Gesture and Body Position Control for Lightweight Drones Using Remote Machine Learning Framework

Amna Hayat[1], Calvin Huan Li[2], Nicholas Prakoso[3], Renzheng Zheng[4], Amogh Wyawahare[5], and Jerry Wu[6]

Department of Electrical and Computer Engineering, University of Maryland, College Park, Maryland 20742, USA
{ahayat13, cli48, nprakoso, rzheng8}@terpmail.umd.edu, {wamogh, jerrywu}@umd.edu

Abstract. This study explores enhancing drone capabilities in low-light environments using gesture and body pose recognition. Deploying a Crazyflie 2.1 drone integrated with sensors and AI modules, we leverage Google's MediaPipe for real-time gesture detection and pose estimation under different lighting conditions. Compared to other traditional image-based methods, our approach with MediaPipe's landmark-based models highlighted significant improvements in accuracy and computational efficiency. Results demonstrate that the landmark-based models not only perform robustly across different lighting conditions but also reduce dependency on high-quality images, making them suitable for real-time applications in challenging visual conditions. The use of these models in the drone's autonomous navigation system shows potential improvements in how it operates and extends their use in important areas like search and rescue, military surveillance, disaster monitoring, and interactive technologies. This paper proves that landmark-based methods are better than traditional image classification, paving the way for more research into smart, responsive unmanned aerial vehicles for important uses including defense and safety.

Keywords: Drone · CNN · machine learning · gesture recognition

1 Introduction

Unmanned aerial vehicles (UAVs) are widely used in various missions, such as scientific research, industrial production inspection, precision agriculture, parcel delivery, and disaster rescue. Many novel drones and functional building blocks are under development with the advent of advanced materials (flexible materials), optimization and miniaturization of electronic and optical components (sensors, microprocessors), onboard processors (Nvidia Jetson, Intel NUC, Raspberry Pi, etc.), energy modules, and localization systems (GPS, SLAM, UWB, etc.). Thus, the functionality of drones is becoming more complex and intelligent due to the rapid advancement of artificial intelligence (AI), sensors, and onboard computation capability.

There are three major future trends for the future development of drones:

1. **Minimization and optimization:** Palm-sized or even micro and nano drones are designed to accomplish tasks at low power, low cost, and without space constraints. These small form factor autonomous drones have raised the interest of researchers to obtain more inspiration from biology and military purposes.
2. **Robust design of drones:** For outdoor usage, especially military purposes, structural and aerodynamic designs enable drones to obtain increased maneuverability and improved flight performance. For example, ensuring flight safety while dealing with challenging weather conditions, such as heavy winds and extreme temperature variations.
3. **Drone autonomy:** Drone autonomy enables autonomous navigation and task execution. This feature requires real-time perception, online planning, learning-based decision-making modules, and onboard computation capabilities.

This work focuses on a minimized lightweight drone system and autonomous exploration and navigation guided by human body recognition. In order to achieve a high recognition rate, we introduced a solution capable of a cloud-based or remote-based computing environment, which integrated advanced machine learning models for gesture recognition and body pose detection in real-time. The prediction results are fed into the drone flight control system, and the minimized bandwidth RF signals are transferred to the drone.

The contributions of this work include:

1. A minimized lightweight drone has been used to realize the goal in our system. The demand for hardware only requires a basic distance sensor and a low quality, low power, small form factor camera.
2. Machine learning models are built in a remote computer system. This implies we do not need to run the computationally expensive algorithms such as deep neural networks (DNNs) on the drone. Thus, the requirements for the processor, neural processing unit (NPU), GPU, memory, and PCB can be further reduced.
3. The landmark-based models not only perform robustly across different lighting conditions but also reduce dependency on high-quality images, making them suitable for real-time applications in challenging visual conditions or even night vision.
4. Enhancement of range anxiety: due to the above reduction in the requirements for hardware and computational power, the battery efficiency can be further increased very significantly. This is a very important consideration, especially for outdoor drone applications such as military, weather condition monitoring, agriculture, and environmental science.