

Real-Time Air Gesture Recognition Using Computer Vision for Touchless Human-Computer Interaction

Meenakshi Yadav¹, Sukhdanishi Varma², Simran Arya³, Sheelu⁴, Sonam⁵
Department of Computer Science & Engineering
Moradabad Institute of Technology

¹meenakshiyadav2309@gmail.com, ²sukhdanishi@gmail.com, ³sarya0558@gmail.com,
⁴shelukumari91@gmail.com, ⁵sonamsingh9045267458@gmail.com

Abstract

The Air Gesture system presents a touchless human-computer interaction (HCI) framework that enables users to control computer operations through dynamic hand gestures performed in mid-air. The system removes dependency on conventional input devices such as a mouse, keyboard, or touchscreen by utilizing real-time hand gesture recognition through a monocular webcam. This approach provides a low-cost and hardware-efficient alternative for natural interaction.

The proposed system employs computer vision and machine learning techniques for hand detection, tracking, and gesture classification. Key stages include video frame acquisition, hand region segmentation, feature extraction based on hand landmarks and motion trajectories, and gesture recognition using trained machine learning and deep learning models. Temporal information is captured to accurately recognize dynamic gestures, ensuring robustness against variations in hand orientation, scale, and illumination. The recognized gestures are mapped to system-level commands such as cursor movement, clicking, scrolling, virtual typing, drawing, and multimedia control.

Real-time inference is achieved with low latency, enabling smooth and responsive interaction. Experimental evaluations demonstrate reliable gesture recognition performance under diverse environmental conditions. The system is hygienic, scalable, and user-friendly, making it well suited for applications in smart classrooms, interactive presentations, public interfaces, and assistive technologies. The proposed air gesture-based interaction model highlights the potential of vision-driven intelligent interfaces in advancing next-generation human-computer interaction systems.

Keywords: Air Gesture Recognition, Human–Computer Interaction, Computer Vision, Hand Landmark Detection, Touchless Interface, Real-Time Systems

1. Introduction

Human–computer interaction (HCI) has undergone significant evolution, transitioning from traditional physical input devices such as keyboards and mice to more natural and intuitive interaction paradigms [1]. With increasing demand for touchless and hygienic interfaces, air gesture–based interaction has gained attention as an effective solution for contactless human–computer communication [2]. Air gesture technology allows users to interact with computing systems through hand movements performed in mid-air, without requiring physical contact or wearable devices [3].

Recent advancements in computer vision and machine learning have enabled real-time hand detection and tracking using conventional cameras [4]. Vision-based air gesture systems typically employ hand landmark detection, finger position analysis, and motion tracking to recognize predefined gestures and map them to computer operations [5]. Such systems provide flexible and user-friendly alternatives to conventional input methods and have shown promising applications in smart classrooms, presentations, virtual reality environments, and assistive technologies [6].

2. Literature Survey

Early research in gesture recognition primarily relied on specialized hardware devices such as data gloves, magnetic sensors, and depth cameras to capture hand movements and gestures. While these systems provided accurate motion tracking, they significantly increased system cost, computational complexity, and user discomfort, thereby limiting their widespread adoption in real-world applications [7]. Moreover, the dependency on dedicated hardware reduced portability and restricted usage to controlled environments.

To overcome these limitations, vision-based gesture recognition approaches using standard RGB cameras have gained substantial attention. These methods leverage advances in computer vision and machine learning to detect and interpret hand gestures without requiring additional sensors. Vision-based systems are cost-effective, portable, and easier to deploy, making them suitable for a wide range of applications including human–computer interaction, virtual environments, and assistive technologies [8].

Recent studies emphasize the use of landmark-based hand tracking techniques, which represent the hand using key anatomical points such as fingertips, joints, and palm centers. Among these approaches, MediaPipe Hand Tracking has emerged as a robust and efficient framework capable of detecting 21 three-dimensional hand landmarks with high accuracy and low latency in real time. MediaPipe employs a lightweight deep learning pipeline that combines palm detection and hand landmark regression, enabling stable tracking even under varying lighting conditions and complex backgrounds [9].

Several researchers have demonstrated that landmark-based hand tracking significantly improves gesture recognition accuracy and responsiveness compared to traditional feature-based methods. These techniques allow precise modeling of finger movements and hand orientation, making them particularly effective for real-time air gesture recognition and touchless interaction systems [10], [11]. Consequently, landmark-driven vision-based approaches have become a preferred solution for developing scalable, user-friendly, and efficient gesture recognition frameworks.

3. Methodology

3.1 Objective

The primary objective of this project is to design and implement a **real-time air gesture recognition system** that enables users to control computer functions without physical contact. The system aims to provide an intuitive and touchless human–computer interaction mechanism by capturing live video input through a standard webcam, detecting and tracking hand landmarks, recognizing predefined gestures, and executing corresponding system-level commands. Emphasis is placed on achieving high accuracy, low latency, and ease of use while maintaining compatibility with commonly available hardware.

3.2 Gesture Identification Process

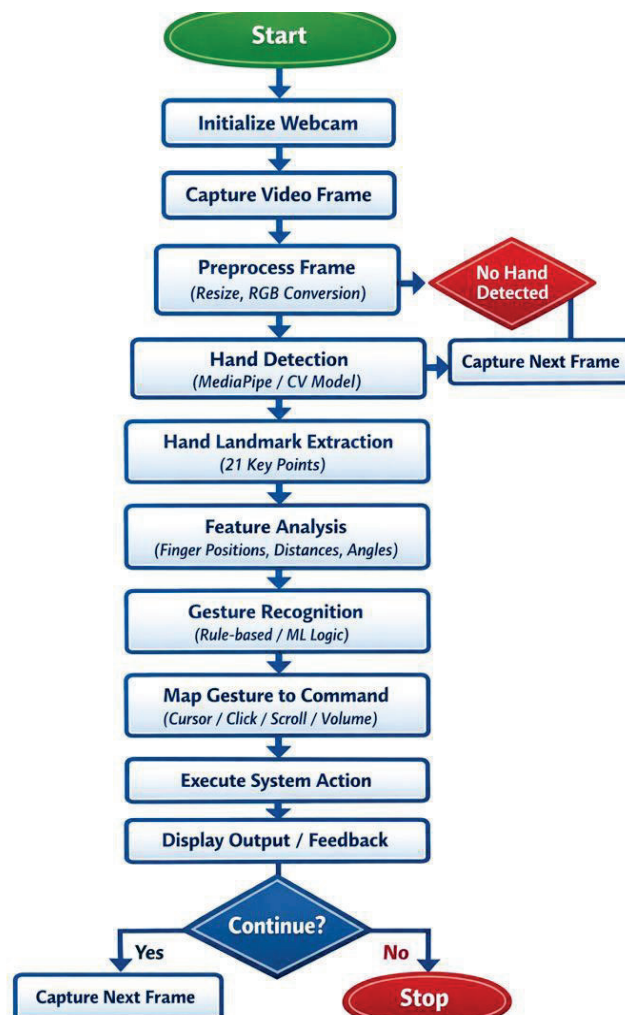
The gesture identification process is based on analyzing the **spatial relationships and relative positions of hand landmarks** detected in each video frame. Using a landmark-based hand tracking model, key points such as fingertips, joints, and palm center are extracted to represent hand posture and motion. Simple geometric computations, including Euclidean distance, relative angles, and positional thresholds, are applied to determine the state of individual fingers (extended or folded).

Specific gestures are defined using these spatial features. For instance, cursor movement is controlled by continuously tracking the position of the index fingertip, while clicking actions

are triggered by detecting pinching gestures formed by reducing the distance between the thumb and index finger below a predefined threshold. This rule-based approach enables efficient and interpretable gesture recognition without requiring complex model training.

3.3 System Responsiveness

To ensure real-time performance, the system processes each video frame independently and performs gesture recognition with minimal computational overhead. The use of lightweight computer vision operations and optimized hand landmark detection enables smooth interaction with low latency. This design allows the system to function efficiently even on devices without dedicated graphics processing units (GPUs). As a result, the proposed system delivers responsive and seamless user interaction, making it suitable for real-time applications such as presentations, smart classrooms, and assistive interfaces.



3.4 Features of the System

The proposed Air Gesture system provides multiple touchless interaction features that enable users to control various computer functions efficiently and intuitively using hand gestures. The key features of the system are described below:

1. irtual Keyboard

The virtual keyboard feature allows users to type text without physical contact with a keyboard. Hand gestures are used to select and confirm characters displayed on a virtual on-screen keyboard. This feature is particularly useful in public or shared environments where hygiene is a concern, as well as for users with physical disabilities.

2. irtual Mouse

The virtual mouse enables complete cursor control through hand movements. Cursor navigation is achieved by tracking the index finger position, while clicking, dragging, and scrolling actions are performed using predefined gestures such as pinching or finger combinations. This feature effectively replaces a conventional mouse, offering a natural and flexible mode of interaction.

3. irtual Painter

The virtual painter feature allows users to draw or write on a digital canvas using finger gestures. By tracking the movement of the index finger, the system enables freehand drawing, sketching, or annotation. This feature is useful for educational purposes, presentations, and creative applications.

4. Virtual Brightness Control

The virtual brightness control feature enables users to adjust the screen brightness using hand gestures. By varying the distance or orientation between specific fingers, the system increases or decreases brightness levels in real time. This touchless control enhances user convenience and reduces dependency on physical controls.

5. irtual Volume Control

The virtual volume control feature allows users to adjust system audio levels through simple hand gestures. Gesture-based control, such as increasing or decreasing the distance between fingers, is mapped to volume adjustment, enabling seamless multimedia interaction without physical input devices.

1.1 Virtual Keyboard

The virtual mouse module allows users to control the keyboard keys using hand movement.

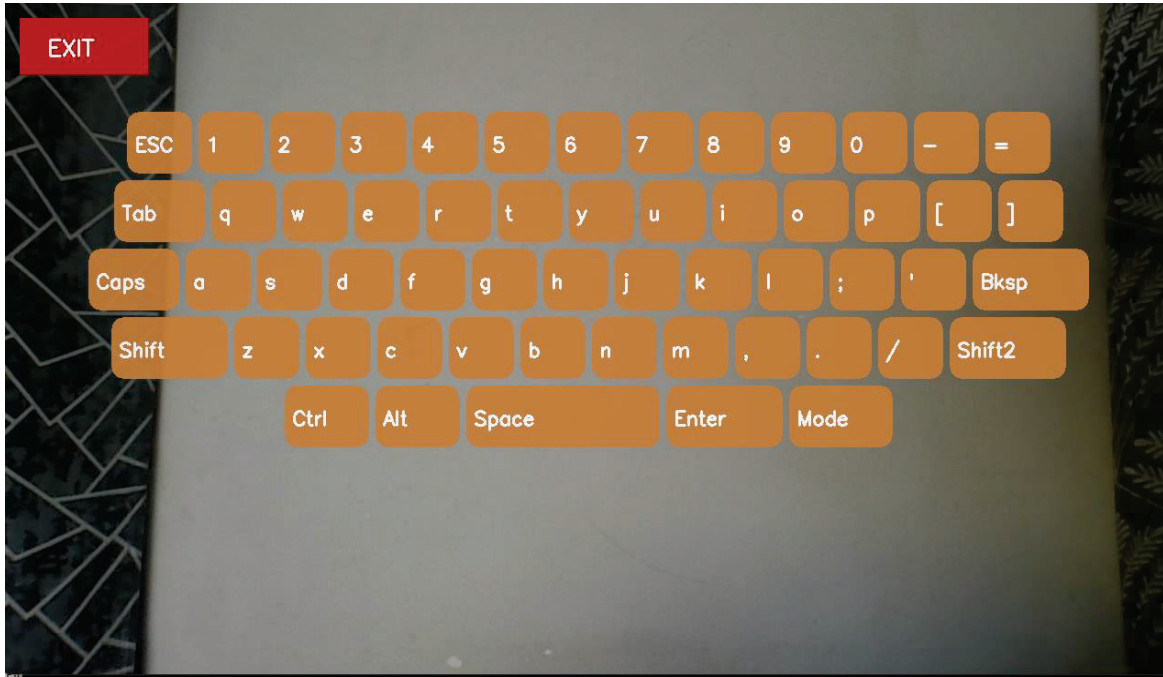


Figure 1: Virtual Keyboard using Hand Gestures

1.2 Virtual Mouse

The virtual mouse module allows users to control the cursor using hand movement.

- Cursor Movement: Index finger tip
- Left Click: Thumb and index finger
- Right Click: Thumb and middle finger
- Scroll: Vertical finger motion

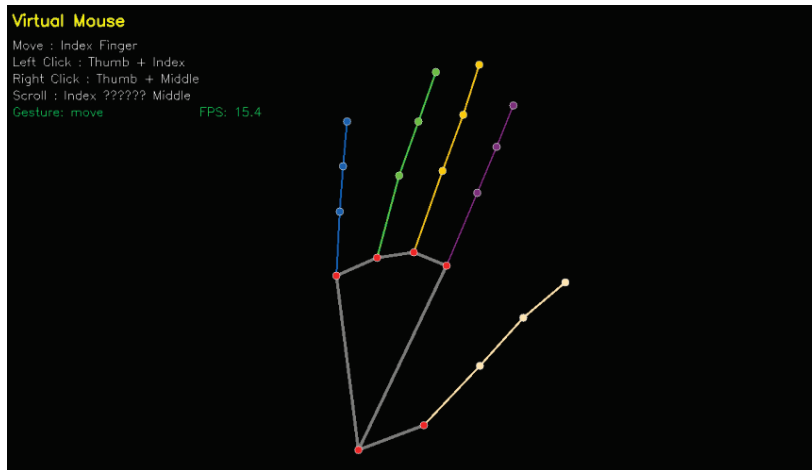


Figure 2: Virtual Mouse using Hand Gestures

1.3 Virtual Painter

The virtual painter feature allows users to draw in the air by tracking fingertip movement.



Figure 3: Painter Using Hand Gestu

1.4 Virtual Brightness Control

Screen brightness is adjusted by measuring the distance between the thumb and index finger.

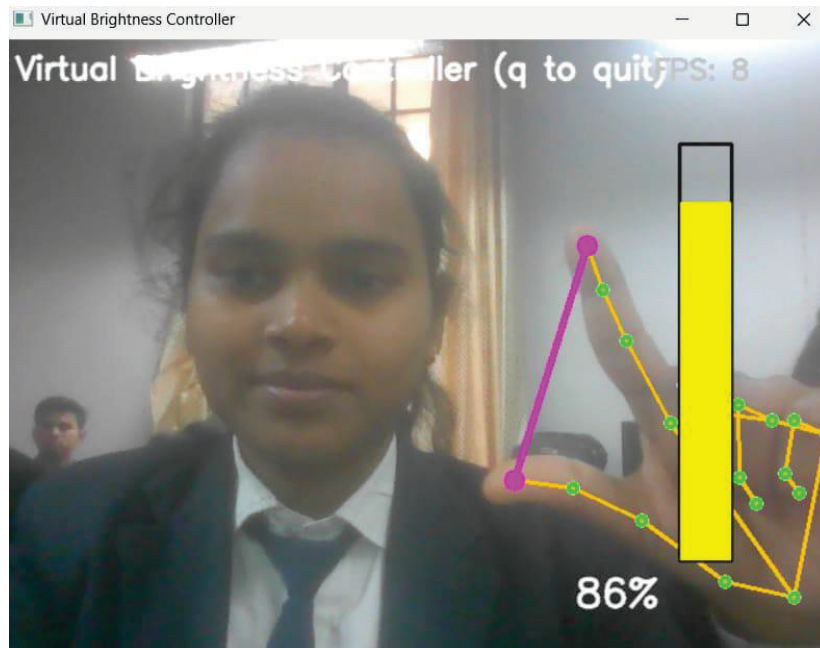


Figure 4: Gesture-Based Brightness Control

1.5 Virtual Volume Control

System audio volume is controlled dynamically using finger distance gestures.

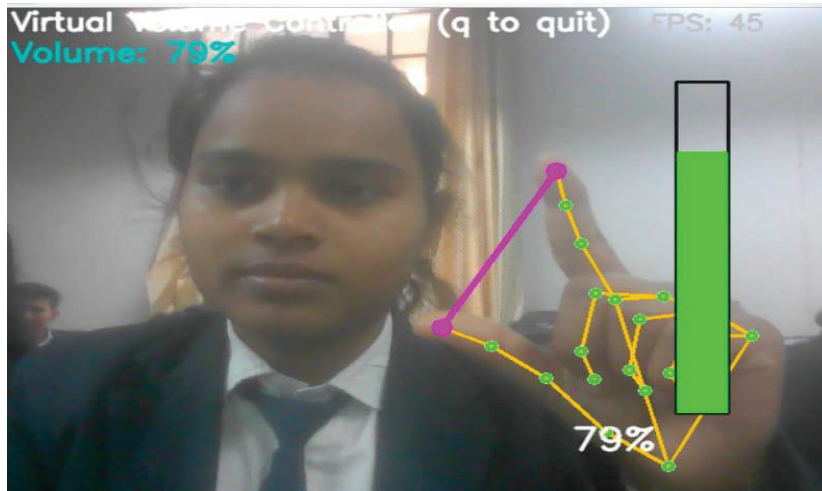


Figure 5: Gesture-Based Volume Control

4. Applications

The proposed air gesture recognition system can be applied in a wide range of real-world scenarios requiring touchless interaction. In **smart classrooms**, the system enables gesture-based control of presentations, digital boards, and learning content, enhancing interactivity and hygiene. It is also suitable for **touch-free public systems** such as kiosks, information terminals, and ticketing machines, where minimizing physical contact is essential.

The system supports **assistive technologies** by providing an alternative interaction mechanism for users with physical impairments, improving accessibility and ease of use. Additionally, it can be utilized in **creative and design applications**, including digital drawing and sketching through virtual painter functionality. The system is further applicable to **interactive presentations**, allowing seamless control of multimedia content using intuitive hand gestures. These applications demonstrate the effectiveness of air gesture-based interfaces for modern touchless human-computer interaction environments.

5. Advantages

- The proposed air gesture recognition system offers several advantages over traditional input methods.
- It enables **touchless and hygienic interaction**, reducing the need for physical contact with devices.

- The system operates using a standard webcam, requiring **no additional hardware**, which lowers deployment complexity.
- It is a **low-cost and portable solution**, making it accessible for widespread use.
- The system is **easy to use and user-friendly**, allowing intuitive interaction without extensive user training.

6. Limitations

Although the proposed air gesture recognition system demonstrates effective real-time performance, certain limitations remain. The accuracy of hand detection and gesture recognition may degrade under **poor or uneven lighting conditions** and **cluttered backgrounds**. Additionally, **rapid or abrupt hand movements** can occasionally result in gesture misclassification due to motion blur or temporary loss of landmark tracking. These factors may affect system robustness in uncontrolled environments.

7. Future Scope

Future work will focus on enhancing the robustness and scalability of the air gesture recognition system. Advanced deep learning models, such as convolutional and recurrent neural networks, can be integrated to improve gesture classification accuracy under varying lighting conditions and complex backgrounds. Temporal modeling of gestures may further reduce misclassification caused by rapid hand movements.

The system can also be optimized for **edge and embedded platforms** to support deployment on low-power devices. Additionally, extending the framework to support **multimodal interaction**, including voice commands and depth sensing, can improve usability and reliability. Expanding the gesture vocabulary and incorporating user-adaptive learning are further directions to enhance personalization and application scope in real-world environments.

8. Conclusion

This paper presented an air gesture recognition system that demonstrates the effective use of computer vision techniques as an alternative to traditional input devices. By enabling touchless interaction through a standard webcam, the system provides an efficient, intuitive, and cost-effective approach to human-computer interaction. The proposed framework supports real-time gesture recognition and control of multiple system functions without the need for additional

hardware. Overall, the results highlight the potential of air gesture-based interfaces to enhance accessibility, usability, and hygiene in modern interactive computing environments.

References

- [1] S. A. Brewster, "The evolution of human-computer interaction," *IEEE Computer Graphics and Applications*, vol. 24, no. 1, pp. 44–45, 2004.
- [2] Y. Wu and T. S. Huang, "Vision-based gesture recognition: A review," *International Gesture Workshop*, Springer, pp. 103–115, 1999.
- [3] J. O. Wobbrock, H. H. Aung, B. Rothrock, and B. A. Myers, "Maximizing the guessability of symbolic input," *Proc. SIGCHI Conference on Human Factors in Computing Systems*, pp. 1869–1878, 2005.
- [4] Z. Cao, G. Hidalgo, T. Simon, S.-E. Wei, and Y. Sheikh, "OpenPose: Realtime multi-person 2D pose estimation using part affinity fields," *IEEE TPAMI*, vol. 43, no. 1, pp. 172–186, 2021.
- [5] F. Molchanov, S. Gupta, K. Kim, and J. Kautz, "Hand gesture recognition with 3D convolutional neural networks," *CVPR Workshops*, pp. 1–7, 2015.
- [6] R. Rautaray and A. Agrawal, "Vision based hand gesture recognition for human computer interaction: A survey," *Artificial Intelligence Review*, vol. 43, no. 1, pp. 1–54, 2015.
- [7] Dipietro, L., Sabatini, A. M., & Dario, P. (2008). A survey of glove-based systems and their applications. *IEEE Transactions on Systems, Man, and Cybernetics*, 38(4), 461–482.
- [8] Mitra, S., & Acharya, T. (2007). Gesture recognition: A survey. *IEEE Transactions on Systems, Man, and Cybernetics*, 37(3), 311–324.
- [9] Zhang, F., Bazarevsky, V., Vakunov, A., et al. (2020). MediaPipe Hands: On-device real-time hand tracking. arXiv preprint arXiv:2006.10214.
- [10] Wachs, J. P., Kölsch, M., Stern, H., & Edan, Y. (2011). Vision-based hand-gesture applications. *Communications of the ACM*, 54(2), 60–71.
- [11] Molchanov, P., Gupta, S., Kim, K., & Kautz, J. (2016). Hand gesture recognition with 3D convolutional neural networks. *IEEE CVPR*.