

Analysis of object identification methods for FPV drones

Denys Chebanchyk¹ Vladyslav Yevsiev²

1. CITAR Department, Kharkiv National University of Radio Electronics, UKRAINE, Kharkiv, Nauki Ave. 14., e-mail: denys.chebanchyk@nure.ua
2. CITAR Department, Kharkiv National University of Radio Electronics, UKRAINE, Kharkiv, Nauki Ave. 14., e-mail: vladyslav.yevsieiv@nure.ua

Abstract: The abstracts of the report consider modern methods of object identification for FPV drones with an emphasis on their application in real time and under conditions of limited computing resources. Classical approaches based on keypoint extraction, deep convolutional neural networks, semantic and instance-segmentation methods, as well as state filters and lightweight optimized models are analyzed. The study shows that each of the methods has its advantages and limitations depending on the accuracy, processing speed and complexity of the environment. Special attention is paid to hybrid approaches that combine the advantages of several methods to ensure stable and effective object identification on board FPV drones. The results obtained emphasize the need to optimize algorithms and adapt models to the resource constraints of drones to ensure reliability and accuracy of operation in dynamic conditions.

Keywords: FPV drones, object identification, computer vision, deep neural networks, segmentation, Kalman filter, lightweight models, hybrid methods, real-time.

I. INTRODUCTION

The modern development of unmanned aerial vehicles, in particular FPV drones, opens up new opportunities for their use in military, civil and industrial spheres, where fast and reliable identification of objects in real time plays a key role. FPV technology provides the operator with the ability to directly visually control the drone, which allows performing complex maneuvers, however, the effectiveness of task performance largely depends on the accuracy of the object recognition system. In conditions of a dynamic environment, changes in lighting, the presence of obstacles or camouflage, target identification becomes a particularly difficult task. This necessitates the use of modern methods of computer vision, machine learning and deep neural networks, which are able to improve the quality of recognition even with incomplete or distorted data. The importance of this direction is determined by the need to increase the autonomy of FPV drones, which will reduce dependence on the operator and increase efficiency in performing reconnaissance, search and rescue or tactical tasks. At the same time, it is relevant to study methods for object identification from the point of view of optimizing computing resources, since FPV drones usually have limited computing power. Thus, analyzing existing approaches and determining their advantages and disadvantages is an important stage for the further development of intelligent decision support systems in the field of FPV technologies.

II. MATHEMATICAL DESCRIPTION OF METHODS FOR CONTROLLING A GROUP OF MOBILE ROBOTS

Feature-based keypoint detection & matching (SIFT / ORB / BRIEF) method. This is a classic approach that extracts local "keypoints" in frames, builds descriptors for them, and compares the descriptors between frames or with a sample database to identify objects; works well with limited resources and changing object appearance. Mathematically, the key steps are: feature detection through local operations on the image (e.g., Difference of Gaussians - DoG) and descriptor construction (gradient histograms in SIFT or binary comparisons in BRIEF/ORB).

- DoG for detection:

$$\begin{aligned} L(x, y, \sigma) &= G(x, y, \sigma) \cdot I(x, y) \\ DoG(x, y, \sigma_k) &= L(x, y, \sigma_k) - L(x, y, \sigma) \end{aligned} \quad (1)$$

Where: G - Gaussian filter with dispersion σ ; I - intensity; σ - scale, detection threshold, number of octaves - ORB/BRIEF descriptor + matching:

$$D(d_a, d_b) = \text{Hamming}(d_a, d_b) \quad (2)$$

Advantages: few computations (especially ORB/BRIEF), invariance to shift/scale/rotation (SIFT), works well with scarce resources and for real-time on edge platforms. Disadvantages: sensitivity to strong lighting changes, texture-poor regions (where there are few keypoints), difficult to scale to semantic identification (class \neq specific keypoint).

Method one-stage detectors based on CNN (YOLO / SSD). These models implement object detection and classification in a single pass of the network, estimating for each anchor rectangle the coordinates and class probabilities; suitable for fast processing of FPV video with sufficient hardware support (GPU / NN-accelerator). Mathematically — convolutions + box regression + loss functions for coordinates and probabilities (combination of local regression and cross-entropy). Model (basic architecture):

$$\begin{aligned} y &= f_{\theta}(I) \\ L &= \lambda_{coord} \sum (x - \hat{x})^2 + \lambda_{coord} \sum (y - \hat{y})^2 \\ &+ \sum CE(p_{class}, \hat{p}_{class}) + \dots \end{aligned} \quad (3)$$

Where: f_{θ} - convolutional network.

Advantages: high speed (especially YOLO-lite versions), ready-made end-to-end solutions, good balance of accuracy and speed. Disadvantages: for FPV drones, requires optimization/hardware acceleration; accuracy drops for very

small or very blurred objects; demanding on the amount of training data.

Method semantic segmentation / instance segmentation (U-Net, DeepLab, Mask R-CNN). The method returns object masks (pixel identification) and is useful when you need to accurately separate the shape of an object in an FPV frame (for example, people, cars, useful details). The mathematics is based on convolutional network transformations with upsampling (decoder) and a loss function per pixel (cross-entropy or IoU-based loss).

- model 1 (U-Net):

$$y = \text{Decoder}(\text{Encoder}(I)) \quad (4)$$

where the encoder compresses the image properties, the decoder restores the mask size; parameters: number of levels, filter size, activation function.

- model 2 (IoU-loss for mask):

$$L_{IoU} = 1 - \frac{\sum p_i g_i}{\sum p_i + \sum g_i - \sum p_i g_i} \quad (5)$$

Where: p_i - predicted pixels; g_i - true.

Advantages: accurate localization and shape of the object, useful for navigating around static/dynamic obstacles. Disadvantages: high computational and memory requirements; slower than box detectors; more difficult in real-time on limited FPV drone hardware.

Tracking + state estimation method (Kalman Filter, Particle Filter + association - SORT/DeepSORT). This approach combines frame-by-frame detection with tracking of object trajectories over time, which allows maintaining object identity during short-term fading or partial overlaps. The mathematics consists of state filter equations and prediction/correction steps and an association algorithm between detections and tracks (e.g., guessing/Hungarian algorithm). Model (Kalman filter, linear):

- state:

$$x_{t+1} = Fx_t + Bu_t + w_t \quad (6)$$

- observation:

$$z_t = Hx_t + v_t \quad (7)$$

Where: $w_t \sim \mathcal{N}(0, Q)$; $v_t \sim \mathcal{N}(0, R)$; H - observation matrix; Q, R - noise covariance. Particle filter is suitable for nonlinear/non-Gaussian cases where the state is approximated by a set of particles $\{x^{(i)}, w^{(i)}\}$. Association: minimization of the cost matrix (IOU or distance) via Hungarian; parameters: IOU threshold, maximum track loss. Advantages: stability of tracks over time, possibility of predicting object motion and recovery after losses. Disadvantages: dependence on detector quality, association errors with strong crowding of objects; Particle filter can be computationally expensive.

Method Lightweight on-device models and optimizations (MobileNet, EfficientNet-Lite + quantization/pruning/distillation). For FPV drones, it is critical to reduce computation and memory, so specially optimized architectures and model compression techniques are used. Mathematically important idea is depthwise

separable convolutions (MobileNet), which decompose the full convolution into depthwise and pointwise stages, significantly reducing the number of operations. Model (depthwise separable conv): the full convolution has a complexity:

$$D_k^2 \cdot M \cdot N \cdot H \cdot W \quad (8)$$

Where: D_k - kernel size; M - input channels; N - output; parameters: network width (width multiplier), input resolution, quantization degree (8-bit, 16-bit). Quantization is also used: $w_q = \text{round}(\frac{w}{s})$ with scale s , and pruning (removal of small weights). Advantages: the ability to perform inference directly on board a drone with limited energy consumption, low latency, maintaining autonomy. Disadvantages: tradeoff of accuracy against size/speed; some optimizations (aggressive pruning, low-bit quant) can spoil stability for delicate classes.

General conclusion: all the described approaches have their strengths and limitations - classical feature-based methods are light and cheap in resources, but weak in semantic classification; CNN detectors provide powerful semantics and speed, but require optimization for FPV drones; segmentation is useful for accurate localization of shapes, but is more resource-intensive; the combined approach "detection + tracking/state filtering" significantly increases the robustness of identification in the video stream; and the use of lightweight architectures and compression techniques is a mandatory practical step for implementing an identification system directly on a drone. For FPV scenarios, hybrid solutions are most effective: for example, a lightweight CNN detector on board + segmentation/more accurate validation on the backend or a hardware accelerator; or a combination of feature-matching for fast matches and CNN for class confirmation. If you want, I can: a) offer examples of architectures (MobileNet-Tiny + SSD configurations) with approximate calculation metrics; b) prepare a flowchart of on-board / off-board processing integration for an FPV drone - immediately in technical form.

III. COMPARATIVE ANALYSIS OF METHODS FOR CONTROLLING A GROUP OF MOBILE ROBOTS

A comparative analysis of the considered methods of object identification for FPV drones shows that each of them has its own implementation specifics, which are determined by both mathematical models and hardware limitations of the platforms themselves. Classical computer vision methods based on feature extraction, in particular SIFT or ORB, are characterized by relative simplicity of implementation and low requirements for computing resources, which makes them suitable for FPV drones with limited on-board processor performance. At the same time, these methods are vulnerable to changes in lighting, noise and partial overlaps, which significantly reduces their reliability in dynamic and unpredictable conditions. Deep convolutional neural networks, on the contrary, provide high accuracy of object recognition even in complex environments, but require significant computing power and optimization, in particular through the use of accelerators such as GPU or TPU. For FPV drones, where energy efficiency and processing speed are critical, CNN models require special optimizations, for

example, by using MobileNet or Tiny-YOLO. Bayesian methods, in particular variations of the Kalman filter or particle filters, provide good performance in conditions of noisy sensor data and partial visibility of objects, allowing to integrate information from multiple sources. However, their effectiveness depends on the correct definition of a priori models and probability distributions, which is sometimes a difficult task in real flight conditions. Methods based on fuzzy logic demonstrate flexibility in cases where the data is imprecise or contradictory, but their accuracy and speed may be inferior to other current approaches. Bio-inspired algorithms, such as swarm models or evolutionary approaches, allow FPV drones to distribute the load and collectively identify objects, but when used alone they do not provide high accuracy without integration with more traditional methods. In general, it can be noted that for FPV drones, the most effective are hybrid systems that combine the simplicity of classical methods for pre-extraction of features with the high accuracy of deep models, as well as support the results with probabilistic and fuzzy methods. This allows you to balance computational costs, energy efficiency and recognition quality, which is critical for performing tasks in real time.

IV. CONCLUSION

The conducted study of object identification methods for FPV drones showed that the efficiency of such systems largely depends on the choice of algorithmic approach and hardware platform. Classical methods based on key points, such as SIFT and ORB, provide relatively low computational complexity and fast operation, but their accuracy is significantly reduced in conditions of noise, lighting changes or partial overlapping of objects. Deep convolutional neural networks allow achieving high recognition accuracy even in complex dynamic scenes, but require optimization for use on board FPV drones due to limited computing resources. Semantic and instance-segmentation methods provide accurate selection of the object shape and are useful for navigation, but are more resource-intensive and slower in real time. The use of state filters, such as Kalman or particle filters, together with tracking allows maintaining the identity of the object over time and predicting its movement, but requires a correct a priori model and parameter settings. Lightweight optimized models such as MobileNet and EfficientNet-Lite with quantization and pruning techniques demonstrate a balance between speed and accuracy, making them suitable for on-board processing on FPV drones. The overall analysis shows that the most effective are hybrid approaches that combine classical feature extraction methods with deep models and probabilistic filters to increase the robustness and accuracy of identification. This approach allows for optimal allocation of computing resources, ensuring fast real-time response and system reliability in dynamic environments. The implementation of hybrid methods is promising for the development of autonomous FPV drones that perform reconnaissance, monitoring and navigation tasks in complex conditions. The results obtained emphasize the need for further improvement of algorithms, integration of optimized models and development of adaptive systems capable of operating in various environments with a high level of uncertainty.

- [1] Zia, A., & Haleem, M. (2025). Bridging research gaps in industry 5.0: Synergizing federated learning, collaborative robotics, and autonomous systems for enhanced operational efficiency and sustainability. *IEEE Access*.
- [2] Dobrzańska, M., & Dobrzański, P. (2025). Simulation Model as an Element of Sustainable Autonomous Mobile Robot Fleet Management. *Energies*, 18(8), 1894.
- [3] Chen, Z., Wang, W., Liu, W., Liu, Y., & Xi, J. (2025). The Effects of Communication Delay on Human Performance and Neurocognitive Responses in Mobile Robot Teleoperation. *arXiv preprint arXiv:2508.18074*.
- [4] Gupta, S., Mamodiya, U., Hossain, A. Z., & Al-Gburi, A. J. (2025). A Framework for Bluetooth-Based Real-Time Audio Data Acquisition in Mobile Robotics. *Signals*, 6(3), 31.
- [5] Delgado Asencio, C., Segura Manzano, F., & Andújar Márquez, J. M. (2025). Hygrobot: Hydrogen-Powered Mobile Robotic Platform for Rough Terrain: Electro-Mechanical Design, Physical Implementation, and Experimental Testing. *Applied Sciences*, 15(9), 5028.
- [6] Pastrana Triana, M. A., Santana, M. S., Mendoza Peñaloza, J. A., Nunes de Oliveira, L. H., & Muñoz Arboleda, D. M. (2025). Comparison of gmdh and perceptron controllers for mobile robot obstacle following/avoidance with hardware-in-the-loop validation. *Journal of Intelligent & Robotic Systems*, 111(1), 1-23.
- [7] Liu, F., & Tahir, R. (2025, May). Optimising FPV Drone Cinematography Experience: Transforming Flight Planning Workflow by Integrating Ground Control Software. In *International Conference on Human-Computer Interaction* (pp. 85-110). Cham: Springer Nature Switzerland.
- [8] Khudov, H., Khudov, R., Khizhnyak, I., Makoveichuk, O., & Khudov, V. (2025). Image Segmentation Methods for Kamikaze FPV Drones Targeting to Aid Critical Energy National Infrastructure Assets Protection. In *Systems, Decision and Control in Energy VII: Volume I: Energy Informatics and Transport* (pp. 139-151). Cham: Springer Nature Switzerland.
- [9] Halytskyi, O., Denysiuk, D., Kozhemiako, Y., & Kvassay, M. (2025). Method of FPV drone stabilization on an automatically determined target and its further observation.
- [10] Xiao, J., Zhang, R., Zhang, Y., & Feroskhan, M. (2025). Vision-based learning for drones: A survey. *IEEE Transactions on Neural Networks and Learning Systems*.
- [11] Joyce, K. J. (2025, March). Drone racing's utility to contemporary operations as FPV kamikaze loitering UAV munitions and the next tech leap. In *AIAC 2025: 21st Australian International Aerospace Congress* (pp. 617-623). Melbourne: Engineers Australia; Royal Aeronautical Society.
- [12] Taylor, B., Allen, M., Henson, P., Gao, X., Malik, H., & Zhu, P. (2025). Enhancing Drone Navigation and Control: Gesture-Based Piloting, Obstacle Avoidance, and 3D Trajectory Mapping. *Applied Sciences*, 15(13), 7340.

- [9] Attar, H., Abu-Jassar, A. T., Yevsieiev, V., Lyashenko, V., Nevliudov, I., & Luhach, A. K. (2022). Zoomorphic mobile robot development for vertical movement based on the geometrical family caterpillar. *Computational intelligence and neuroscience*, 2022(1), 3046116.
- [10] Nevliudov, I., Yevsieiev, V., Baker, J. H., Ahmad, M. A., & Lyashenko, V. (2020). Development of a cyber design modeling declarative Language for cyber physical production systems. *J. Math. Comput. Sci.*, 11(1), 520-542.
- [11] Nevliudov, I., & et al.. (2020). Method of Algorithms for CyberPhysical Production Systems Functioning Synthesis. *International Journal of Emerging Trends in Engineering Research*, 8(10), 7465-7473.
- [12] Lyashenko, V., Abu-Jassar, A. T., Yevsieiev, V., & Maksymova, S. (2023). Automated Monitoring and Visualization System in Production. *International Research Journal of Multidisciplinary Technovation*, 5(6), 9-18.
- [13] Mustafa, S. K., Yevsieiev, V., Nevliudov, I., & Lyashenko, V. (2022). HMI Development Automation with GUI Elements for Object-Oriented Programming Languages Implementation. *SSRG International Journal of Engineering Trends and Technology*, 70(1), 139-145.
- [14] Nevliudov, I., Yevsieiev, V., Lyashenko, V., & Ahmad, M. A. (2021). GUI Elements and Windows Form Formalization Parameters and Events Method to Automate the Process of Additive Cyber-Design CPPS Development. *Advances in Dynamical Systems and Applications*, 16(2), 441-455.
- [15] Kragic, D., Gustafson, J., Karaoguz, H., Jensfelt, P., & Krug, R. (2018, July). Interactive, Collaborative Robots: Challenges and Opportunities. In *IJCAI* (pp. 18-25).
- [16] Nevliudov, I., Yevsieiev, V., Maksymova, S., Gopejenko, V., & Kosenko, V. (2025). Development of mathematical support for adaptive control for the intelligent gripper of the collaborative robot manipulator. *Advanced Information Systems*, 9(3), 57-65.
- [17] Maksymova, S., Yevsieiev, V., Chala, O., & Ababneh, J. (2025). DECISION-MAKING MODEL FOR CONTROLLING A COLLABORATIVE ROBOT-MANIPULATOR BASED ON THE SENSOR FUSION METHOD AND THE RULES OF RULE-BASED SYSTEMS. *Multidisciplinary Journal of Science and Technology*, 5(6), 526-538.
- [18] Невлюдов, І. Ш., Євсєєв, В. В., & Гурін, Д. В. (2025). Model development of dynamic representation a model description parameters for the environment of a collaborative robot manipulator within the industry 5.0 framework. *Системи управління, навігації та зв'язку. Збірник наукових праць*, 1(79), 42-48.
- [19] Yevsieiev, V., Abu-Jassar, A., Maksymova, S., & Demska, N. (2025). Development of a model for recognizing various objects and tools in a collaborative robot workspace. *ACUMEN: International journal of multidisciplinary research*, 2(1), 224-239.
- [20] Yevsieiev V. Mobile Robots and Autonomous Vehicles in the Mobility as a Service (MAAS) Concept / V. Yevsieiev // Sustainable smart cities and communities: business and innovation solutions 2025 : Theses of Reports of I st I International Conference, April 21, 2025. - Kharkiv, 2025. - P.7-8.
- [21] Yevsieiev V. Using Multi-Agent Systems in the Management of Collaborative Robots / V. Yevsieiev // Computer-integrated technologies, automation and robotics 2025 : Theses of Reports of II st All-Ukrainian Conference, May 16-17, 2025. - Kharkiv, 2025. - P. 13-17
- [22] Yevsieiev, V., Maksymova, S., Gurin, D., & Alkhalaileh, A. (2024). HR data visualization of the distance to the object in the collaborative robot workspace based on hc-sr04 sensor. *ACUMEN: International journal of multidisciplinary research*, 1(4), 388-401.
- [23] Yevsieiev, V., Maksymova, S., Abu-Jassar, A., & Ababneh, J. (2025). MATHEMATICAL MODEL OF LOCAL DECISION-MAKING FOR COLLABORATIVE ROBOTS USING EDGE COMPUTING. *Multidisciplinary Journal of Science and Technology*, 5(6), 34-46.
- [24] Yevsieiev, V. Comparative Analysis of the Characteristics of Mobile Robots and Collaboration Robots Within INDUSTRY 5.0. / V. Yevsieiev, D. Gurin // Sectoral research XXI : characteristics and features : collection of scientific papers "SCIENTIA" with proceedings of the VI International Scientific and Theoretical Conference, September 8, 2023. - Chicago : European Scientific Platform, 2023. - P. 92-94.
- [25] Gurin, D., Yevsieiev, V., Maksymova, S., & Abu-Jassar, A. (2024). Effect of Frame Processing Frequency on Object Identification Using MobileNetV2 Neural Network for a Mobile Robot. *Multidisciplinary Journal of Science and Technology*, 4(8), 36-44.
- [26] Gurin, D., Yevsieiev, V., Maksymova, S., & Alkhalaileh, A. (2024). MobileNetv2 Neural Network Model for Human Recognition and Identification in the Working Area of a Collaborative Robot. *Multidisciplinary Journal of Science and Technology*, 4(8), 5-12.
- [27] Yevsieiev, V. Using the Triangulation Method to Measure the Distance to Objects in the Working Area of a Collaborative Manipulator Robot / V. Yevsieiev, S. Starikova // Manufacturing & Mechatronic Systems 2024 : Proceedings of VIII st International Conference, October 25-26, 2024. - Kharkiv, 2024. - P.107-109.
- [28] Yevsieiev V. Simulation of the operation of the sensor system of a mobile robot in the Autodesk tinkercad environment / V. Yevsieiev, S. Starikova // Комп'ютерні ігри і мультимедіа як інноваційний підхід до комунікації-2023 : матеріали III Всеукр. наук.-техн. конф. молодих вчених, аспірантів і студентів, Одеса, 28-29 жовтня 2023 р. - Одеса : ОНТУ, 2023 . - С. 21-23.
- [29] Gurin, D., Yevsieiev, V., Maksymova, S., & Abu-Jassar, A. (2024). Effect of Frame Processing Frequency on Object Identification Using MobileNetV2 Neural Network for a Mobile Robot. *Multidisciplinary Journal of Science and Technology*, 4(8), 36-44.
- [30] Nevliudov, I., Yevsieiev, V., Maksymova, S., Demska, N., Kolesnyk, K., & Miliutina, O. (2023, September). Mobile Robot Navigation System Based on Ultrasonic Sensors. In *2023 IEEE XXVIII International Seminar/Workshop on Direct and Inverse Problems of Electromagnetic and Acoustic Wave Theory (DIPED)* (Vol. 1, pp. 247-251). IEEE.