand size, or gesture orientation.
4) User Dependency: Accuracy may vary significantly from one user to another, especially when models are not trained on diverse datasets.
5) Low Real-Time Performance: Some systems have high processing time, making real-time communication inefficient or impractical.

IV.PROPOSED SYSTEM

The proposed system aims to develop a real-time, camera-based sign language recognition model that converts hand gestures into readable text or speech. By using convolutional neural networks (CNN), the system will identify and classify hand gestures from input images without relying on external devices or sensors. It will be trained on a dataset of sign language alphabets to ensure accuracy across various lighting conditions and hand shapes. This system enhances communication between hearing-impaired individuals and the general public, providing a portable and user-friendly solution that works effectively in real-world scenarios

A. *Advantages of the Proposed System*

1) Device-Free Interaction: Uses computer vision techniques without requiring external hardware like gloves or sensors, making it more user-friendly and cost-effective.
2) Supports Real-Time Recognition: Capable of recognizing hand gestures instantly, facilitating smooth and quick communication.
3) Deep Learning Integration: Utilizes Convolutional Neural Networks (CNNs) for accurate gesture classification and feature extraction.
4) User-Independent Design: Designed to work effectively across different users regardless of hand shape, size, or skin tone.
5) Enhanced Accuracy: Combines both hand-crafted features and deep learning to achieve better accuracy in sign detection and interpretation.
6) Assistive Technology: Acts as a valuable communication bridge between hearing-impaired individuals and the general public.
7) Scalable Architecture: Easy to expand and train for new gestures or additional languages over time.
8) Multimodal Output Support: Converts signs into both text and speech, offering dual-mode communication.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue VIII Aug 2025- Available at www.ijraset.com

V. MODEL

The proposed system operates through five structured modules to accurately interpret hand gestures from input images, as illustrated in

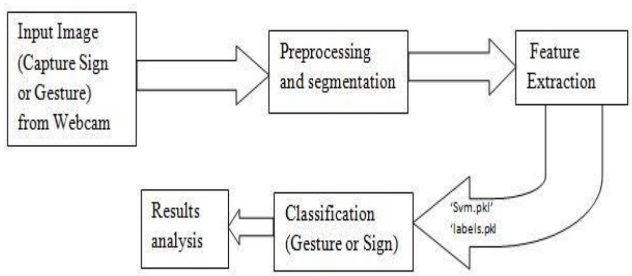


Figure No 1: Model

A. *Input Image Capture (Gesture Acquisition)*

The system utilizes either a built-in laptop webcam or an external camera to capture hand gesture images. Capturing high-resolution visuals is crucial to maintain image clarity, which plays a key role in ensuring precise gesture recognition.

B. *Preprocessing and Segmentation*

Preprocessing gives the image best quality. The image undergoes filtering and smoothing, followed by segmentation to detect the largest skin-colored region, typically the hand. This helps eliminate background noise and irrelevant skin-like regions.

C. *Feature Extraction*

Key features that uniquely define each gesture are extracted from the segmented hand image. These features include shape, texture, and spatial data. Extracted features are stored in a serialized file (svm.pkl) while the corresponding gesture labels are stored in labels.pkl, based on the training and test datasets.

D. *Classification*

The classification module uses machine learning (SVM) to match live input gestures with trained patterns. A binary sequence is derived from the features—significant peaks are encoded as '1', and insignificant ones as '0'. These are then matched against the trained model to recognize the correct gesture.

E. *Result Analysis*

The system performs real-time gesture recognition with high accuracy (up to 97%). During live detection, the system compares the input gesture with stored feature data and displays the corresponding result upon a successful match.

VI. SYSTEM ARCHITECTURE

The architectural design phase plays a crucial role in establishing the foundational structure of any software system. It begins with identifying the main components or modules of the system and determining how these parts interact and exchange information. This involves defining the responsibilities of each subsystem and outlining how they communicate to ensure seamless functionality. The goal is to create a blueprint that guides the development process and ensures that each component integrates well within the overall framework.

In this project, the proposed system architecture defines how various modules—such as image input, preprocessing, feature extraction, classification, and result display—work together in a coordinated manner. The architecture provides clarity on the flow of data between these modules, illustrating how raw input is transformed step-by-step into meaningful output. The diagram below presents a high-level overview of the system's structure and its core operational logic, helping to visualize the internal processes and their interactions.

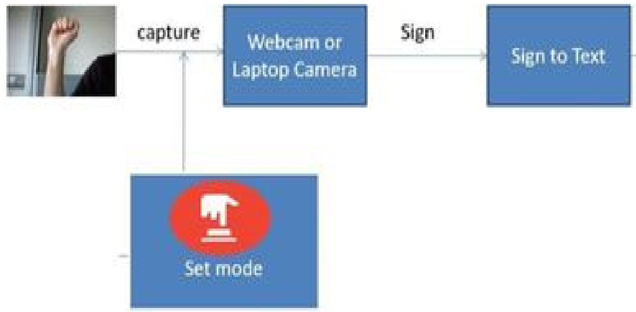


Figure No 2: System Architecture

VII. RESULTS

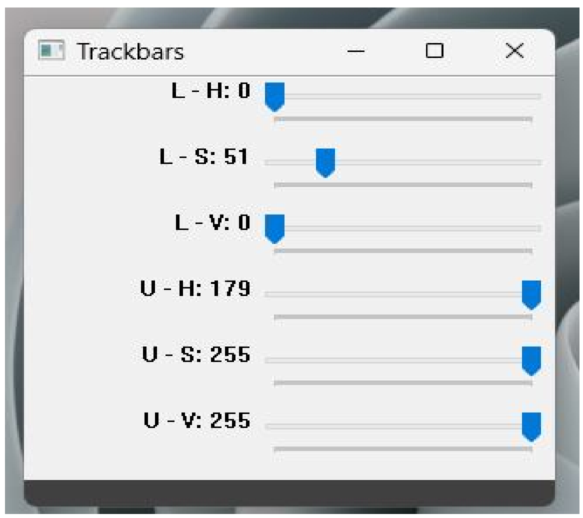


Figure No 3: Ajusting the task bar

Figure No 4: Mask Screen

Figure No 5: Detection of Alphabet

The system identifies the unknown sign language gesture based on the highest prediction probability among all trained classes. For instance, when the gesture corresponding to the letter "O" is given as input, the model processes the image and returns a probability distribution over all possible gestures. The picture with the highest probability is selected as the final result.

VIII. CONCLUSION

Communication plays a vital role in daily life. However, individuals with hearing and speech impairments face considerable challenges when interacting with others who are not familiar with sign language.Hand gestures and sign language are most commonly used by the deaf and mute community. This project presents a step toward bridging the communication gap between people with such impairments and those unfamiliar with sign language. By using Convolutional Neural Networks (CNNs), the system successfully recognizes hand gestures corresponding to English alphabets with commendable accuracy.

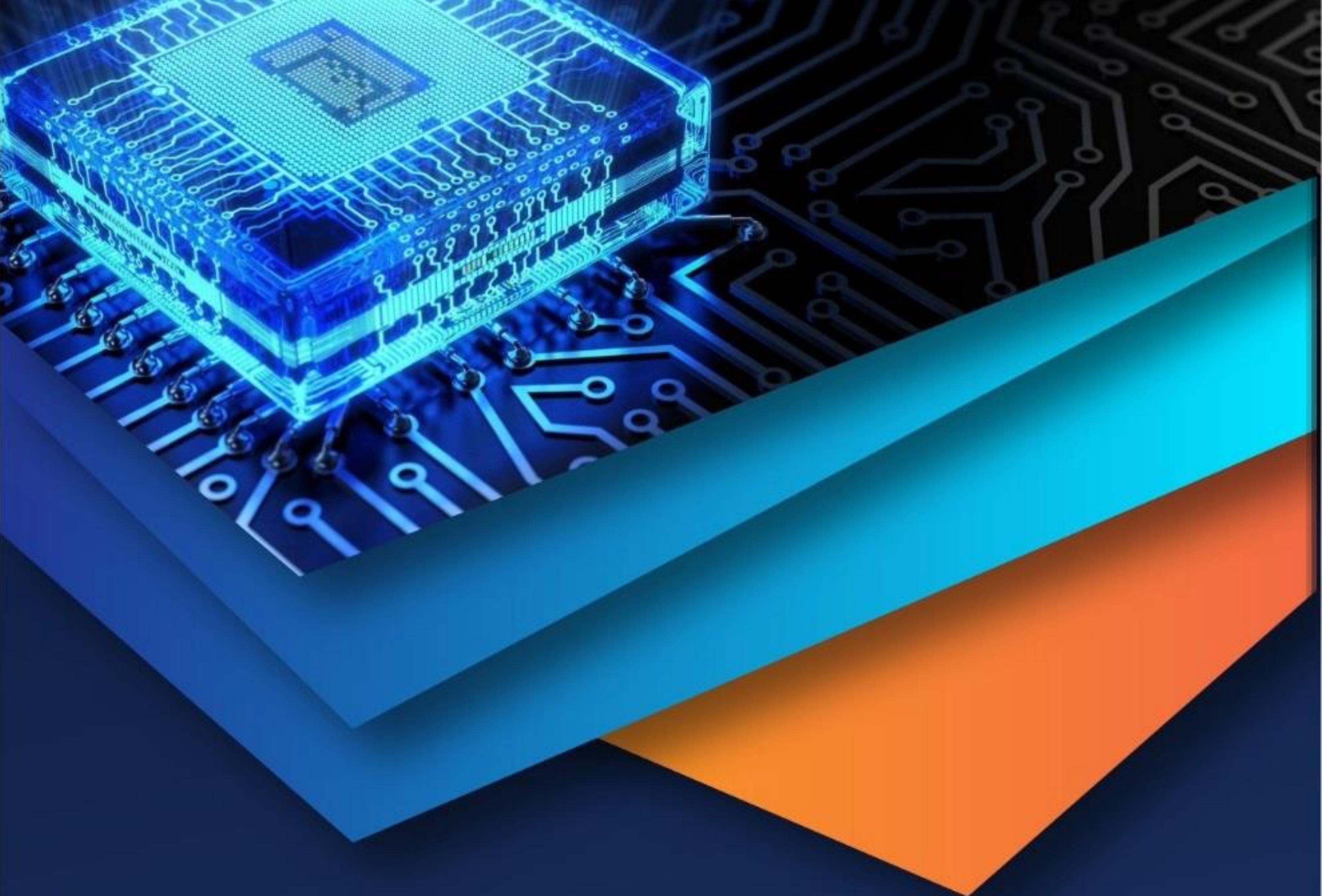
The implemented model demonstrates that machine learning, especially deep learning techniques like CNNs, can be effectively used to identify sign language gestures, making communication more inclusive and accessible for everyone. Future Scope: In real-world applications, sign language recognition needs to operate on live video streams rather than static images. Future development of this system can focus on incorporating real-time video processing, where continuous hand movements are captured frame by frame for gesture detection and interpretation.

Key enhancements for future implementation include:

1) Robust Background Handling: The system should be made adaptable to different lighting conditions, backgrounds, and hand orientations to ensure accurate recognition in varied environments.
2) Gesture Localization and Face Positioning: Improving the model to dynamically detect the position of hands in relation to the face will enhance precision in complex gestures and reduce misinterpretations.
3) Reduction in Processing Time: Optimization techniques can be applied to minimize latency, enabling faster gesture recognition suitable for real-time communication.
4) Mobile Application Integration: A user-friendly mobile app can be

REFERENCES

[1] O'reilly, Learning OpenCV, Computer Vision in C++ with the OpenCV Library, Adrian Kaebler and Gary Bradski.
[2] Thresholdinghttp://www.cse.unr.edu/~bebis/CS791E/Not es/Thresholding.pdf.
[3] AsanterabiMalima, ErolOzgur, and Mujdat Cetin, "A Fast Algorithm for Vision-Based Hand Gesture Recognition for Robot Control", IEEE International Conference on Computer Vision.
[4] Chaudhary, J.L. Raheja, K. Das, "A Vision based Real Time System to Control Remote Robotic hand Fingers", In proceedings of the IEEE International Conference on Computer Control and Automation, South Korea, 1-3 May, 2011, pp. 118-122
[5] A.Chaudhary, J.L. Raheja, "ABHIVYAKTI: A Vision Based Intelligent System for Elder and Sick Persons", In proceedings of the 3rd IEEE International Conference on Machine Vision, Hong Kong, 28-30 Dec, 2010, pp. 361-364. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep Face Recognition," British Machine Vision Conference (BMVC), pp. 1–12, 2015.
[6] J.Rekha, J. Bhattacharya and S. Majumder, "Hand Gesture Recognition for Sign Language: A New Hybrid Approach", 15th IPCV'11, July, 2011, Nevada USA.
[7] http://en.wikipedia.org/wiki/Principal_component_analysis.
[8] https://chatgpt.com/
[9] https://www.researchgate.net/publication/326972551_American_Sign_Language_Recognition_System_An_Optimal_Approch.
[10] Hemakasiny Visuwalingam, Ratnasingam Sakuntharaj, Janaka Alawatugoda, Roshan Ragel. "Deep Learning Model for Tamil Part-of-Speech Tagging", The Computer Journal, 2024.
[11] A.W. Fitzgibbon, and J. Lockton, "Hand Gesture Recognition Using Computer Vision", BSc. Graduation Project, Oxford University.
[12] T.S. Huang, and V.I. Pavlovic, "Hand gesture modeling, analysis and synthesis", Proceedings of international workshop on automatic face and gesture recognition, 1995, pp.73-79.



10.22214/IJRASET

INDEX COPERNICUS 45.98

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY