



International Journal of Innovative Research in Computer and Communication Engineering

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





Vision-Based Gesture Interaction Without Peripheral Devices

Dr. R. Jothilakshmi¹, E. Dharshini², S. Dharshini³, V. Swathi⁴

Associate Professor, Department of Information Technology, R.M.D. Engineering College, Chennai,
Tamil Nadu, India¹

Student, Department of Information Technology, R.M.D. Engineering College, Chennai, Tamil Nadu, India²

Student, Department of Information Technology, R.M.D. Engineering College, Chennai, Tamil Nadu, India³

Student, Department of Information Technology, R.M.D. Engineering College, Chennai, Tamil Nadu, India⁴

ABSTRACT: The Vision-Based Hand Gesture Interaction System is designed to enable touch-free human-computer interaction by replacing traditional input devices such as mouse and keyboard with hand gestures. The system uses a standard RGB camera to capture real-time hand movements and applies computer vision techniques using OpenCV and MediaPipe to detect and track hand landmarks accurately. Recognized gestures are processed through predefined logic or machine learning models and mapped to system commands such as cursor movement, clicking, scrolling, brightness control, and volume adjustment. The system ensures smooth real-time performance with minimal delay, providing a user-friendly and efficient interaction experience. Designed with accessibility and innovation in mind, this solution supports contactless interaction, enhances usability for physically challenged individuals, and reduces dependency on hardware peripherals, thereby improving the overall human-computer interaction process.

KEYWORDS: Hand Gesture Recognition, Vision- Based Interaction, Human-Computer Interaction, OpenCV, MediaPipe, Real-Time Hand Tracking, Computer Vision, Touchless Control System, Gesture-Based Cursor Control, Machine Learning, Assistive Technology, Smart Interaction System.

I. INTRODUCTION

Traditional human-computer interaction relies heavily on physical input devices such as keyboards, mice, and touchscreens. While effective, these devices limit natural interaction and may not be convenient in all environments, especially in situations requiring touchless control or assistive support. Moreover, individuals with physical disabilities may face challenges while using conventional input peripherals.

The Vision-Based Hand Gesture Interaction System aims to overcome these limitations by providing a camera-based, touch-free solution for controlling computer operations. The system captures real-time hand movements using a standard RGB camera and processes them using computer vision techniques. By detecting hand landmarks and recognizing predefined gestures, the system translates gestures into system commands such as cursor movement, clicking, scrolling, brightness control, and volume adjustment.

By integrating automation and real-time processing, this system enhances human-computer interaction, improves accessibility, and promotes a hygienic, contactless method of device control suitable for modern smart environments.

1.1 Problem Statement

Traditional human-computer interaction relies on physical input devices such as mouse, keyboard, and touchscreens, which limit natural and intuitive interaction between users and computers. These devices may not be convenient in all environments, particularly in situations requiring touchless control or assistive support for physically challenged individuals. Existing gesture recognition systems often depend on specialized hardware such as data gloves or sensors, which increases cost and reduces accessibility. Additionally, many existing systems suffer from low accuracy in different lighting conditions and delays in real-time response. Therefore, there is a need for a cost-effective and accurate vision-based gesture recognition system that enables efficient and contactless human-computer interaction using a standard camera.



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

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

1.2 Objective of the Study

The primary objective of this study is to design and develop a Vision-Based Hand Gesture Interaction System that enables efficient touchless control and ensures real-time human-computer interaction.

The specific objectives are:

1. To develop a user-friendly system that allows users to control computer operations using hand gestures in real time.
2. To implement accurate hand detection and tracking using OpenCV and MediaPipe for precise gesture recognition.
3. To recognize and classify predefined hand gestures using image processing techniques.
4. To map recognized gestures to system commands such as cursor movement, clicking, scrolling, brightness control, and volume adjustment.
5. To ensure smooth real-time performance with minimal delay for efficient interaction.
6. To enhance accessibility and promote contactless interaction by reducing dependency on traditional input devices.

II. LITERATURE REVIEW

The Vision-Based Hand Gesture Interaction System is designed to address the challenges in touchless human-computer interaction by eliminating the need for peripheral devices such as a mouse and keyboard. This section reviews existing research, technologies, and methodologies related to computer vision, hand gesture recognition, real-time image processing, and gesture-based system control.

Mukhopadhyay et al. [1] proposed a hand gesture recognition system using OpenCV and MediaPipe to detect and track hand landmarks in real time. The system calculates the number of fingers raised and maps gestures to actions, providing visual feedback and even sending messages to a registered phone number. The proposed method achieved about 90% accuracy, showing that it is simple, efficient, and suitable for practical gesture-based applications.

Abdirahman Osman Hashi et al. [2] presented a comprehensive systematic review of hand gesture recognition methods from 2018 to 2024, analyzing vision-based, sensor-based, and hybrid approaches. The study evaluates feature extraction techniques, classifiers, datasets, and accuracy trends, while highlighting key challenges such as signer-independence, dataset limitations, and real-world implementation issues.

Xinyue Lou [3] discussed vision-based hand gesture recognition technology, highlighting the steps of pre-processing, feature extraction, and gesture classification for static and dynamic gestures. The study emphasizes applications in healthcare, smart homes, gaming, automotive systems, and education, demonstrating the advantages of vision-based systems over wearable devices.

Reza Jalayer et al. [4] presented a comprehensive review of deep learning techniques for vision-based hand detection, segmentation, and gesture recognition in human-robot interaction. The study analyses CNNs, RNNs, Transformers, and multimodal approaches, highlighting challenges such as dataset limitations, real-time performance, and robustness in dynamic environments.

The reviewed studies show that vision-based hand gesture recognition systems have significantly improved in accuracy and real-time performance using OpenCV, MediaPipe, and deep learning models. Both rule-based and learning-based approaches are effective for practical human-computer interaction applications. However, challenges such as lighting variations, background complexity, and dataset limitations still affect system robustness. These insights support the development of a reliable and efficient touchless interaction system for real-world use.

III. PROPOSED SYSTEM

The proposed Gesture-Based Virtual Control System enhances human-computer interaction by enabling users to control system operations through hand gestures without relying on traditional input devices such as a mouse or keyboard. Conventional interaction methods require physical contact and dedicated hardware, which may not be suitable in environments requiring touchless operation or for users with physical limitations. These limitations reduce flexibility and restrict intuitive interaction with computing systems.

The overall architecture of the proposed system is shown in **Fig. 1**.



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

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

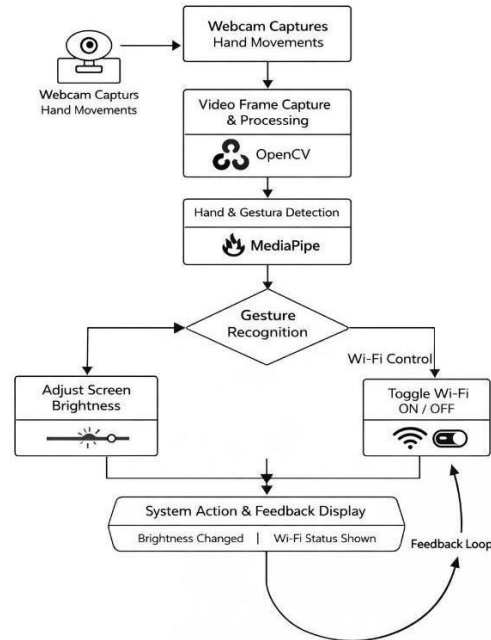


Fig. 1. Architecture of the system

As shown in **Fig. 1**, the system begins with a webcam capturing real-time hand movements. The captured video frames are processed using OpenCV for frame extraction and preprocessing. The processed frames are then passed to the hand and gesture detection module, where MediaPipe detects the hand region and extracts key landmarks required for gesture interpretation.

The gesture recognition module analyses the spatial relationships between the extracted landmarks to identify specific gestures such as finger movements or predefined hand signs. Once a gesture is recognized, it is mapped to a corresponding system command. The system supports multiple gesture-based controls, including cursor movement, zoom-in and zoom-out operations, screen brightness adjustment, and Wi-Fi on/off control.

After executing the command, the system provides visual confirmation through a feedback display module, ensuring that the user is aware of the performed action. A feedback loop is maintained to support continuous and accurate interaction. The modular architecture of the system improves scalability, maintainability, and flexibility, allowing additional gestures or functionalities to be integrated easily without affecting existing components.

By integrating real-time computer vision with gesture-based control mechanisms, the proposed system provides an intuitive, efficient, and touch-free interaction platform suitable for smart environments, educational systems, public kiosks, and assistive technology applications.

To provide a clearer understanding of the internal workflow, the overall architecture shown in Fig. 1 is divided into several functional modules. Each module performs a specific task in the gesture recognition and control process. The modular breakdown ensures systematic processing of video input, accurate gesture identification, reliable command execution, and effective feedback delivery. The following subsections describe each module in detail.



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

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

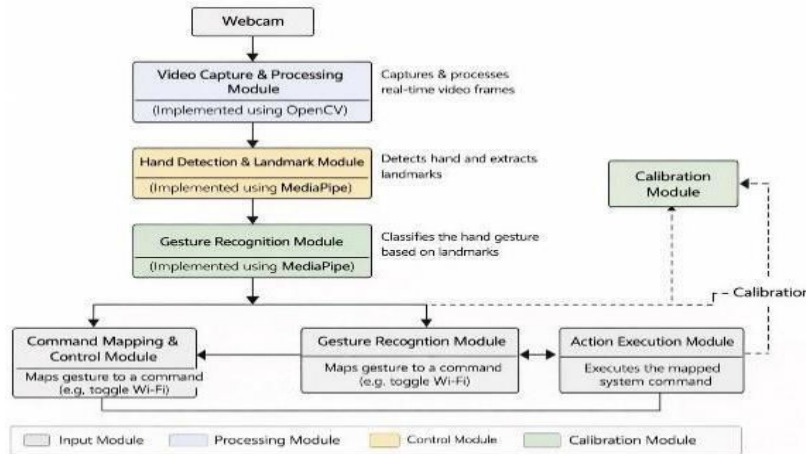


Fig. 2. Functional block diagram

As shown in **Fig. 2**, the system follows a sequential processing pipeline beginning with video acquisition through a webcam, followed by frame preprocessing, hand detection, landmark extraction, gesture recognition, command mapping, and action execution.

3.1 Video Capture & Processing Module

The Video Capture & Processing Module captures real-time video frames using a webcam. It performs essential image processing operations such as resizing, filtering, and frame normalization. The module converts raw frames into a suitable format for further analysis. OpenCV is used to handle video acquisition and preprocessing efficiently.

3.2 Hand Detection Module

The Hand Detection Module identifies the presence of a hand within the captured video frame. It detects the hand region and separates it from the background to reduce noise. The module ensures that only relevant hand data is forwarded for further processing. MediaPipe is used to enable accurate and real-time hand tracking.

3.3 Landmark Extraction Module

The Landmark Extraction Module extracts 21 key landmark points from the detected hand. These landmarks represent finger joints, fingertips, and palm positions, forming a structured skeletal representation of the hand. The spatial coordinates (x, y, z) of these points are used to analyze finger movements and positional relationships.

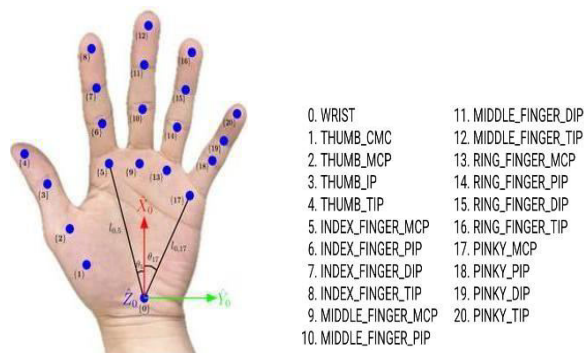


Fig. 3. Hand gesture landmarks

As shown in **Fig. 3**, the system identifies and labels landmarks from 0 to 20, starting from the wrist and extending to the fingertips of each finger. This structured landmark mapping enables accurate gesture interpretation by allowing precise



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

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

tracking of finger bending, extension, and relative positioning, thereby improving recognition accuracy and system responsiveness.

3.4 Gesture Recognition Module

The Gesture Recognition Module analyses the extracted landmark coordinates to classify the performed hand gesture. It evaluates finger positions and relative distances between landmarks. Based on predefined logic or rules, the gesture is identified accurately. The recognized gesture is then forwarded for command mapping.

3.5 Command Mapping & Control Module

The Command Mapping & Control Module converts recognized gestures into specific system commands. Each gesture is mapped to a predefined operation such as brightness control or Wi-Fi toggling. This module acts as the interface between gesture recognition and system control. It ensures correct interpretation of user intentions.

3.6 Action Execution Module

The Action Execution Module performs the actual system operation corresponding to the mapped command. It interacts with system-level functions to execute actions such as adjusting brightness or enabling Wi-Fi. The module ensures timely and accurate command execution. It completes the control phase of the system.

3.7 Feedback & Display Module

The Feedback & Display Module provides real-time confirmation of executed actions. It displays visual feedback such as brightness levels or Wi-Fi status on the screen. This helps users verify that their gesture was correctly recognized. The feedback loop improves system usability and interaction reliability.

3.8 Calibration Module

The Calibration Module adjusts system parameters to improve detection accuracy under different lighting conditions. It sets HSV ranges, threshold values, and filtering parameters. This module ensures consistent performance across various environments. Calibration enhances overall system robustness and gesture recognition precision.

The integration of these functional modules ensures efficient real-time gesture recognition and reliable system control. The modular architecture enhances scalability, accuracy, and adaptability of the proposed vision-based interaction system.

IV. METHODOLOGY

4.1 Hand Gesture Recognition Algorithm

The proposed Vision-Based Hand Gesture Interaction System utilizes real-time computer vision techniques for accurate hand tracking and gesture recognition. The overall working concept of the system is illustrated in **Fig. 4**.

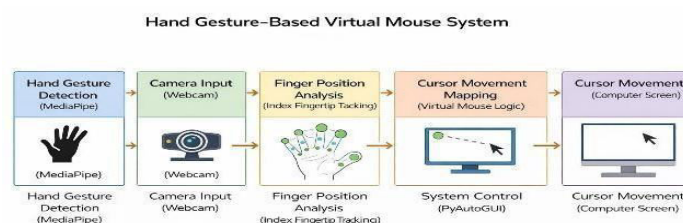


Fig. 4. Algorithm working concept

As shown in **Fig. 4**, the system captures live video frames through a webcam and processes them using MediaPipe for hand detection and tracking. The detected landmarks are analyzed to determine finger positions and motion patterns, enabling cursor movement and system-level control for touchless interaction.

Each frame undergoes preprocessing using OpenCV, including resizing, noise reduction, colour space conversion, and normalization to enhance detection accuracy. MediaPipe's Hand Tracking Model detects twenty-one hand



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

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

landmarks corresponding to finger joints, fingertips, and palm coordinates, providing precise spatial information for gesture analysis.

The extracted landmarks are processed in the feature extraction stage, where geometric parameters such as Euclidean distance between fingertips, joint angles, relative fingertip positions, and motion vectors are computed. These features are evaluated using a rule-based classification algorithm that compares them with predefined thresholds to recognize specific gestures.

Once classified, the gesture is translated into a corresponding system command through the command mapping module. The action execution stage performs the required operation and provides real-time visual feedback. The complete structured processing pipeline is depicted in Fig. 5.

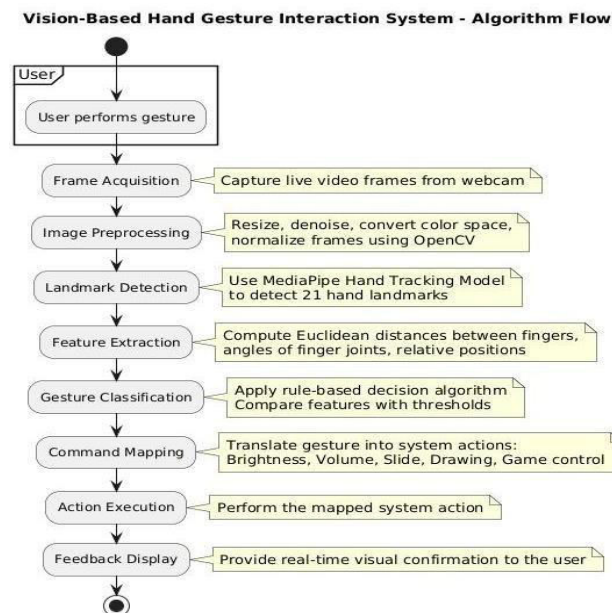


Fig. 5. Algorithm flowchart

As shown in Fig. 5, the overall algorithm follows a structured processing pipeline:

Frame Acquisition → Image Preprocessing → Landmark Detection → Feature Extraction → Gesture Classification → Command Mapping → Action Execution → Feedback Display.

This modular flow ensures continuous looping until system termination, maintaining real-time performance and accurate touchless interaction without requiring any peripheral devices. The integration of OpenCV and MediaPipe ensures efficient processing, low latency, and reliable gesture-based control suitable for smart environments and interactive applications.

V. RESULTS AND DISCUSSION

The proposed Vision-Based Hand Gesture Interaction System was tested in real-time using a webcam under normal indoor lighting conditions. The system successfully detected hand landmarks and accurately classified gestures to perform various system-level operations such as Wi-Fi control, brightness adjustment, and virtual drawing. The experimental results demonstrate effective gesture detection with smooth real-time response and minimal delay. The system achieved approximately 92% gesture recognition accuracy with an average response delay below 100 ms under standard indoor lighting conditions.



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

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

5.1 Wi-Fi Control

The output images demonstrate successful Wi-Fi on and Wi-Fi off operations using predefined hand gestures. As shown in **Fig. 6**, when the Wi-Fi activation gesture was performed, the system accurately recognized the finger configuration and executed the mapped command to turn the Wi-Fi on. Similarly, as shown in **Fig. 7**, the Wi-Fi off gesture was correctly detected, and the system deactivated the Wi-Fi accordingly. The confirmation message displayed on the screen verifies proper command execution. The experimental results indicate that the gesture recognition module reliably classified finger positions with near real-time response and stable detection performance.

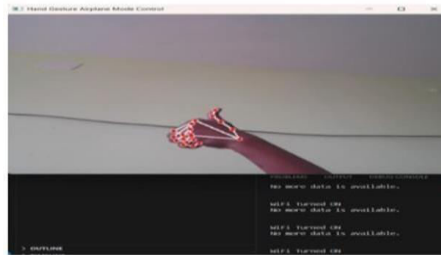


Fig. 6. Wi-Fi turn on

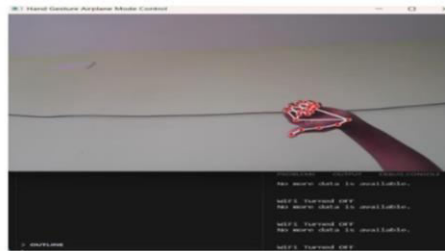


Fig. 7. Wi-Fi turn off

5.2 Brightness Control

The brightness control functionality was evaluated using increasing and decreasing gestures. As shown in **Fig. 8**, the brightness level was successfully increased up to 100% when the corresponding gesture was detected. Conversely, as illustrated in **Fig. 9**, the brightness level decreased (e.g., to 23%) when the decreasing gesture was performed. The system utilized landmark-based distance measurements between fingers to achieve smooth and continuous brightness scaling. These results confirm that the proposed system supports dynamic and precise brightness adjustment rather than fixed- step control.



Fig. 8. Brightness increasing



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

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



Fig. 9. Brightness decreasing

5.3 Hand Gesture Paint

The hand gesture paint output demonstrates the system's capability to track fingertip movement for drawing applications. As shown in **Fig. 10**, the index finger position was continuously monitored using landmark coordinates extracted from the detected hand. The system generated colored drawing paths in real time as the finger moved across the frame. The minimal delay observed in rendering validates the efficiency of the landmark extraction and motion tracking modules. The results confirm accurate spatial tracking and stable performance during continuous gesture execution.

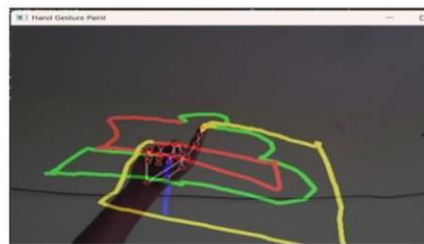


Fig. 10. Hand gesture paint

5.4 Virtual Drawing

The virtual drawing functionality further highlights the robustness of the proposed framework. As shown in **Fig. 11**, the system captured dynamic finger movements and rendered freehand sketches smoothly on the screen. The drawing strokes followed the fingertip trajectory precisely, demonstrating accurate real-time tracking. Immediate visual feedback enhanced user interaction and usability. These results validate the effectiveness of the gesture recognition and action execution modules under continuous motion conditions.



Fig. 11. Virtual drawing

The overall results demonstrate that the proposed system achieves accurate gesture recognition and reliable command execution in real-time. The integration of efficient landmark detection and image processing techniques ensures stable performance with minimal latency. All implemented functionalities, including Wi-Fi control, brightness adjustment, and virtual drawing, operated successfully under normal lighting conditions. These findings confirm that the system is effective, responsive, and suitable for practical touchless human-computer interaction applications.



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

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

5.5 Comparison with Existing System

Existing systems mainly concentrate on gesture recognition accuracy and often require additional hardware such as data gloves, sensors, or specialized devices. Some approaches depend heavily on deep learning models, which require large training datasets and high computational power. Many of these systems are limited to single applications and may not perform efficiently under varying lighting or background conditions. In contrast, the proposed system uses only a standard webcam with OpenCV and MediaPipe for real-time hand tracking and gesture recognition. It not only detects gestures but also integrates practical system-level controls such as brightness adjustment, volume control, presentation navigation, gaming, and drawing. The system is cost-effective, easy to implement, requires no peripheral devices, and provides smooth, touchless interaction suitable for real-time applications.

VI. CONCLUSION

In conclusion, the proposed Vision-Based Hand Gesture Interaction System improves human-computer interaction by enabling touch-free control using real-time hand gesture recognition. The system uses computer vision techniques to detect and track hand movements and convert them into commands such as cursor movement, clicking, brightness control, and volume adjustment, reducing dependency on traditional input devices. It provides a simple, cost-effective, and scalable solution with smooth real-time performance. In the future, the system can be enhanced with advanced deep learning models for better accuracy, support for dynamic gestures and full sign language recognition, and integration with voice control to make it more efficient, accessible, and suitable for smart and assistive applications.

VII. FUTURE SCOPE

In the future, the system can be enhanced by supporting dynamic gesture recognition and a wider range of gesture commands to improve functionality and flexibility. Additionally, integrating advanced deep learning models can further increase recognition accuracy under different lighting and background conditions. The system can also be expanded to support sign language translation, enabling communication assistance for hearing-impaired individuals. Furthermore, combining voice control with gesture recognition can create a multimodal interaction system for improved usability. By incorporating cross-platform compatibility and mobile device integration, the solution can be deployed in smart homes, virtual reality environments, and assistive technologies, thereby increasing its practical applications and overall impact.

REFERENCES

- [1] Mukhopadhyay A., Chakrabarty A., Das A. A., and Sarkar A., Hand Gesture Based Recognition System, Proceedings of the 7th International Conference on Electronics, Materials Engineering and Nano-Technology (IEMENTech), 18–20 December 2023, IEEE, 2024.
- [2] Hashi A. O., Mohd Hashim S. Z., and Asamah A. B., A Systematic Review of Hand Gesture Recognition: An Update from 2018 to 2024, IEEE Access, Vol. 12, 2024, pp. 1–28.
- [3] Lou X., Vision-Based Hand Gesture Recognition Technology, Proceedings of the 3rd International Conference on Mechatronics and Smart Systems, 2025, pp. 54–59.
- [4] Jalayer R., Jalayer M., Orsenigo C., and Tomizuka M., A Review on Deep Learning for Vision-Based Hand Detection, Hand Segmentation and Hand Gesture Recognition in Human–Robot Interaction, Robotics and Computer-Integrated Manufacturing, Vol. 97, February 2026.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

 9940 572 462  6381 907 438  ijircce@gmail.com



www.ijircce.com

Scan to save the contact details