

YOLOv8-Based Autonomous Ball Detection and Tracking for Rotorcraft UAVs Using Onboard Vision Sensors [†]

Oumaima Gharsa ^{1,*}, Mostefa Mohamed Touba ¹, Mohamed Boumehraz ², Nesrine Abderrahmani ¹,
Imam Barket Ghiloubi ¹ and Abdel Hakim Drid ¹

¹ Identification, Command, Control & Communication Laboratory (LI3CUB), Mohamed Khider University, 07000 Biskra, Algeria; email1@email.com (M.M.T.); email2@email.com (N.A.); email1@email.com (I.B.G.); email2@email.com (A.H.D.)

² Modeling of Energy Systems Laboratory (LMSE), Mohamed Khider University, 07000 Biskra, Algeria; email1@email.com (M.B.)

* Correspondence: oumaima.gharsa@univ-biskra.dz

[†] Presented at the 12th International Electronic Conference on Sensors and Applications (ECSA-12), 12–14 November 2025; Available online: <https://sciforum.net/event/ECSA-12>.

Abstract

Drones equipped with onboard cameras offer promising potential for modern digital media and remote sensing applications. However, effectively tracking moving objects in real time remains a significant challenge. Aerial footage captured by drones often includes complex scenes with dynamic elements such as people, vehicles, and animals. These scenarios may involve large-scale changes in viewing angles, occlusions, and multiple object crossings occurring simultaneously, all of which complicate accurate object detection and tracking. This paper presents an autonomous tracking system that leverages the YOLOv8 algorithm combined with a re-detection mechanism, enabling a quadrotor to effectively detect and track moving objects using only an onboard camera. To regulate the drone's motion, a PID controller is employed, operating based on the target's position within the image frame. The proposed system functions independently of external infrastructure such as motion capture systems or GPS. By integrating both positional and appearance-based cues, the system demonstrates high robustness, particularly in challenging environments involving complex scenes and target occlusions. The performance of the optimized controllers was assessed through extensive real-world testing, involving various trajectory scenarios to evaluate the system's effectiveness. Results confirmed consistent and accurate detection and tracking of moving objects across all test cases. Furthermore, the system exhibited robustness against noise, light reflections, and illumination interference, ensuring stable object tracking even when implemented on low-cost computing platforms.

Keywords: unmanned aerial vehicle (UAV); YOLOv8; tracking; camera sensors; moving target

Academic Editor(s): Name

Published: date

Citation: Gharsa, O.; Touba, M.M.; Boumehraz, M.; Abderrahmani, N.; Ghiloubi, I.B.; Drid, A.H. YOLOv8-Based Autonomous Ball Detection and Tracking for Rotorcraft UAVs Using Onboard Vision Sensors. *Eng. Proc.* **2025**, volume number, x. <https://doi.org/10.3390/xxxxx>

Copyright: © 2025 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Autonomous drones often called commercial micro aerial vehicles (MAVs) have long been integral to robotics and are widely deployed for surveying, surveillance, search-and-rescue, military operations, and even package delivery. Yet achieving fully autonomous

quadrotor navigation including stable flight, reliable tracking, and robust obstacle avoidance remains a difficult challenge.

In recent years, unmanned aerial vehicles (UAVs) have seen growing adoption across many applications, driven by their low operating costs and strong performance. Cameras, in particular, have emerged as a practical, low-cost sensing option that enables the estimation of rich, information-dense representations of the environment [1]. In computer vision, visual tracking focuses on automatically estimating an object's bounding box in each subsequent frame [2]. Despite strong progress, many high-performing visual tracking methods remain too computationally heavy for real-time use on UAVs, where long-term tracking must robustly handle fast motion, scale changes, heavy occlusion, and out-of-view events while operating within strict onboard processing limits [3]. Many studies on unmanned aerial vehicles (UAVs) have shown substantial progress toward systems that operate without human intervention. Consistent with human- and society-centered notions of autonomy, technical systems claiming autonomy must be able to make decisions and respond to events independently of direct human control [4]. Our system targets GPS-denied environments (e.g., tunnels, basements) and avoids external infrastructure such as motion capture or GPS by leveraging image-based detection and tracking to both perform target tracking and explore vision-driven control of UAVs [5]. By employing detection and tracking techniques, drones can capture target imagery via onboard cameras and conduct real-time analysis to accurately determine the target's position and orientation [6]. Deep learning has driven significant progress in detection methods over the past few years. R-CNN [7] applies high-capacity CNNs to bottom-up region proposals for object localization and segmentation. SPP-net [8] introduces spatial pyramid pooling to remove the fixed-size input constraint. Building on both, Faster R-CNN [9] unifies proposal generation and detection, yielding faster training/inference and improved accuracy. Many detectors most notably SSD [10] tend to underperform at small-object detection. The YOLO family (YOLO [11], YOLOv2 [12], YOLOv3 [13], and later variants such as YOLOv4 [14]–YOLOv8 [15]) is a prominent line of one-stage detectors. YOLOv8 achieves state-of-the-art accuracy and speed, building on prior YOLO advances to suit a wide range of object-detection tasks. For our drone, we adopt a streamlined YOLOv8 variant to cut computational load, enabling deployment on severely resource-constrained onboard hardware. In recent years, numerous studies have advanced quadrotor guidance and tracking. Still, designing a robust flight system remains difficult due to the platform's strongly non-linear, coupled dynamics. To improve performance on resource-constrained airframes, researchers have introduced efficient, real-time tracking techniques specifically tailored to the limited onboard computational capacity of UAVs [16,17]. Bertinetto et al. [18] proposed an end-to-end trainable, fully convolutional Siamese network for visual tracking. In [19], a quadcopter vision system was developed to track a ground-based moving target, with switching controllers used to maintain flight stability. Li et al. [20] introduced a tracker built on a Siamese Region Proposal Network, trained offline on large-scale image pairs. A ResNet-based Siamese tracker is presented in [21]. We achieve real-time DATMO [22] on a quadrotor with a single camera by estimating inter-frame UAV motion to synthesize an artificial optical flow, contrasting it with the real optical flow to isolate and cluster dynamic pixels into moving objects, then tracking them over time to suppress noise while acknowledging limitations from dependence on accurate motion estimation and good texture/lighting, sensitivity to parallax, vibration/rolling-shutter, fast maneuvers and occlusion, and tight onboard computer. DB-Tracker is a detection-based multi-object tracker for drone videos that combines RFS-based position modeling (Box-MeM-Ber) with hierarchical OSNet appearance features and a joint position–appearance cost matrix, achieving robust results in complex, occluded scenes [23].

In [6], the authors combine a YOLOv4 detector, a SiamMask tracker, and PID attitude control to deliver efficient, real-time detection and stable target tracking on resource-constrained drones indoors and outdoors, experimentally validating systematic model pruning for practical control.

This work uses a UAV with a single monocular RGB camera to develop an accurate, precise autonomous braking system for indoor environments. The approach pairs a YOLOv4 object detector with a Kalman filter (KF)-based tracker to enable real-time object tracking. This study focuses on detecting a ball as the target object. Our system comprises three core modules (1) object detection, (2) target tracking, and (3) Proportional-Integral-Derivative (PID) control and we implement and evaluate it on a DJI Tello drone for end-to-end detection and tracking.

The structure of the remainder of this paper is as follows: Section 2 presents the materials and approach used in this study. An explanation of the control strategy is available in Section 3. Then, Section 4 we explain the experiments results. Section 5 provides the final conclusions.

2. Materials and Methods

The UAV used in this study is a DJI Tello, shown in Figure 1a (<https://www.ryzero-botics.com/tello>). Figure 1b depicts the ball used as the target in our system. The ball moves along a predefined trajectory, and the UAV's task is to autonomously detect and continuously track it.

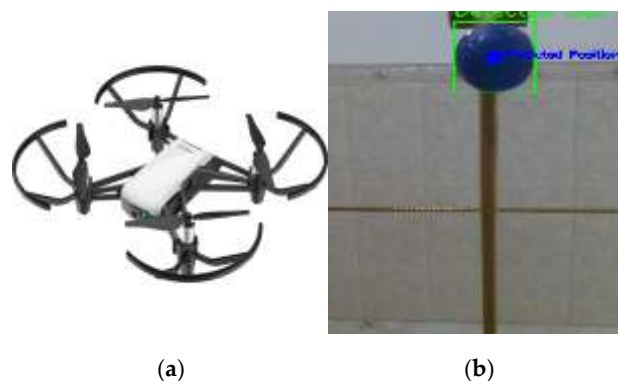


Figure 1. UAV and Ball object. (a) UAV with camera; (b) spherical object (Ball).

The scenario involves a UAV and a single target (a ball) captured by the onboard camera and followed along a predefined path. And A laptop computer is connected to the Tello drone via Wi-Fi for communication. For this study, the onboard camera is used to detect and track the target (ball) in real time. The sensor captures 5-MP still images and streams 720p video; with electronic image stabilization (EIS), the captured images and streamed video remain stable even under airframe vibrations [24].

2.1. Target Detection

UAV To track the target with a UAV, we first use YOLOv8 to detect it and produce a 2D image-plane bounding box, which then serves as the measurement for state estimation. The ball in aerial imagery is detected using YOLOv8, whose architecture is shown in Figure 2 [25]. Before inference, all input images regardless of original size are resized to 640×480 . We selected YOLOv8 as the architecture most likely to succeed for this task, as it is widely regarded as state-of-the-art offering higher (mAP) and faster inference (lower latency) on the COCO dataset [26].

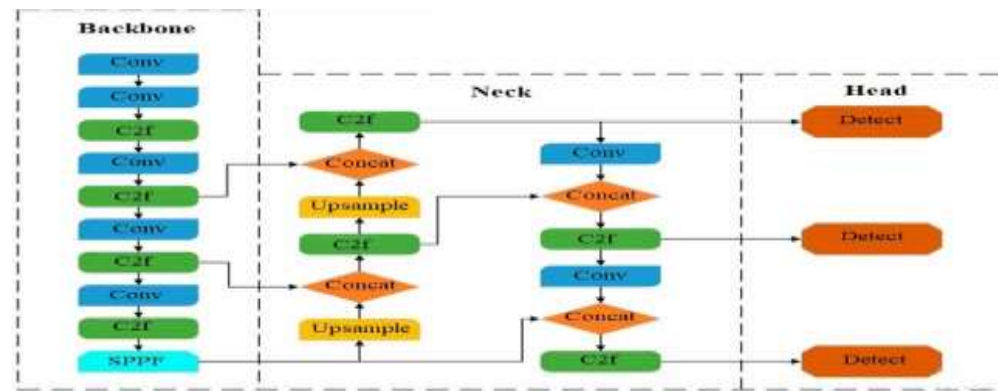


Figure 2. YOLOv8 architecture.

The performance and speed of various YOLO versions are compared in Table 1. The data illustrates a clear trade-off between speed and accuracy, with YOLOv8n being one of the fastest real-time detectors available, delivering impressive accuracy for its size. In contrast, the larger YOLOv8x model achieves 16 mAP higher accuracy, demonstrating a powerful option for tasks where precision is paramount, without sacrificing real-time capability.

Table 1. A Comparative Analysis of YOLO Model Performance.

Model	mAP	FPS	Dataset
YOLOv1	63.4	45	VOC
YOLOv1-Tiny	57.1	155	VOC
YOLOv2	78.6	40	VOC
YOLOv2-Tiny	57.1	244	VOC
YOLOv3	55.3	20	COCO
YOLOv3-Tiny	33.1	220	COCO
YOLOv4	65.7	60	COCO
YOLOv4-Tiny	40.2	400	COCO
YOLOv8n	37.3	500	COCO
YOLOv8x	53.9	100	COCO

2.2. Target Tracking

The algorithm uses a single onboard camera to estimate the target's position in the environment and track it along its trajectory. Using a Kalman filter, we estimate the ball's image-plane center ($\hat{x}, \hat{y}$); Figure 3 and guide the UAV for precise tracking; the filter also predicts the target's next position, improving robustness during brief detection dropouts.

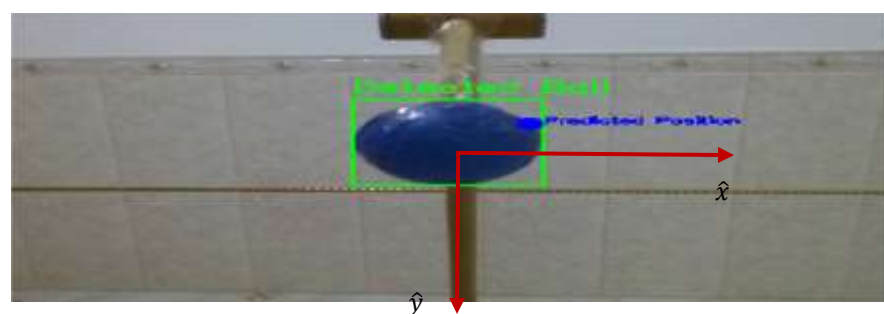


Figure 3. Ball Definition in the Image Coordinate.

3. Control Method

This section describes the Proportional–Integral–Derivative (PID) controller used to steer the drone, enabling smooth and safe tracking of the ball. Leveraging state estimates from the Kalman filter, the UAV autonomously adjusts its motion to follow the ball's trajectory. Figure 4 provides an overview of the control architecture.

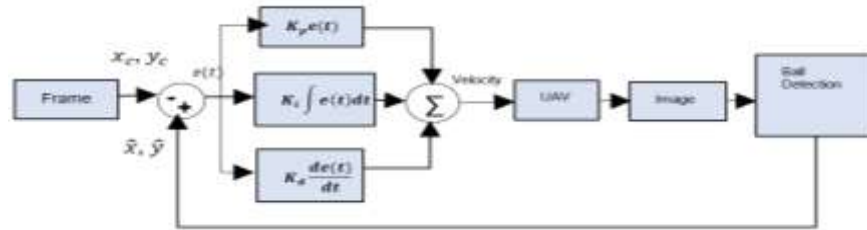


Figure 4. PID Control System for UAV.

From the camera stream, the Kalman filter provides the Ball's estimated image-plane center $(\hat{x}, \hat{y})$; the tracking errors relative to the image center (x_c, y_c) are

$$e_x = \hat{x} - x_c \quad (1)$$

$$e_y = \hat{y} - y_c \quad (2)$$

which feed the PID law

$$u(t) = K_p \cdot e(t) + K_i \int_0^t e(t)dt + K_d \cdot \frac{de(t)}{dt} \quad (3)$$

where K_p , K_i and K_d are the proportional, integral, and derivative coefficients of the PID controller, respectively.

4. Experiment Results

This section describes a vision-based guidance method integrated with a PID controller for visual detection and tracking. We evaluate the system on a DJI Tello drone connected to a laptop with a single NVIDIA GeForce GTX 1650 GPU to assess robustness and reliability; the experiments demonstrate stable, successful tracking.

After detection, the UAV aligns to keep the ball at the image center and flies toward it. Figure 5 illustrates this process: in Figure 5a, the ball appears in the upper-left of the camera view; in Figure 5b, after tracking, the ball is centered in the frame.

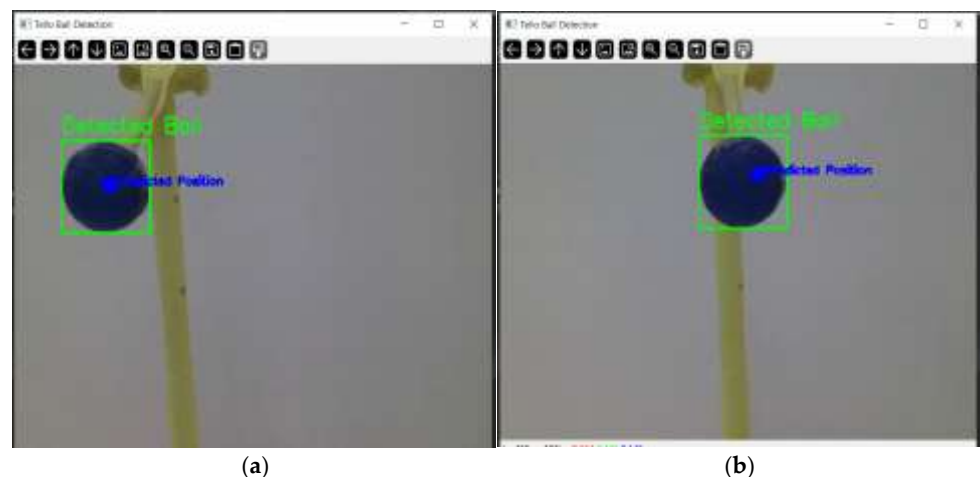


Figure 5. Autonomous ball detection and tracking.

Figure 6 evaluates vertical tracking by moving the ball up and down, showing that the drone can follow these changes. In Figure 6a the ball appears low; after it shifts upward (Figure 6b), the PID controller drives the drone up to track and recenter the target. Figure 7 shows the response to each maneuver. When a change occurs, the tracking error briefly increases, then the controller quickly reduces it and returns the system to a stable state. These results indicate that the designed control system corrects disturbances effectively.

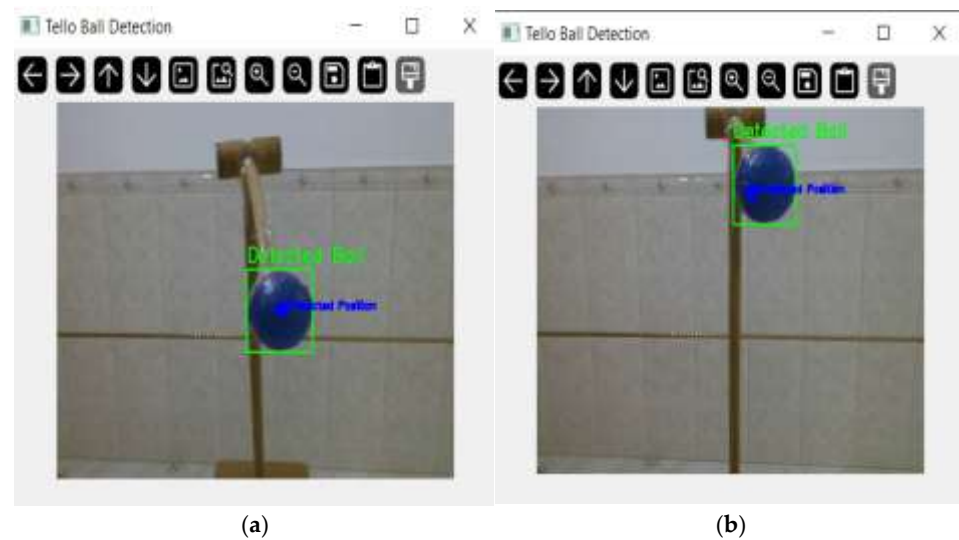
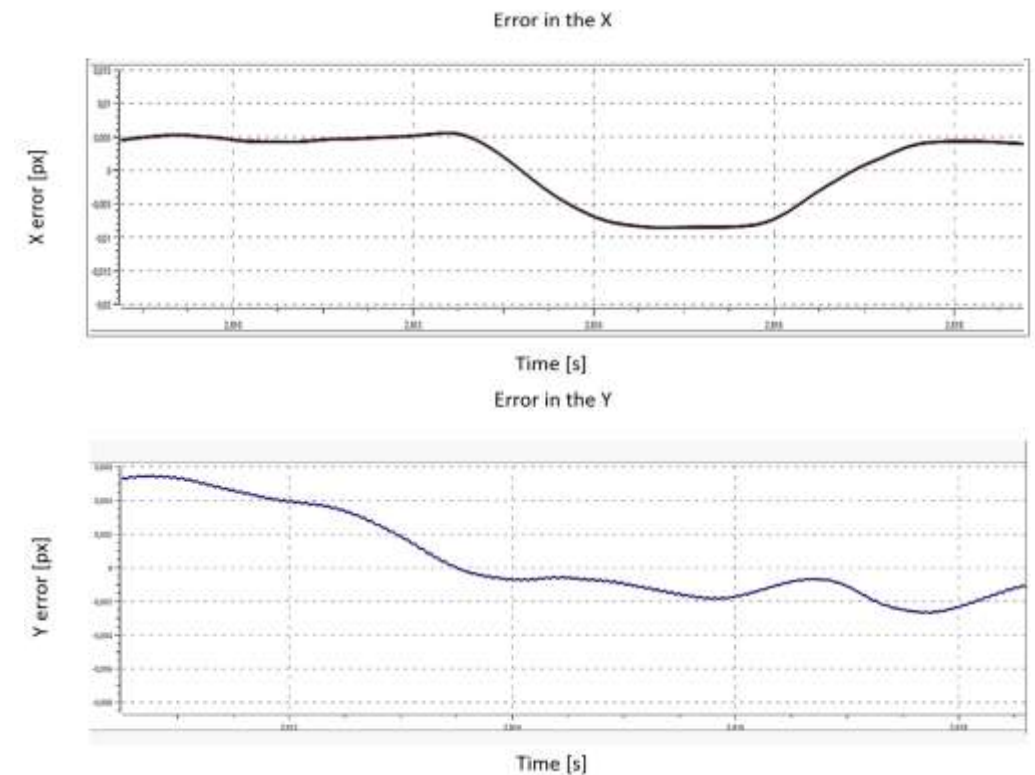


Figure 6. Ball Tracking Movement of UAV in Up/Down.



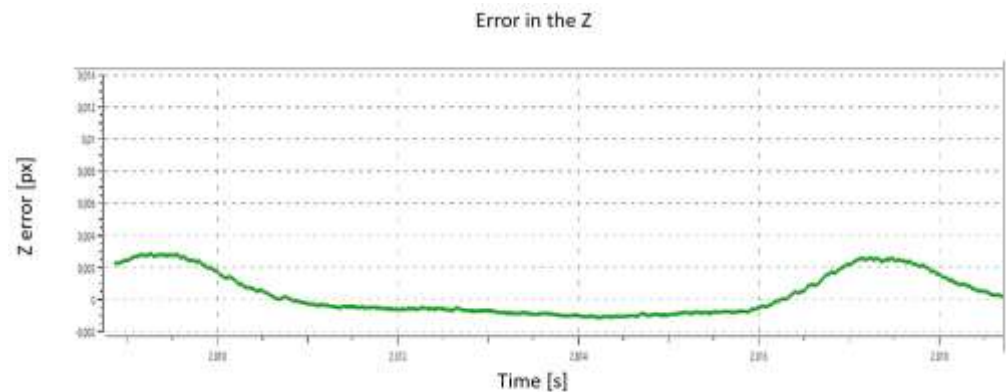


Figure 7. Error in X, Y, and Z over time.

Our method is designed to balance maximum accuracy with real-time speed. We chose the large and powerful YOLOv8x model for the computationally intensive task of initial object detection, ensuring the blue ball is located with high precision. For the subsequent frame-to-frame tracking, we selected a Kalman filter, which is ideal for speed-critical tasks. This hybrid system, when attached to a DJI Tello's camera, creates an effective and engaging tracker that demonstrates high-performance capabilities on a low-cost drone.

5. Conclusions

The paper details the design and implementation of a UAV-mounted, vision-based system for autonomous air-target tracking. We use YOLOv8 to detect the moving ball in real time, even when it occupies only a few pixels. Once the target ball is identified, a Kalman filter-based tracking algorithm is used for visual tracking. We further incorporate a PID module to compute tracking errors and adjust the drone's attitude. PID is a classic feedback controller that derives commands from the proportional (current error), integral (accumulated error), and derivative (rate of error change) terms to drive the system output toward the desired setpoint. The PID module uses the ball's location relative to the image center and the drone's state to compute errors and update attitude control for steady target tracking. Overall, flight tests validate the system's effectiveness and robustness: YOLOv8 delivers fast, real-time detection, the Kalman filter maintains accurate ball tracking, and the PID controller computes attitude corrections that keep the UAV stably locked on the target.

Author Contributions:

Funding:

Institutional Review Board Statement:

Informed Consent Statement:

Data Availability Statement:

Conflicts of Interest:

References

1. Moali, O.; Mezghani, D.; Mami, A.; Oussar, A.; Nemra, A. UAV trajectory tracking using proportional-integral-derivative-type-2 fuzzy logic controller with genetic algorithm parameter tuning. *Sensors* **2024**, *24*, 6678.
2. Kang, H.C.; Han, H.N.; Bae, H.C.; Kim, M.G.; Son, J.Y.; Kim, Y.K. HSV color-space-based automated object localization for robot grasping without prior knowledge. *Appl. Sci.* **2021**, *11*, 7593.

- Hirano, S.; Uchiyama, K.; Masuda, K. Controller Design Using Backstepping Algorithm for Fixed-Wing UAV with Thrust Vectoring System. *Adv. Sci. Technol. Eng. Syst. J.* **2020**, *5*, 284–290.
3. Cheng, H.; Lin, L.; Zheng, Z.; Guan, Y.; Liu, Z. An autonomous vision-based target tracking system for rotorcraft unmanned aerial vehicles. In Proceedings of the 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Vancouver, BC, Canada, 24–28 September 2017; pp. 1732–1738.
4. Shakhathreh, H.; Sawalmeh, A.H.; Al-Fuqaha, A.; Dou, Z.; Almaita, E.; Khalil, I.; Othman, N.S.; Khreishah, A.; Guizani, M. Unmanned aerial vehicles (UAVs): A survey on civil applications and key research challenges. *IEEE Access* **2019**, *7*, 48572–48634.
5. Gharsa, O.; Touba, M.M.; Boumehraz, M.; Abderrahman, N.; Bellili, S.; Titaouine, A. Autonomous landing system for A Quadrotor using a vision-based approach. In Proceedings of the 2024 8th International Conference on Image and Signal Processing and Their Applications (ISPA), Biskra, Algeria, 21–22 April 2024; pp. 1–5.
6. Yang, S.Y.; Cheng, H.Y.; Yu, C.C. Real-time object detection and tracking for unmanned aerial vehicles based on convolutional neural networks. *Electronics* **2023**, *12*, 4928.
7. Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 580–587.
8. He, K.; Zhang, X.; Ren, S.; Sun, J. Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2015**, *37*, 1904–1916.
9. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards real-time object detection with region proposal networks. *Adv. Neural Inf. Process. Syst.* **2015**, *28*.
10. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.Y.; Berg, A.C. Ssd: Single shot multibox detector. In *European Conference on Computer Vision*; Springer International Publishing: Cham, Switzerland, 2016; pp. 21–37.
11. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788.
12. Redmon, J.; Farhadi, A. YOLO9000: Better, faster, stronger. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 7263–7271.
13. Redmon, J.; Farhadi, A. Yolov3: An incremental improvement. *arXiv* **2018**, arXiv:1804.02767.
14. Bochkovskiy, A.; Wang, C.Y.; Liao, H.Y.M. Yolov4: Optimal speed and accuracy of object detection. *arXiv* **2020**, arXiv:2004.10934.
15. Varghese, R.; Sambath, M. Yolov8: A novel object detection algorithm with enhanced performance and robustness. In Proceedings of the 2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS), Chennai, India, 18–19 April 2024; pp. 1–6.
16. Saavedra-Ruiz, M.; Pinto-Vargas, A.M.; Romero-Cano, V. Monocular visual autonomous landing system for quadcopter drones using software in the loop. *IEEE Aerosp. Electron. Syst. Mag.* **2021**, *37*, 2–16.
17. Lee, D.; Park, W.; Nam, W. Autonomous landing of micro unmanned aerial vehicles with landing-assistive platform and robust spherical object detection. *Appl. Sci.* **2021**, *11*, 8555.
18. Bertinetto, L.; Valmadre, J.; Henriques, J.F.; Vedaldi, A.; Torr, P.H. Fully-convolutional siamese networks for object tracking. In *European Conference on Computer Vision*; Springer International Publishing: Cham, Switzerland, 2016; pp. 850–865.
19. Gomez-Balderas, J.E.; Flores, G.; García Carrillo, L.R.; Lozano, R. Tracking a ground moving target with a quadrotor using switching control: Nonlinear modeling and control. *J. Intell. Robot. Syst.* **2013**, *70*, 65–78.
20. Li, B.; Yan, J.; Wu, W.; Zhu, Z.; Hu, X. High performance visual tracking with siamese region proposal network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 8971–8980.
21. Li, B.; Wu, W.; Wang, Q.; Zhang, F.; Xing, J.; Yan, J. Siamrpn++: Evolution of siamese visual tracking with very deep networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 4282–4291.
22. Rodríguez-Canosa, G.R.; Thomas, S.; Del Cerro, J.; Barrientos, A.; MacDonald, B. A real-time method to detect and track moving objects (DATMO) from unmanned aerial vehicles (UAVs) using a single camera. *Remote Sens.* **2012**, *4*, 1090–1111.
23. Yuan, Y.; Wu, Y.; Zhao, L.; Chen, J.; Zhao, Q. DB-Tracker: Multi-Object Tracking for Drone Aerial Video Based on Box-MeMBer and MB-OSNet. *Drones* **2023**, *7*, 607.
24. Iskandar, M.; Bingi, K.; Prusty, B.R.; Omar, M.; Ibrahim, R. Artificial intelligence-based human gesture tracking control techniques of Tello EDU Quadrotor Drone. In *IET Conference Proceedings CP837*; The Institution of Engineering and Technology: Stevenage, UK, 2023; Volume 2023, pp. 123–128.

25. Elhanashi, A.; Dini, P.; Saponara, S.; Zheng, Q. TeleStroke: Real-time stroke detection with federated learning and YOLOv8 on edge devices. *J. Real-Time Image Process.* **2024**, *21*, 121.
26. Reis, D.; Kupec, J.; Hong, J.; Daoudi, A. Real-time flying object detection with YOLOv8. *arXiv* **2023**, arXiv:2305.09972.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.