ISSN2277-2685

IJESR/April. 2026/ Vol-16/Issue-21/776-781

Full Length Article

Real-Time Hand Gesture Controlled UAV Using Computer Vision Techniques.

Mrs. Kiran Pakmode[1], Thakur Tarun Singh[2], Sanjay Kumar Anugula[3], Merugu Swapnil Reddy[4], Kundarapu Dikshitha[5]

[1]Associate Professor; Department Of Electronics & Computer Engineering J.B. Institute Of Engineering & Technology

[2,3,4,5]B.Tech Student's; Department Of Electronics & Computer Engineering J.B. Institute Of Engineering & Technology

Accepted 27-04-2026

Abstract

Unmanned Aerial Vehicles (UAVs) are widely deployed in applications such as surveillance, inspection, and environmental monitoring. Conventional UAV control methods depend on handheld controllers or mobile devices, which may be impractical in situations requiring hands-free or rapid response interaction. This study proposes a real-time, vision-driven hand gesture control system that enables intuitive and touchless UAV operation. The system captures live video using a camera and processes it through a computer vision pipeline to detect and track hand movements. Using landmark-based hand modelling , key features are extracted and analyzed to identify predefined gestures based on spatial relationships among hand joints. The implementation leverages Python along with OpenCV and MediaPipe Hands for efficient gesture recognition. Recognized gestures are mapped to specific flight commands and transmitted to a DJI Tello drone via its Software Development Kit (SDK), allowing real-time execution of actions such as takeoff, landing, directional navigation, and altitude adjustment. To enhance operational safety, the system integrates mechanisms such as gesture validation delays, emergency stop controls, and automatic landing triggered by low battery levels or communication disruptions.Experimental evaluation demonstrates that the system achieves reliable gesture recognition with minimal latency, resulting in smooth and stable drone maneuvering. The proposed framework eliminates the need for traditional input devices and offers a practical solution for natural human–drone interaction. This approach is particularly beneficial for applications in search and rescue, surveillance, and advanced human–machine interfaces, demonstrating its potential as a scalable and effective UAV control paradigm.

Keywords

Hand Gesture Recognition, UAV Control, Computer Vision, Human–Drone Interaction, MediaPipe Hands, OpenCV, Real-Time Systems, DJI Tello, Contactless Control, Intelligent Interfaces

INTRODUCTION

Unmanned Aerial Vehicles (UAVs), widely referred to as drones, have experienced rapid growth in recent years due to their versatility across multiple domains such as surveillance, agriculture, disaster response, and aerial imaging. Among the various platforms available, the DJI Tello drone has emerged as a popular choice in academic and experimental settings because of its compact structure, affordability, and programmability using languages like Python. Conventional drone control methods typically rely on handheld transmitters or mobile-based applications, which require continuous physical interaction and may not provide an intuitive experience for all users, especially in time-critical or hands-free environments. With advancements in computer vision and artificial intelligence, alternative interaction mechanisms have been explored to enhance usability and efficiency. One such promising approach is hand gesture recognition, which allows users to communicate commands through natural hand movements without requiring physical interfaces.This work, titled "Real-Time Hand Gesture Controlled UAV Using Computer Vision Techniques," focuses on designing and implementing a vision-based control system that enables intuitive drone navigation using real-time hand gestures. The system utilizes camera input to capture live video, which is processed using computer vision frameworks such as OpenCV for image processing and MediaPipe Hands for precise hand landmark detection. By analyzing spatial relationships between detected hand landmarks, predefined gestures are identified and mapped to corresponding drone commands. These commands enable the drone to perform movements such as forward, backward, lateral navigation, and altitude adjustment. This approach enhances human–computer interaction by offering a more natural and contactless method of controlling UAV systems.The development of the Tello drone is closely associated with collaboration between Ryze Tech, DJI, and Intel, and it was officially introduced during the Consumer Electronics Show 2018. Designed as a lightweight quadcopter, it features a 5-megapixel camera capable of capturing images and recording 720p video, along with flight stabilization technology that ensures smooth operation, particularly in indoor environments. The integration of an Intel Movidius processor enables basic onboard vision processing, making it suitable for entry-level computer vision applications. Over time, the platform gained widespread acceptance in STEM education, leading to the release of an enhanced version, Tello EDU, which supports programming environments such as Python, Scratch, and Swift. This programmability makes it highly suitable for research-oriented applications, including gesture-based control systems.The methodology of the proposed system begins with establishing a Wi-Fi connection between the drone and a computing device, allowing real-time video streaming. The captured frames are processed to extract relevant visual features, after which hand landmarks are detected using MediaPipe. Based on geometric analysis of these landmarks, gestures are classified into predefined categories corresponding to drone actions. The interpreted commands are then transmitted to the drone through a Python-based interface, enabling real-time execution of flight operations such as takeoff, landing, and directional movement. The system continuously processes incoming frames to maintain responsiveness and ensure smooth control.The successful implementation of this system requires coordinated efforts across multiple roles, including software development, computer vision engineering, system integration, testing, and safety monitoring. Each role contributes to ensuring that the gesture recognition system operates accurately, the communication with the drone remains stable, and overall system performance meets reliability standards. Additionally, proper documentation and validation are essential to evaluate system behavior under different environmental conditions and to ensure safe operation during testing phases.This project contributes to the advancement of intuitive UAV control by introducing a gesture-based interface that replaces traditional controller-dependent mechanisms. By leveraging computer vision techniques, it demonstrates how real-time visual data can be effectively used to interact with physical devices. The system not only simplifies drone operation for beginners but also provides a foundation for developing advanced human–machine interaction systems. Furthermore, it highlights the potential of contactless control technologies in applications where conventional input methods may be impractical, thereby offering a scalable and efficient solution for next-generation UAV interfaces.

LITERATURE SURVEY

The literature survey examines prior research related to gesture-based control systems and Unmanned Aerial Vehicles (UAVs), highlighting the evolution of human–machine interaction techniques. With the advancement of computer vision and artificial intelligence, researchers have increasingly focused on developing intuitive and contactless control mechanisms that improve usability and reduce reliance on traditional input devices. Gesture recognition, in particular, has gained attention as a natural interaction method, enabling users to communicate commands through simple hand movements. Modern frameworks such as OpenCV and MediaPipe Hands have made it feasible to implement real-time gesture recognition systems with improved accuracy and efficiency.Existing drone control systems predominantly rely on handheld controllers or smartphone applications. For platforms such as the DJI Tello drone, users typically interact through mobile interfaces featuring virtual joysticks or dedicated remote controllers. While these methods are functional, they require continuous manual input and demand a certain level of skill, making them less accessible for beginners. Controlling parameters such as pitch, roll, yaw, and throttle simultaneously increases cognitive load and may lead to operational errors. Furthermore, conventional systems lack the ability to interpret natural human gestures in real time, limiting their adaptability and user-friendliness. Although some advanced research prototypes incorporate computer vision, these solutions are often complex, costly, or not suitable for educational purposes, thereby restricting widespread adoption.Several challenges are associated with existing approaches, including the dependence on manual controls, increased operational complexity, and the absence of intuitive interaction mechanisms. Users must continuously manage multiple control inputs, which can reduce efficiency and increase the likelihood of mistakes, potentially resulting in unstable flight or crashes. Additionally, traditional systems do not fully exploit modern vision-based technologies, limiting their capability to support advanced features such as gesture-driven control. These limitations emphasize the need for a more accessible and intelligent control framework.To address these issues, the proposed system introduces a real-time gesture-based control mechanism that replaces conventional input methods with natural hand interactions. The system captures live video input and processes it using OpenCV for preprocessing tasks such as frame enhancement and noise reduction.

Subsequently, MediaPipe is employed to detect and track hand landmarks with high precision. By analyzing spatial relationships between these landmarks, the system identifies predefined gestures and maps them to corresponding drone commands, including takeoff, landing, and directional movements. These commands are transmitted to the drone through a Python-based interface, enabling real-time and responsive control.The proposed approach offers several advantages, including improved usability, contactless interaction, and enhanced accuracy in gesture recognition. By eliminating the need for physical controllers, the system simplifies drone operation and makes it more accessible to users with minimal training. The integration of efficient vision-processing techniques ensures low latency and smooth control, while reducing the probability of human error. Overall, this system demonstrates a practical and innovative application of computer vision in UAV control, providing a scalable solution for future human–machine interaction systems.

REQUIREMENT ENGINEERING

Requirement engineering defines the essential software and hardware components necessary for developing and deploying the gesture-controlled UAV system. The software environment includes an operating system such as Windows, Linux, or macOS, which supports program execution and communication with the drone via Wi-Fi. The system is developed using Python due to its simplicity and extensive support for libraries related to computer vision and real-time processing. OpenCV is utilized for video capture and image processing, enabling frame manipulation and feature extraction, while MediaPipe facilitates accurate hand landmark detection and gesture tracking.To enable communication with the drone, the system employs the Djitellopy library, which provides an interface for sending commands such as takeoff, landing, and directional movement. Development is carried out using integrated environments like Visual Studio Code or PyCharm, which support efficient coding, debugging, and project management. Additional Python libraries, including NumPy and time modules, assist in performing numerical computations and managing real-time execution.The hardware requirements include a computer or laptop with sufficient processing capability to handle real-time video processing, preferably with at least 4 GB of RAM (8 GB recommended). A camera, either built-in or external, is required to capture hand gestures, while a stable Wi-Fi connection ensures communication between the system and the drone. The drone itself operates on a rechargeable battery, and maintaining adequate power levels is essential for stable flight and testing. Together, these hardware components ensure accurate gesture detection, efficient processing, and reliable drone operation.

IMPLEMENTATION

The implementation of the gesture-controlled UAV system is structured into multiple functional modules that work together to achieve real-time performance. The process begins with the video capture module, which continuously acquires frames from a camera source. These frames are then processed using OpenCV in the image processing module, where operations such as resizing, color conversion, and noise filtering are performed to enhance input quality.The hand detection and tracking module utilizes MediaPipe to identify key hand landmarks, including finger joints and tips, which are essential for gesture recognition. Based on these landmarks, the gesture recognition module interprets hand movements by comparing spatial configurations with predefined patterns. Once a gesture is identified, the command mapping module translates it into a corresponding drone instruction.Communication with the drone is handled through a dedicated module that establishes a Wi-Fi connection and transmits commands using a Python interface. The drone control module executes these commands, enabling actions such as takeoff, landing, and directional movement. Additionally, a feedback module provides visual output by displaying the processed video stream along with detected landmarks and recognized gestures, allowing users to monitor system performance in real time.The system also incorporates safety mechanisms such as gesture confirmation delays, command cooldown intervals, and emergency overrides to prevent unintended actions. These features ensure stable and secure operation, making the system suitable for practical deployment.

TESTING METHODOLOGIES

Testing plays a critical role in ensuring the reliability, accuracy, and performance of the gesture-controlled UAV system. Multiple testing strategies are employed to validate system functionality under different conditions. Unit testing is conducted to verify individual components, such as gesture detection and command transmission, ensuring that each module operates correctly in isolation. This approach helps identify errors early and simplifies debugging.

Integration testing is performed by combining different modules, including video capture, gesture recognition, and drone communication, to ensure seamless interaction between system components. This phase verifies that data flows correctly across modules and that gestures are accurately translated into drone movements. Performance testing evaluates system responsiveness by measuring parameters such as frame processing rate, gesture recognition speed, and command execution latency. The system is tested under varying environmental conditions to ensure consistent performance.Validation testing ensures that the system meets user requirements and behaves as expected in real-world scenarios. It involves verifying the accuracy of gesture recognition and confirming that each gesture triggers the correct drone action. Functional testing further examines whether all system features operate according to specifications, including gesture detection, command mapping, and drone control.Finally, security testing is conducted to ensure safe communication between the control system and the drone. This includes verifying that only authorized commands are accepted and that the system is resistant to issues such as signal interference or unauthorized access. Together, these testing methodologies ensure that the system is robust, reliable, and ready for real-world applications.

Results

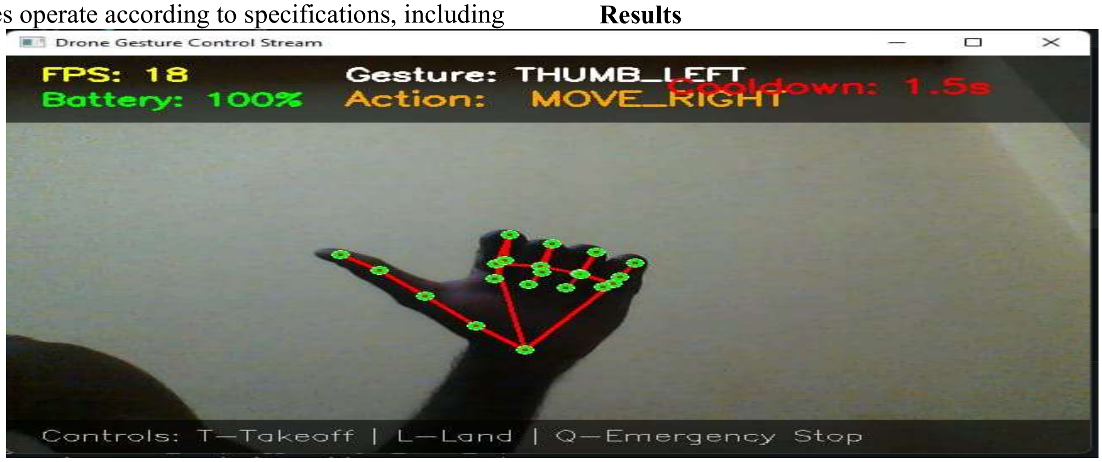

Fig 1 Move Right

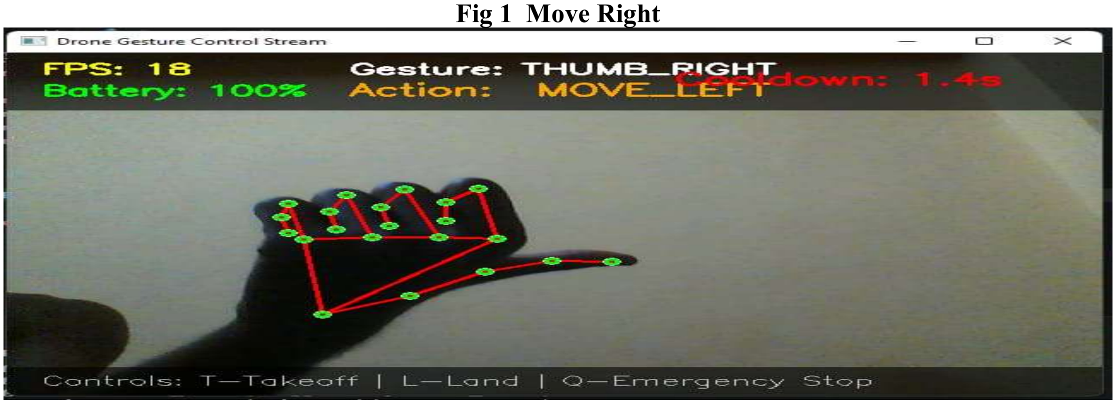

Fig 2 Move Left

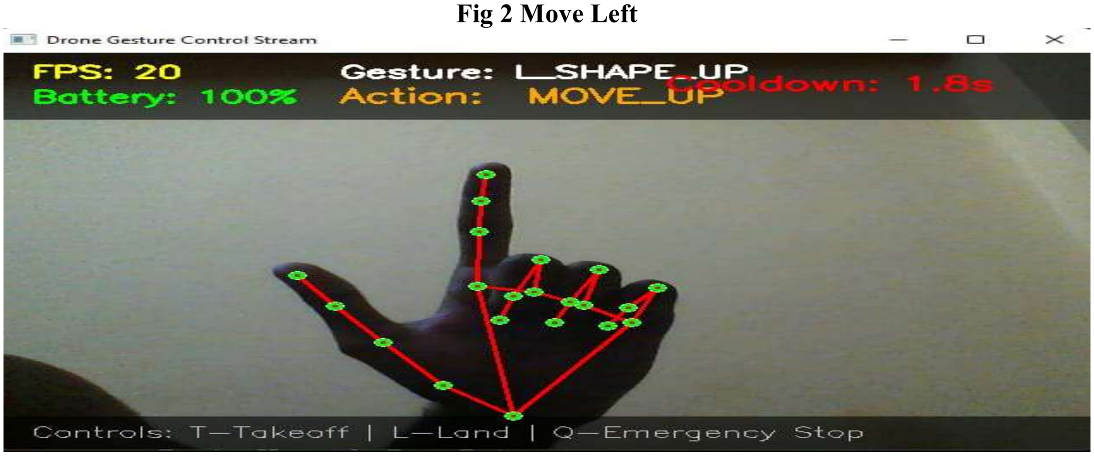

Fig 3 Move Up

ISSN2277-2685
IJESR/April. 2026/ Vol-16/Issue-21/776-781
Mrs. Kiran Pakmode *et.al.,*/International Journal of Engineering & Science Research

Fig 4 Move Back

Fig 5 Move Forward

Conclusion

The developed gesture-controlled UAV system demonstrates an effective approach for enabling natural and intuitive interaction between humans and aerial platforms. By leveraging real-time computer vision techniques, the system replaces traditional input devices with hand gesture-based commands, significantly improving usability and accessibility. The integration of OpenCV and MediaPipe Hands facilitates accurate detection and interpretation of hand movements, which are subsequently translated into control instructions for the DJI Tello drone.Experimental implementation confirms that the system performs efficiently in real-time scenarios, maintaining low latency and stable drone response under standard operating conditions. This work highlights the practical application of vision-based human–computer interaction in robotics and demonstrates how gesture recognition can serve as a viable alternative to conventional control interfaces. Furthermore, the project establishes a foundation for future advancements in intelligent UAV systems, contributing to the broader field of smart and interactive technologies.

Challenges Faced

Despite its effectiveness, the system encounters several practical limitations that influence overall performance. Real-time gesture recognition may become less reliable under varying lighting conditions, complex backgrounds, or limited camera quality, which can affect the accuracy of hand landmark detection. Processing delays may also occur when handling continuous video streams, particularly on systems with lower computational capabilities, resulting in reduced responsiveness. Additionally, maintaining a stable wireless connection with the drone can be challenging, as signal interruptions or interference may disrupt command transmission.

Another difficulty lies in designing robust gesture mappings, since visually similar hand configurations can sometimes be misclassified, leading to unintended drone actions. Ensuring operational safety is also critical, as incorrect gesture interpretation or system delays may cause unstable movements, especially in confined indoor environments. These challenges emphasize the importance of optimization, environmental control, and enhanced robustness in future implementations.

Future Scope

The proposed system provides a strong foundation for further research and development in gesture-based UAV control. Future enhancements may include the integration of advanced deep learning models to improve gesture recognition accuracy under diverse environmental conditions, including low-light or cluttered backgrounds. Expanding the system to support dynamic and multi-hand gestures could enable more complex command structures and smoother interaction.Incorporating additional input modalities, such as voice commands, can lead to a hybrid control system that improves flexibility and user experience. The integration of artificial intelligence techniques for object detection and tracking using OpenCV can further extend system capabilities, enabling features such as obstacle avoidance, autonomous navigation, and target following. Moreover, the framework can be scaled to support coordinated control of multiple drones, which is particularly useful in applications like surveillance, agriculture, and disaster response.Future developments may also focus on enhancing portability through mobile-based implementations, allowing gesture control via smartphones instead of dedicated computing systems. Strengthening communication security and reliability will be essential for deployment in critical environments. Overall, these advancements can transform the system into a more intelligent, adaptive, and scalable solution for next-generation UAV applications.

REFERENCES

1. K. Yaseen et al. (2024). *Next-Generation Dynamic Hand Gesture Recognition using MediaPipe, Inception-v3, and LSTM*. Electronics, 13(16), 3233.
2. G. Yun, H. Kwak, & D. H. Kim (2024). *Single-Hand Gesture Recognition for Drone Motion Control using RGB Cameras*. Applied Sciences, 14(22), 10230.
3. G. H. Kumar et al. (2024). *Human–Computer Interaction for Drone Control using MediaPipe-Based Gesture Recognition*. Proceedings of ICVTTS Conference.
4. A. H. Althubiti & H. Algethami (2024). *Dynamic Gesture Recognition using Transformer Models and MediaPipe*. International Journal of Advanced Computer Science and Applications.
5. B. J. Jo (2024). *Hand Gesture Classification for Mobile Virtual Reality using MediaPipe*. International Journal of Intelligent Systems and Applications in Engineering.
6. O. El Salibi & M. K. Uyguroğlu (2025). *Dual-Modality Drone Control using Vision-Based Gesture Recognition and Autonomous Navigation*. ICCRE Conference.