Tech Science Press

Object Tracking-Based "Follow-Me" Unmanned Aerial Vehicle (UAV) System

Olubukola D. Adekola[1,*], Onyedikachi Kenny Udekwu[2], Oluwatobi Tolulope Saliu[2], Damilola Williams Dada[2], Stephen O. Maitanmi[1], Victor Odumuyiwa[3], Olujimi Alao[2], Monday Eze[2], Funmilayo Abibat Kasali[4] and Ayokunle Omotunde[2]

[1]Department of Software Engineering, Babcock University, Ilisan-Remo, Ogun State, Nigeria
[2]Department of Computer Science, Babcock University, Ilisan-Remo, Ogun State, Nigeria
[3]Department of Computer Science, University of Lagos, Akoka, Lagos State, Nigeria
[4]Department of Computer Science and Mathematics, Mountain Top University, Ogun State, Nigeria
*Corresponding Author: Olubukola D. Adekola. Email: adekolao@babcock.edu.ng
Received: 20 June 2021; Accepted: 29 July 2021

Abstract: The applications of information technology (IT) tools and techniques have, over the years, simplified complex problem solving procedures. But the power of automation is inhibited by the technicality in manning advanced equipment. To this end, tools deliberately combating this inhibition and advancing technological growth are the Unmanned Aerial Vehicles (UAVs). UAVs are rapidly taking over major industries such as logistics, security, and cinematography. Among others, this is a very efficient way of carrying out missions unconventional to humans. An application area of this technology is the local film industry which is not producing quality movies primarily due to the lack of technical know-how in utilizing these systems. This study therefore aim to devise an autonomous object tracking UAV system that would eliminate the complex procedure involved in stabilizing an aerial camera (aerial bot) midair and promote the creation of quality aerial video shooting. The study adopted Unified Modeling Language (UML) tools in modeling the system's functionality. The traditional Server-Client model architecture was adopted. The OpenCV library employed proved highly efficient in aiding the tracking procedure. The system provided a usable web controller which provides easy interaction between the pilot and the drone. Conclusively, investments in UAVs would enhance creation of quality graphic contents.

Keywords: Artificial intelligence; cinematography; unmanned aerial vehicles; drones; object tracking

1 Introduction

Artificial intelligence (AI) as a term was first coined by John McCarthy in 1956 at the first academic conference on the subject [1]. Reference [2] in a seminar "As We May Think" proposed a system which amplifies people's own knowledge and understanding. A few years later, Alan Turing wrote on the ability of machines to simulate human beings and perform intelligent operations, for instance, play Chess game

[3]. Reference [4] says "Technology goes beyond mere tool making; it is a process of creating ever more powerful technology using the tools from the previous round of innovation." Technology advancement is not proportionate to general expectations but AI as a research area attempts to revolutionize and enlarge the scope of functional capabilities. AI could mean building technologically advanced machines that perceive and respond to the world around them [5]. Specifically, UAVs are helping modern-day videography in augmenting traditional procedures of graphic content creation to extend the possibilities of fully eliminating lack of precision caused by human error in video content creation. A drone is an unmanned aerial vehicle (UAV) that can be remotely controlled from a ground control station for a vast number of purposes [6]. Due to man's never-ending quest to stretch the boundaries of photography by querying the constraints imposed upon the various axis with which images were captured, man has thus resorted to asking the question, "what would the world look like from the eyes of a bird?" That is, looking at the world from above? This gave birth to Aerial Photography and its fulfillment with a rather more advanced technology-Unmanned Aerial Vehicles (UAV-Drone technology). Reference [7] suggested that aerial view be understood in a fluid relational context. Aerial Photography has been used for numerous disciplines such as mapping of geographic locations, artistic endeavours, search and rescue missions [8]. Drones have traversed the limitations of a traditional on-ground camera, launching a new era in photography; eliminating the need to incur exorbitant bills in using helicopters.

The aim of this study is to devise an efficient means of creating graphic content through augmenting existing UAV technology with artificial intelligence. This explores a possibility of revolutionizing modern-day cinematography, taking advantage of the rapid technological progression by specifically creating and deploying specially equipped UAVs capable of autonomous cinematography. More specifically, the UAV works based on the principle of computer vision as a branch of artificial intelligence where the movement of an object precisely a face of an individual is tracked with the help of object tracking algorithms implemented in the interface of the drone. The drone provides an on-point video feed or snapshots of the face being tracked. The video feed would then be returned to a monitor where video editors would analyze it and modifications can be implemented to the video if necessary. As a result, smart unmanned aerial vehicles would open new horizons of data utilization.

2 Problem Statement

The entertainment and filming industries do successfully present events or acting for the world to see but the question is "are these stories told precisely or are these moments captured correctly?" The quality of cinematographic contents has declined because it has not met the rising demands of the consumers. Many local film industries are at low ebb because they want to maximize "profit" rather than going the extra mile to achieve "high quality" contents. Although Aerial cinematography is being implemented to upgrade the quality of content produced by the entertainment industry but safely piloting a drone while filming a moving target in the presence of obstacles; counter-acting the forces of nature like wind so as to get precise camera shots is highly demanding and requires a lot of concentration, often requiring multiple human expert operators. Therefore, this article devises an efficient UAV that represents the creation of a smart unmanned hovering vehicle that would foster autonomous creation of content requiring live and dynamic camera viewpoints, capturing highlight moments in exact, with on-point precision.

3 Literature Review

Recent upsurge in AI research as a result of increasing computing power, Big Data, and powerful optimization algorithms have culminated into creating a new epoch in the world by making intelligent machines. Andrew Ng of Stanford called it the new electricity [9]. The first commonly adopted drone emerged in 1935 as a full-size retooling of the de Havilland DH82B "Queen Bee" biplane [10]. The term

drone dates to the present first use, a play on the "Queen Bee" terminology. In December 2016, Amazon delivered its first Prime Air parcel in Cambridge, England. In March of 2017, it demonstrated a Prime Air drone delivery in California [11]. Autonomous flying machines or drones use computer vision systems to float in the air, dodging obstacles to stay going in the right direction. AI is being used in drones to make this flying machine smarter. These devices were first created manually, and remotely controlled.

3.1 Technology Involved in AI Drones

While drones serve a variety of purposes, such as recreational, photography, commercial and military, their two basic functions are flight and navigation. To achieve flight, drones consist of a power source, such as battery or fuel, rotors, propellers, and a frame. The frame of a drone is typically made of lightweight, composite materials, to reduce weight and increase maneuverability during flight. Drones require a controller, which is used remotely by an operator to launch, navigate and land it. Controllers communicate with the drone using radio waves, including Wi-Fi. The propeller and the motor function to lift the craft the antenna to facilitate communication between the drone and the controller, the GPS to locate the drone's position and to facilitate tracking procedure [12]. Sensors are used to capture all data processed by drone systems, including visual, positioning, and environmental data. On machine perception, several AI-related activities for drones deal with image detection and recognition, the unmanned aerial vehicle must be able to interpret and absorb information from the environment or objects. This is normally achieved with electro-optical, stereo-optical, and LiDAR (Light Detection and Ranging) sensors. This mechanism is called the Vision Machine [13]. Computer vision is about the drone processing raw sensor data collected. That is the automated retrieval, interpretation, and comprehension of useful information from one or more images. Computer vision plays a crucial role in identifying different types of objects when flying in mid-air. High-performance onboard image processing and a drone neural network are used to identify, recognize and track objects when flying in the air. Machine Learning methods can be used to refine differentiable parameters. Unlike software that has been programmed manually, machine learning algorithms are designed to learn and evolve when introduced to new data. Deep learning is a specialized technique of information processing and a branch of machine learning that uses neural networks and a vast volume of data for decision-making. Motion Planning is a valuable weapon for situational awareness; to capture the atmosphere. The drone visualizes the environment, e.g., with SLAM (Simultaneous Localization and Mapping) technology. In the sense of Motion Planning, Deep Learning is used to identify and classify artefacts such as humans, bikers, or vehicles and to establish an appropriate flight path.

3.2 Review of Existing Technologies

The following is a list of notable technologies related to the subject matter:

  (i) *Ardupilot*: Ardupilot is an open-source UAV ground control station developed by Michael Oborne in C# programming language and planned specifically to support the Ardupilot Mega Autopilot project [14].

 (ii) *FA6 Drone:* FA6 Drone is a face-recognition software that detects suspects, missing persons, and pedestrians from a drone camera. FA6 Drones is run in real-time, can process captured images and has ability to instantly distinguish individuals from the air.

(iii) *Sensory Technology:* Aerial platforms are heavily influenced by their mounted sensor loads. The sensors used should then satisfy various criteria, such as weight constraints and ease of installation, when performing the required tasks [15].

(iv) *Generalized Autonomy Aviation System:* Generalized Autonomy Aviation System (GAAS) is an open-source autonomous aeronautical software application designed for fully autonomous drones

and flying cars. GAAS was developed to provide a popular computer-based drone intelligence infrastructure.

(v) ***Drone Volt:*** Drone Volt provides object detection, counting, segmentation, and tracking. It offers two solutions to run AI-based applications on the ground station during or after the flight, namely Online: real-time video processing and Offline: post-flight video processing.

3.3 Review of Related Works

Reference [16] presented a UAV scenario characterized by search, detect, geo-locate and automatic landing components. In this context, for relatively high speed, a stabilizing system is required to protect the UAV from flipping or crashing. Reference [17] proposed a multi-staged framework that employed two computer vision algorithms that extract the position and size of an aircraft from a video frame. The study saw the need to further design a framework with a collision avoidance algorithm for safety purposes. Reference [18] established that enabling indoor follow-me function is very difficult due to the lack of an effective approach to track users in indoor environments, therefore, the researchers designed and implemented a system that accurately tracks drone and controls its movement to maintain a specified distance and orientation to the user for auto-videotaping. Reference [19] registered that inaccuracy of the GPS data makes targeting performance unsatisfactory. The study considered visual tracking but the reliability of visual tracking on long-term tracking in unexpected operating environments could be problematic. The study proposed a hybrid method combining the high accuracy of a visual tracking algorithm in short-term tracking and the reliability of GPS-based type in long-term tracking. The implemented prototype has limitations for commercial use. Reference [20] presented a two-stage UAV tracking framework comprising a target detection stage based on multi-feature discrimination and a bounding-box estimation stage based on the instance-aware attention network. There is difficulty of tracking the target robustly in UAV vision task. The future direction is in improving the loss function to suppress the expansion of bounding-box anomalies in the optimization process. The work of [21] focused on considering trackers' capabilities to be deployed in the dark. It depicts a discriminative correlation filter-based tracker with illumination adaptive and anti-dark capability. Reference [22] reported a considerable progress in the development of highly manoeuvrable drones for applications like monitoring and aerial mapping. The challenging parts are real-time response, workload management, and complex control.

4 Materials and Methods

This study embodies case studies; systematic literature reviews such that relevant documents obtained were qualitatively analyzed for convergence, and useful details extracted. The design of various components of the system was done using 2-tier architecture and modelling done using unified modelling language (UML) tools. The system followed iterative development model. An effective object tracking algorithm was implemented to aid computer vision. Open CV tool was used to facilitate smooth deployment of the object tracking algorithm. A web platform for managing data transmission between the pilot and the drone and vice versa was implemented using sockets networking in Python Programming language to maximize speed, ensuring little or no latency in transmitting information.

The functional requirements include pilot being able to have a real-time update on flight data, receive live video feed from the drone. Drones should be able to easily switch between modes without crashing. Drones should be able to automatically manipulate its position i.e., devoid of human interaction with the system.

The non-functional requirements include the latency time for delivering pilot instructions to the drone as well as transmitting telemetry information from drone to pilot which must not exceed five seconds. Server

WebSocket should be able to route data via UDP (User Datagram Protocol) and TCP (Transmission Control Protocol) where needed, effectively. There should be at most a one second delay in relaying telemetry data to and from the flight controller. Pixhawk flight controller was used.

4.1 The Logical Design

The logical design in Fig. 1 represents the interaction between the components of the drone and how data flows from the input layer, processing till designated action by the drone. The drone perceives the environment through its sensor (camera); feeding this to the computer onboard which then executes the face tracking program or algorithm on the perceived data, serving the video feed to the client via the Flask server. The onboard computer also receives commands sent to the drone from the web application via the Flask server; executing these commands by sending movement controls to the flight controller of the drone which basically is the brain of the drone's locomotion system thus making the drone move on command. However, commands could be sent to the drone directly without having to go through the onboard companion computer. This is because these commands are already in a form already understandable by the flight controller's firmware. The commands are from the radio transmitter, received by the receiver and then sent to the flight controller, finally instructing the actuators to move respectively. An Object Detection and Avoidance system was implemented as well in this build by using proximity sensors, in this case, LiDARs that aid the detection of obstacles within the path of travel of the vehicle. The obstacle avoidance procedure is executed in parallel with the object tracking procedure. This way, the UAV tracks efficiently via avoiding obstacles along its path the target object to be recorded.
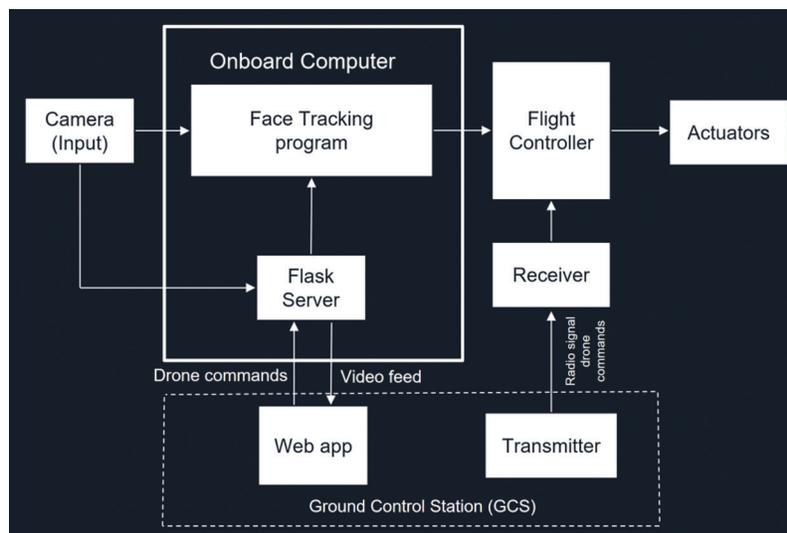

Figure 1: Block diagram showing the logical design of the proposed system

4.2 The Physical Design

Fig. 2 represents the mode of operation of the "Follow me" procedure. A proportional Integral and Differential (PID) controller is used where the Proportional controller keeps record of the "Present error", Pe, which is the current distance away from the target (face). The Integral keeps record of the "Past error", Pu, while the Differential controller makes calculations to determine the future error Pv. For every movement in 3-D (three-dimensional) space made by the drone, there is a PID controller that controls the velocity of the drone; increasing and decreasing the velocity when necessary. Considering the movement

along the Y-axis, the drone is to get to a particular point which is the Y coordinate of the face's center, the initial error becomes like 20 cm representing the vertical displacement of the face's center from the camera frame's center. The following was used to calculate the *speed* with which the drone should approach its target:

$$speed = (PIDm * Pe) + (PIDm * (Pe - Pu)) \tag{1}$$

where,

speed represents the velocity of arrival

PIDm represents the PID multiplier variable

Pe and Pu represent the present and past error respectively
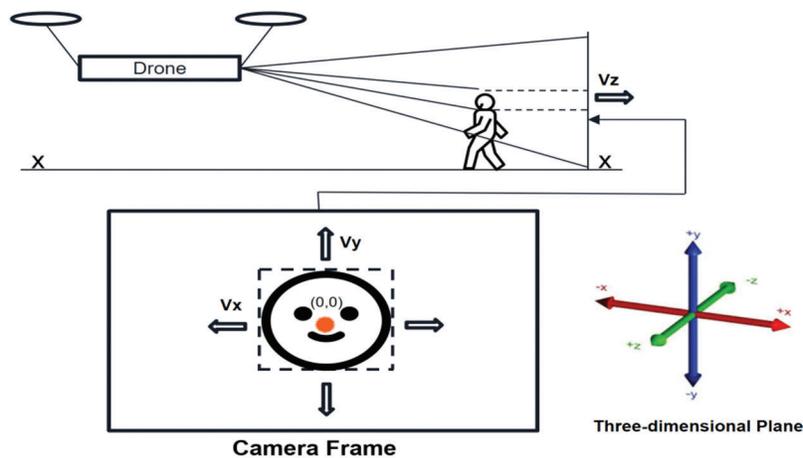
(Pe-Pu) represents Pv the future error

Figure 2: Face tracking procedure schematics

A good point to note here is that the present error Pe for a particular time becomes the past error Pu for a future time. In other words, Pe could refer to Pu. So, it is only logical to infer that upon initial closure to the target the speed of the drone will be at its max; subsequently reducing as the drone approaches the target due to the Integral controller reducing the value it's been adding to the original speed value generated by the Proportional controller. Thus, leading to a controlled approach to the target–autonomous deceleration. The same principle seen on the Y-axis applies to every other axis in 3-D space.

4.3 System Modeling Using Unified Modeling Language (UML)

Fig. 3 shows the use case of the pilot; specifying the functions the pilot can toggle on the system. These functions include:

Manual Maneuvers: In this case, the pilot is able to send commands (PWM values that control the rotation of the motors of the drone) to different channels of a radio communication transmitter.

Setup: The primary objective of the drone is to track an object. Here, the pilot is expected to have an object (human face) ready; having the camera point towards the direction of the object.

Initialization: This process entails having the pilot turn on the drone; send the takeoff command via the GUI; enable the "face detect and tracking" feature.

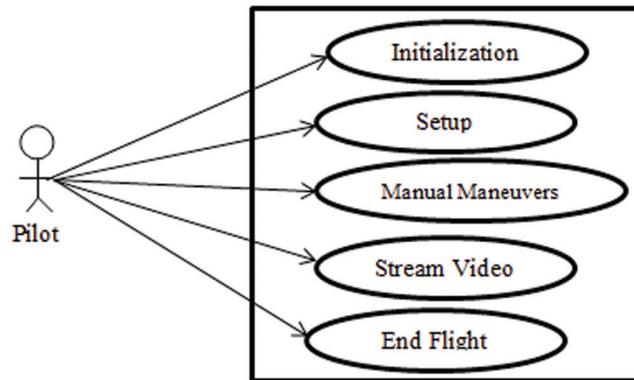**Stream video:** The pilot receives and streams video feed transmitted via a video streaming server.

Figure 3: Pilot's use case diagram

End flight: The pilot disables the autonomous tracking feature of the drone; issuing a command to the drone to return to its original takeoff location or land, having the drone vertically descends to the ground.

The drone's use case diagram in Fig. 4 specifies the function the drone performs as a part of the system; responding by executing commands received from the pilot. These functions include:

Auto takeoff: Here, the drone is able to automatically ascend to a default altitude on command issued by the pilot via the web application.

Detect and track object: The drone autonomously with the help of AI-facilitated hard-coded programs adjust its position in 3-D space to accurately move in response to movement of the object being tracked thus following the object.
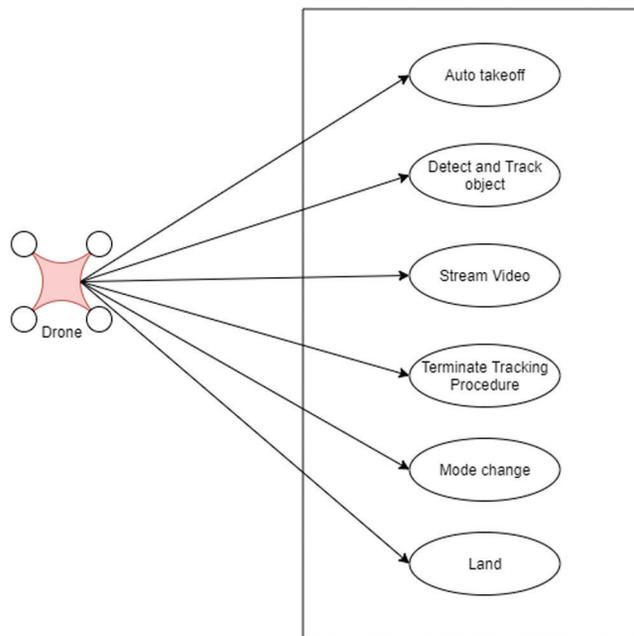

Figure 4: Drone's use case diagram

Land: As instructed by the pilot upon termination of the flight/mission, the drone automatically descends gradually by a defined vertical speed; safely returning to land.

Stream video: As requested by the pilot via the "detect and track" function, the drone returns continuously frames captured of the moving object being tracked to the pilot via a streaming server. This video could also be returned even without having the pilot toggle the "detect and track" feature.

Terminate tracking procedure: The drone having received the command to seize executing the tracking function, returns to "stabilize mode"; enabling the pilot take manual control of the drone.

Mode change: The drone will be able to switch between modes as commanded by the pilot via the RC transmitter. Autopilot modes such as RTL (Return To Launch), Loiter, Alt Hold and so on will be deployable by the drone.

Fig. 5 represents the use case diagram for the object being tracked; standing in as one of the primary actors, for without it there would be nothing to track, hence the system becomes impractical.

Setup: To initiate the tracking procedure, the face to be tracked would have to be aligned in such a way that it is within the line of sight of the camera.

Tracking: This use case delineates actions to be executed by the object to instigate the "follow me" procedure of the drone, e.g., having the drone move in response to the changing location of the object.

Object out of range: Here, the tracking procedure is automatically terminated due to the inability of the drone's camera to perceive the object. In other words, when the object moves to a location deemed as far away from the drone or out of the drone's camera range, the tracking procedure will as earlier stated be automatically terminated.
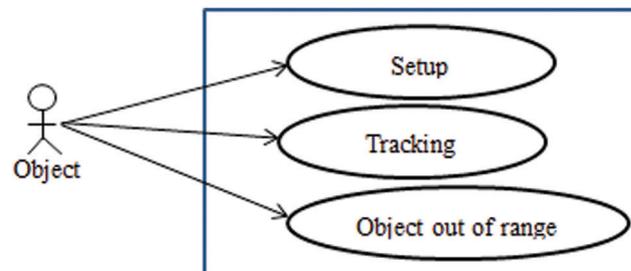

Figure 5: Object's use case diagram

The following Tab. 1 represents a further illustration on the major use cases in the system design.

The following diagram Fig. 6 depicts interactions and relationships between the various actors involved: the drone, the pilot and the object being tracked.

4.4 Sequence Diagram for the Flow of Activities in the System

The sequence diagram here shows the flow of interaction between various objects of the system arranged in a time sequence. The Tracking Sequence Diagram in Fig. 7 shows the sequence of activities needed to be executed by the pilot/the object to be followed as well as the actions executed by the drone to facilitate the overall tracking procedure.

The Mode Switch Sequence Diagram in Fig. 8 shows the sequence of activities that a pilot can perform to have the drone toggle between modes; switching from autonomous mode to manual mode.

The pilot streaming live drone camera broadcast sequence diagram in Fig. 9 shows the sequence of activities executed by the pilot in order to receive live real-time video feed from the drone and take snapshots; having the aerial photographs displayed back to the pilot.

Table 1: Use case narrative for most common use cases in the system

Use case	Description	Actors involved
Manual maneuvers	The pilot manually controls the drone using an RC transmitter that sends PWM values to the drone; controlling the speed of the motors.	Pilot, drone
End flight	The pilot initiates a command; asking the drone stops all running functions and gradually descends to land.	Pilot, drone
Auto takeoff	The drone gradually ascends to a target altitude upon receipt of command from the pilot.	Pilot, drone
Detect and track object	The drone detects and follows a target object by constantly adjusting its position in such a way that the object in the image is kept at the center of the camera frame.	Pilot, drone
Land	The drone gradually descends to the ground upon receipt of command from the pilot.	Pilot, drone
Stream video	A drone broadcasts live video feed from its camera while the pilot streams it.	Pilot, drone
Terminate tracking procedure	The drone stops following an object.	Drone
Mode change	The pilot *via* the RC transmitter can request to change the current mode of the drone from say "Guided" to "Stabilize", thus switching controls from automatic to manual.	Pilot, drone

The drone broadcasting camera feed sequence diagram in Fig. 10 shows the sequence of activities a drone performs to broadcast image frames captured continuously to a video streaming server for a pilot to stream.

5 System Architecture

The system architecture adopted is the traditional Server-Client model. The client tier represents the layer of interaction between the pilot(s) and the autonomous drone. It comprises a single page application (for the pilots) well defined and Application Programmable Interface (API) utilized by the drone. The pilots constitute the primary users as they would interact with the system using PCs (Personal Computers) running on any modern-day Operating System. The Server/Application tier comprises both the business logic responsible for the full operational activities of the drone as well as how commands are executed by the drone via the server on board the drone. It contains stored procedures to transmit information to-and-fro the drone such as video feeds which is facilitated by the Flask server and then commands as well also leveraging on the Flask server. It also contains the algorithm responsible for quality dissemination of information to the client interacting with the drone. The Flask server basically routes commands received from the frontend to the respective actions (functions) to execute such commands in the backend. OpenCV utilized is an open-source software library built on the grounds of the Artificial Intelligence paradigm Computer Vision and Machine Learning. It comprises algorithms that facilitate certain intricate procedures like face recognition, object identification, classification of various human actions in videos, camera movement tracking, object tracking, extraction of 3D models of objects, image stitching for production of high-resolution images, even more specifically, retrieve similar images from an image

database, eye movement tracking and their likes. For this study, opencv-python version 3.4.1.15 proved highly efficient in aiding the tracking procedure.
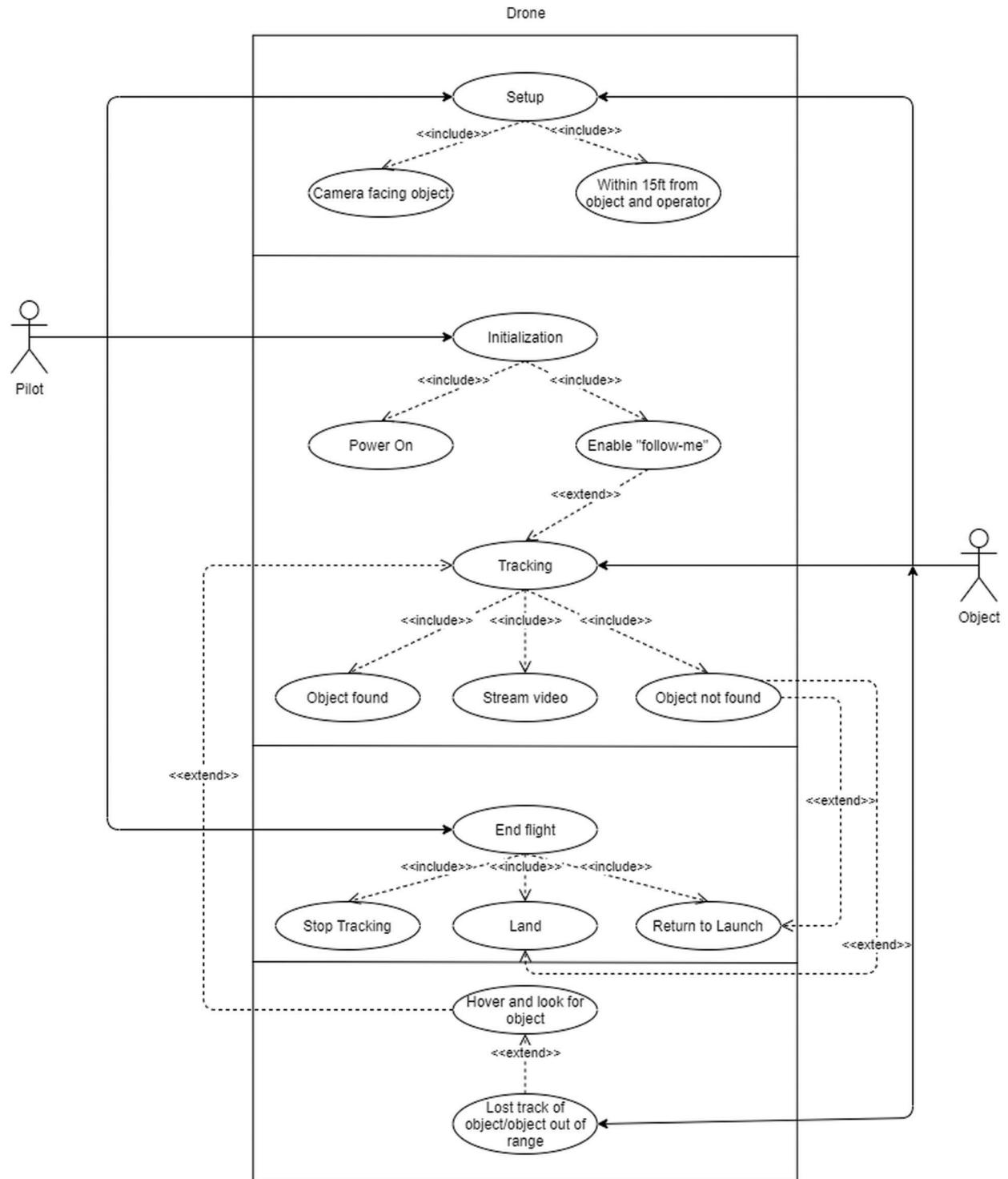

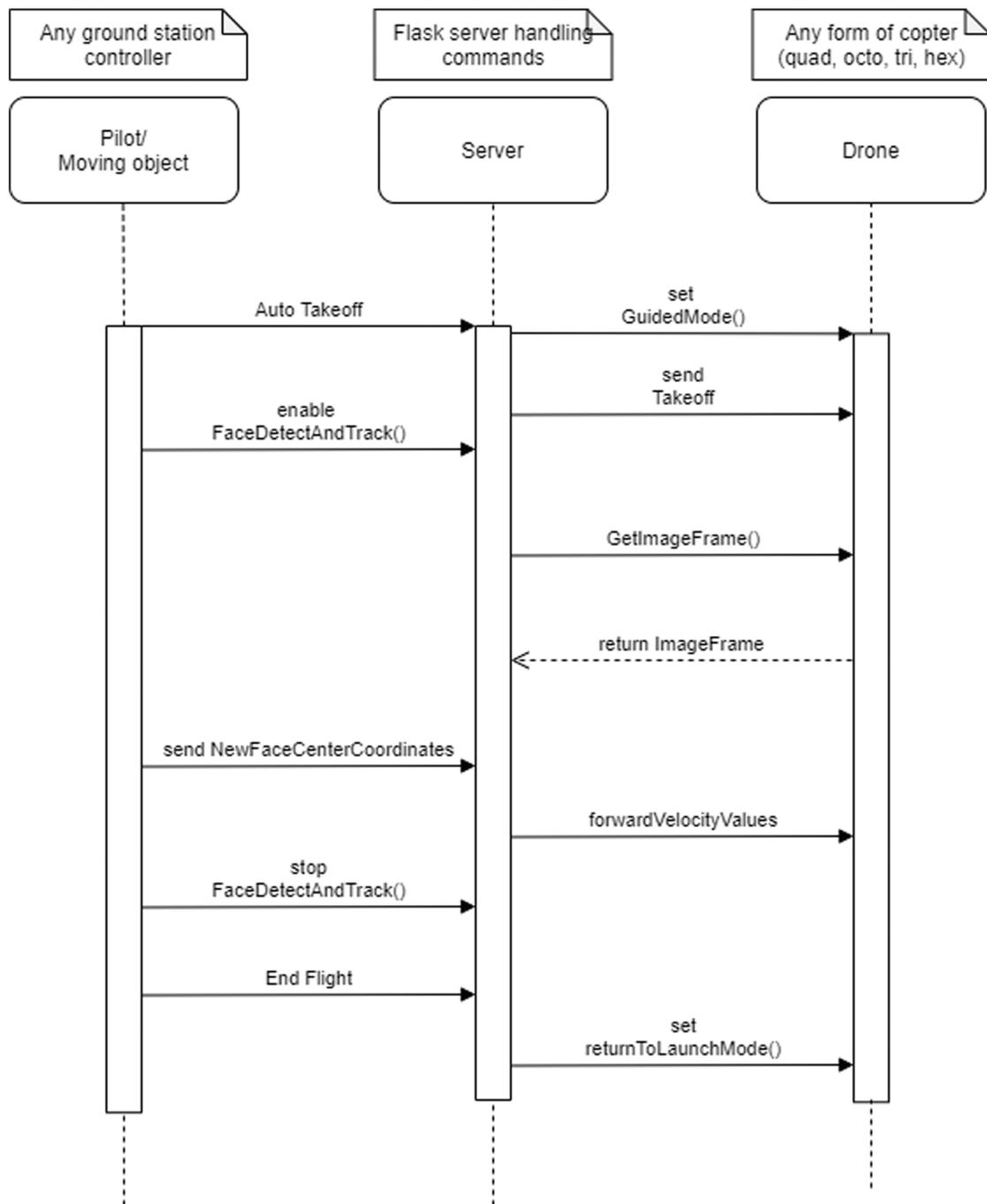**Figure 6:** Interaction between the drone, the pilot and the tracked object

Figure 7: Object tracking sequence diagram

6 Results and Discussion

This section presents an explanation of how a pilot can interact with the UAV via a simple single webpage application (web controller); "Drone World". By toggling the "Takeoff" switch as seen in the following Fig. 11, the pilot is able to initiate a controlled levitation of the vehicle to a target altitude of about 2 m effortlessly; eliminating the complex activities required to manually levitate and stabilize the

drone in midair. Just like in the "Initiate Takeoff" procedure, the land procedure is done the same way. The pilot seamlessly instructs the drone to execute the land function by toggling the "Land" switch. Once toggled, the drone immediately starts to descend at a controlled vertical speed.
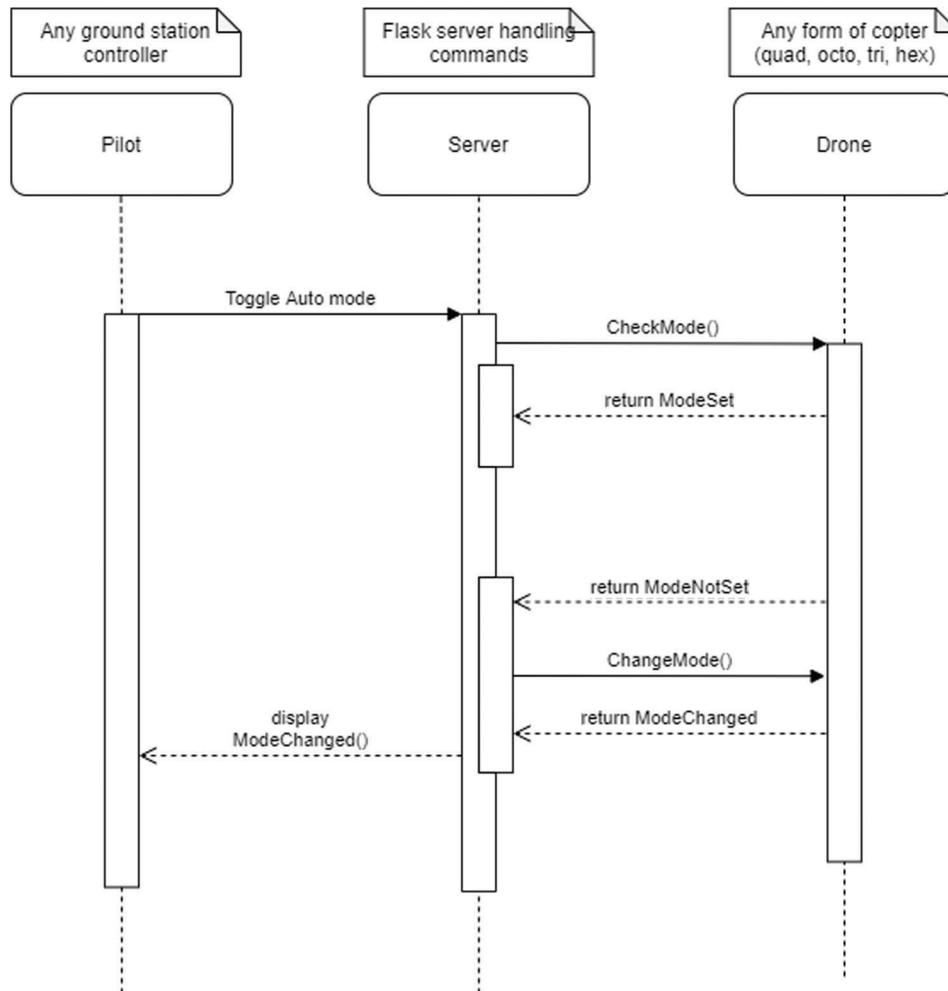

Figure 8: Mode change sequence diagram

The movement of the drone while airborne could be controlled with the soft buttons seen on the web application where the direction 'up' and 'down' buttons control the rise and descent of the drone respectively; the 'left' and 'right' direction buttons control the roll of the drone; the 'forward' and 'backward' buttons control the pitch of the drone. Lastly, the 'clockwise' and 'anticlockwise' direction buttons control have the drone rotates clockwise or anticlockwise. To have the drone fly autonomously, the "Follow face" or "Follow body" switch is to be toggled. The switches are mutually exclusive from the other. The drone upon receiving the 'Follow face' command would detect and track an individual's face while in the "Follow body" mode, the UAV would track a different object: an individual's entire body. Both operations are deactivated using the "stop Follow face" and "stop Follow body" switches.
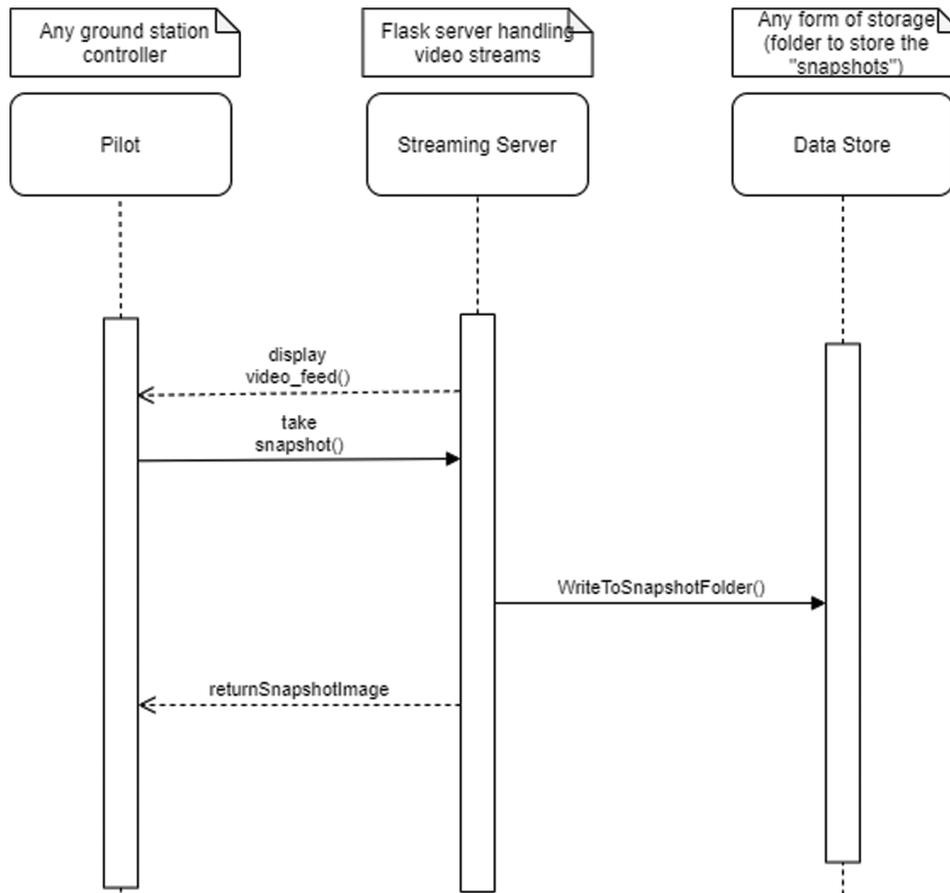
Figure 9: Pilot streaming live drone broadcast sequence diagram

The UAV has three modes incorporated into the web controller namely: Manual mode, Auto mode and Return to home mode; each triggering a specific movement of the drone. The Manual mode has the drone maintain a particular position in 3-D space. The drone autonomously adjusts its position depending on how it is displaced from its initial location. The Auto mode creates an avenue through which the drone would execute specific autonomous activities like the "Follow face" and "Follow body" procedures. The Return to home mode commands the drone to return to its launch location. In execution of this command, the drone climbs to an altitude of 15 meters to clear any obstacle in close proximity before proceeding to the initial launch location. Interestingly, drone is able to switch between modes effortlessly.

The drone is however able to return live video feed via HTTP (Hyper Text Transfer Protocol) using a technology called "mimetype". This works by sending video binary data over HTTP, decoding and then displaying the decoded data which is now a continuous stream of image frames on the UI (User Interface) for the pilot to make do. The video stream automatically starts once the "controller" page is requested for whereas aerial photographs need be taken manually by toggling the "Snapshot" switch. Once the Snapshot switch has been toggled, the drone stores a single frame of the stream of frames returned as the video feed and then displays the captured frame just below the Snapshot button while having that same frame stored locally on the drone's onboard computer.
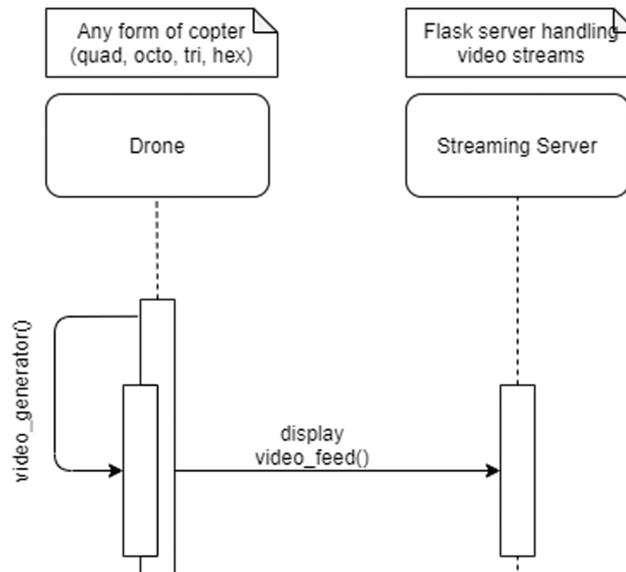
Figure 10: Drone broadcasting camera feed sequence diagram

Figure 11: Toggling auto takeoff mode showing interface sample view

7 Conclusion and Recommendation

The Drone system was built using a handful of technologies that ensure scalability and accessibility via the browser, with a huge reduction in latency time. This is a pointer to developing tools with sustainable advantages in cinematic and photography works. This study adopted UML tools in modeling the system's functionality. Systematic review of literature enhanced the conceptualization of the system design.

The state-of-the-art in drone technology shows a wide applicability which is quite promising. Investment in UAVs would serve as a primary tool for cinematographers in creating quality controlled graphic contents with ease. The researchers recommend the development of a web infrastructure to facilitate authorization of pilots intending to control the drone from anywhere around the globe. Also, addition of more advanced sensors including ultrasonic sensors, thermal sensors, and collision avoidance sensors would make a difference. Addition of a gyroscope and vibration dampeners would aid a better focus on the object being or to be tracked, thus improving the efficiency of the "follow-me" procedure. To have drones with potentials of becoming monitoring mechanisms, there is a need to create designs that can thrive long hours. Also, solar panels or other energy sources could be incorporated to generate power.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] J. McCarthy, "Review of artificial intelligence: A general survey," 2000. [Online]. Available: http://wwwformal. stanford.edu/jmc/reviews/lighthill/lighthill.html.

[2] V. Bush, "As we may think," *The Atlantic Monthly*, vol. 176, no. 1, pp. 101–108, 1945.

[3] A. Turing, "Computing machinery and intelligence," *Mind*, vol. 49, no. 236, pp. 433–460, 1950.

[4] R. Kurzweil, "The law of time and chaos," in: *The Age of Spiritual Machines: When Computers Exceed Human Intelligence*. New York: Penguin Books, Viking Press, pp. 24–37, 1999.

[5] S. J. Russell and P. Norvig, "Solving problems by searching," in: *Artificial Intelligence: A Modern Approach*, 3rd ed., Upper Saddle River, New Jersey: Prentice Hall, pp. 64–75, 2010.

[6] J. Hu and A. Lanzon, "An innovative tri-rotor drone and associated distributed aerial drone swarm control," *Robotics and Autonomous Systems*, vol. 103, no. 3, pp. 162–174, 2018.

[7] P. Amad, "From god's-eye to camera-eye: Aerial photography's post-humanist and neo-humanist visions of the world," *History of Photography*, vol. 36, no. 1, pp. 66–86, 2012.

[8] J. L. Morgan, S. E. Gergel and N. C. Coops, "Aerial photography: A rapidly evolving tool for ecological management," *BioScience*, vol. 60, no. 1, pp. 47–59, 2010.

[9] S. Lynch, "Why AI is the new electricity, Stanford Graduate School of Business. Stanford: CA., Retrieved from, 2017. [Online]. Available: https://www.gsb.stanford.edu/insights/andrew-ng-why-ai-new-electricity.

[10] B. Zimmer, "The flight of 'drone' from bees to planes," *Wall Street Journal*, 2013. [Online]. Available: https://www.wsj.com/articles/SB10001424127887324110404578625803736954968, Retrieved: April 30, 2021.

[11] D. Attard, "The history of drones: A wonderful, fascinating story over 235+ years, Dones Buy.Net," 2017. [Online]. Available: https://www.dronesbuy.net/history-of-drones, Retrieved: October 25, 2019.

[12] G. D. Maayan, "How do AI-based drones work? The main benefits and potential of using A.I. with drone systems, Heartbeat by Fritz AI," 2020. [Online]. Available: https://heartbeat.fritz.ai/how-ai-based-drones-work-a94f20e62695#, Retrieved: April 30, 2021.

[13] V. S. Bisen, "What is the difference between AI, machine learning and deep learning?," 2020. [Online]. Available: https://www.vsinghbisen.com/technology/ai/difference-between-ai-machine-learning-deep-learning/, Retrieved: April 30, 2021.

[14] Ardupilot, "APM 2.0 planner," 2014. [Online]. Available: http://planner2.ardupilot.com/, Retrieved: April 19, 2021.

[15] D. Roca, S. Lagüela, L. Díaz-Vilariño, J. Armesto and P. Arias, "Low-cost aerial unit for outdoor inspection of building façades," *Automation in Construction*, vol. 36, no. 4, pp. 128–135, 2013.

[16] S. Kyristsis, A. Antonopoulos, T. Chanialakis, E. Stefanakis, C. Linardos *et al.,* "Towards autonomous modular UAV missions: The detection, geo-location and landing paradigm," *Sensors*, vol. 16, no. 1844, pp. 1–13, 2016.

[17] K. R. Sapkota, S. Roelofsen, A. Rozantsev, V. Lepetit, D. Gillet *et al.,* "Vision-based unmanned aerial vehicle detection and tracking for sense and avoid systems," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, Daejeon, South Korea, pp. 1556–1561, 2016.

[18] W. Mao, Z. Zhang, L. Qiu, J. He, Y. Cui *et al.,* "Indoor follow me drone," in *Proc. of the 15th Annual Int. Conf. on Mobile Systems, Applications, and Services-MobiSys '17*, Niagara Falls, NY, USA, pp. 345–358, 2017.

[19] T. Tuan Do and H. Ahn, "Visual-GPS combined 'follow-me' tracking for selfie drones," *Advanced Robotics*, vol. 32, no. 19, pp. 1047–1060, 2018.

[20] S. Zhang, L. Zhuo, H. Zhang and J. Li, "Object tracking in unmanned aerial vehicle videos via multifeature discrimination and instance-aware attention network," *Remote Sensing*, vol. 12, no. 2646, pp. 1–18, 2020.

[21] B. Li, C. Fu, F. Ding, J. Ye and F. Lin, "All-day object tracking for unmanned aerial vehicle," 2021. [Online]. Available: https://arxiv.org/pdf/2101.08446.pdf.

[22] V. Kangunde, R. S. Jamisola and E. K. Theophilus, "A review on drones controlled in real-time," *International Journal of Dynamics and Control*, vol. 2021, pp. 1–15, 2021.