УДК 632.45

RESEARCH OF METHODS FOR PROCESSING ANALOG VIDEO SIGNALS OF FPV DRONES FOR IDENTIFICATION SYSTEMS BASED ON YOLOV9

Kyrylo Luchaninov
Kharkiv Kharkiv National University of Radio Electronics
Ukraine, 61166, Kharkiv, Nauky av, 14
E-mail: kyrylo.luchaninov@nure.ua

Abstract: The article addresses the problem of reduced performance of identification systems based on YOLOv9 when using analog video signals from FPV drones operating in the 5.8 GHz band. An analysis of the impact of noise, transmission artifacts, and digitization processes on input data quality and object detection accuracy is conducted. An approach to improving the informativeness of the video stream is proposed through the application of signal preprocessing methods and adaptation of the neural network to analog channel conditions. The obtained results confirm the feasibility of using a comprehensive preprocessing pipeline to enhance the reliability of computer vision systems for FPV drones.

Keywords: FPV drone, analog video signal, YOLOv9, computer vision, signal processing, object detection, preprocessing.

ДОСЛІДЖЕННЯ МЕТОДІВ ОБРОБКИ АНАЛОГОВОГО ВІДЕОСИГНАЛУ FPV ДРОНІВ ДЛЯ СИСТЕМ ІНДІФІКАЦІЇ НА БАЗІ YOLO v.9

Кирило Лучанінов
Харківський національний університет радіоелектроніки
Україна, 61166, Харків, пр. Науки 14
E-mail: kyrylo.luchaninov@nure.ua

Анотація: У статті розглянуто проблему зниження ефективності систем ідентифікації на базі YOLOv9 при використанні аналогового відеосигналу FPV-дронів у діапазоні 5.8 ГГц. Проведено аналіз впливу шумів, артефактів передачі та процесів оцифрування на якість вхідних даних і точність детекції об'єктів. Запропоновано підхід до підвищення інформативності відеопотоку шляхом застосування методів попередньої обробки сигналу та адаптації нейронної мережі до умов аналогового каналу. Отримані результати підтверджують доцільність використання комплексного preprocessing pipeline для підвищення надійності систем комп'ютерного зору FPV-дронів.

Ключові слова: FPV-дрон, аналоговий відеосигнал, YOLOv9, комп'ютерний зір, обробка сигналів, детекція об'єктів, попередня обробка.

The relevance of studying methods for processing analog video signals at a frequency of 5.8 GHz is due to the rapid development of FPV drones, which are widely used in monitoring, reconnaissance, search and autonomous navigation tasks in conditions of limited resources and complex environments. Despite the active transition to digital data transmission systems, analog FPV cameras remain dominant due to low signal latency, energy efficiency and ease of implementation, which is critically important for real-time tasks. At the same time, the use of an analog channel leads to significant video distortions, which significantly reduce the efficiency of modern computer vision systems, in particular those based on neural networks such as YOLO v9. This creates a contradiction between the hardware simplicity of FPV systems and the requirements for highly accurate object identification, which requires the development of specialized signal preprocessing methods. Converting an analog video stream into a high-quality digital format with minimal information loss is a key step in ensuring the stable operation of detection algorithms in conditions of noise, interference and dynamic changes in the environment. This is of particular importance in autonomous control tasks for FPV drones, where recognition errors can lead to critical consequences. Research in this area allows us to increase the reliability of computer vision systems without the need to completely replace the existing analog infrastructure. In addition, the integration of effective methods for filtering, restoring and stabilizing video opens up opportunities for implementing intelligent functions even on limited computing platforms. This corresponds to current trends in the development of edge AI and autonomous robotic systems. Thus, the study of analog video signal processing methods is an important scientific and practical task aimed at increasing the efficiency of intelligent FPV drones in real operating conditions. YOLOv9-based identification systems do not work well with images from an analog 5.8 GHz camera not because YOLOv9 is "weak", but because the input of the neural network is already a partially corrupted video signal.

YOLOv9, like most modern detectors, is designed for fairly stable digital RGB frames with clear contours, textures, colors and scale of the object; the YOLOv9 architecture itself is focused on preserving useful information inside the neural network through PGI and GELAN, but it cannot restore information that was lost before the signal was digitized.

The main problem with the 5.8 GHz analog video channel is that the image is not transmitted as a digital stream of pixels, but as an analog signal, which is sensitive to noise, multipath, signal level drop, interference, synchronization instability and receiver quality. As a result, "snow", stripes, flicker, line breaks, color shift, contrast drop, local geometric distortions and temporary image drops appear in the frame. For a person, such a frame may still be understandable, but for YOLO this means the loss of features by which the model recognizes an object: contours, angles, textures, borders, color transitions and stable shape. The second important reason is the low effective resolution of the analog camera. Even if the receiver or USB capture device outputs a frame, for example 720×576 or 640×480, the actual detail is often lower due to PAL/NTSC limitations, sensor quality, analog modulation, radio channel and noise. In digital video processing, it is explicitly stated that video digitized from analog sources is limited by the resolution and artifacts of the analog signal itself. This is critical for YOLOv9, because a small object, after scaling to the input size of the network, can occupy a very small number of useful pixels, and its shape becomes fuzzy. The third problem is interlacing and deinterlacing. Many analog video signals are formed as interlaced video, where one frame consists of two fields captured at slightly different times. If the object or camera moves, "combs", contour bifurcations, and microshifts appear after digitization. This is especially harmful for YOLO, because the boundaries of the object become not a physical form, but an artifact of the transformation. In computer vision tasks, such artifacts can be perceived as false signs or, conversely, hide the real object. The fourth reason is the instability of frames in time. YOLOv9 usually analyzes a single frame or a sequence of frames, but expects the image to have a relatively stable structure. In a 5.8 GHz analog channel, when the antenna position changes, reflections from walls, metal objects, or platform movement, the quality can change every second. One frame may be clear, the next one may be noisy, the third one may have brightness dips or synchronization breaks. Because of this, the confidence score in YOLOv9 becomes unstable: the model sometimes finds an object, sometimes loses it, sometimes confuses it with the background. The fifth problem is color and brightness distortion. Neural detection networks often use not only geometric features, but also color, contrast, and texture statistics. An analog camera, due to automatic exposure, white balance, matrix noise, and losses in the transmitter and receiver, can produce changed colors, overexposed areas, dark areas, or "floating" contrast. If YOLOv9 was trained on high-quality digital images, then the analog FPV image becomes a different data domain for it, that is, a domain shift occurs: the model sees the wrong data from which it was trained. The sixth reason is digitization artifacts. To feed an analog signal to YOLOv9, it must be captured through a video receiver, AV-to-USB converter, or frame grabber. At this stage, additional losses occur: incorrect scaling, compression, delay, frame drops, deinterlacing errors, unstable FPS, YUV/RGB color space conversion, and additional noise. That is, YOLOv9 does not work with the “camera frame,” but with the frame after several imperfect transformations. The seventh problem is motion blur and rolling/field distortion during movement. Analog FPV cameras are often used on mobile robots, drones, or mobile platforms where there are vibrations, turns, changes in lighting, and rapid movement of objects. In this mode, the image loses sharpness, and YOLO determines the bounding box, class and center of the object worse. This is especially pronounced for small objects, thin contours, QR/markers, people at a distance, road signs or technical details. The eighth reason is a poor match between the training dataset and real video. If YOLOv9 is trained or retrained on pure digital frames, it will not work well on analog 5.8 GHz video without special adaptation. To increase accuracy, it is necessary to form a dataset from a real analog channel: with noise, stripes, signal drops, different RSSI levels, different antennas, lighting, movement and range. Otherwise, the model learns on one type of image, but works on a completely different one.

According to the analysis, the following system architecture is proposed for preprocessing the video stream from the FPV drone camera to improve processing using the YOLOv9 neural network, which is presented in Figure 1.

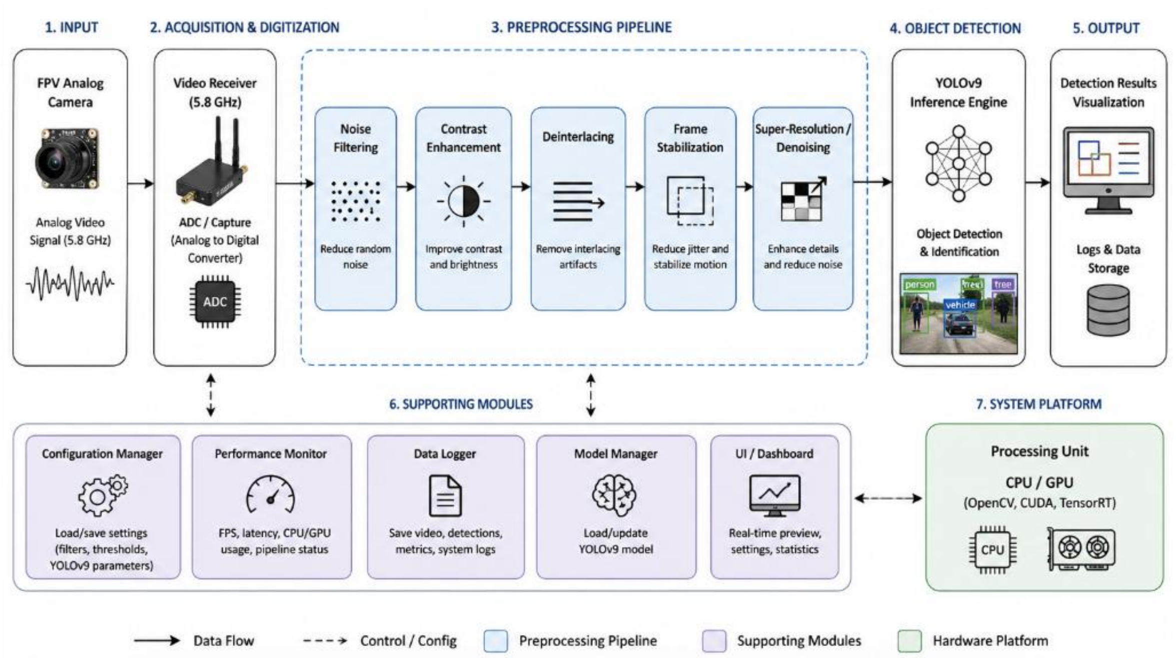

Figure 1 - System architecture for pre-processing a video stream from an FPV drone camera

Let us describe the purpose of each element of the developed architecture of the pre-processing system for the video stream from the FPV drone camera (Fig. 1):

FPV Analog Camera – a video signal source that forms an analog image of the scene in real time. The camera transmits the signal with minimal delay, but without digital processing, which makes it sensitive to noise and distortion;

Video Receiver (5.8 GHz) – an analog signal receiver that provides its demodulation after transmission over the radio channel. At this stage, the video quality largely depends on the signal propagation conditions and the level of interference;

ADC / Capture (Analog to Digital Converter) – a digitization module that converts the analog video signal into a digital format for further processing. This is where additional quality losses occur due to sampling, quantization and possible synchronization errors;

Noise Filtering – a pre-processing stage aimed at reducing random noise and artifacts that arise during signal transmission. This allows you to improve image quality and improve the visibility of objects;

Contrast Enhancement – a module to increase contrast and brightness, which improves the visibility of scene details. This is especially important for highlighting objects in difficult lighting conditions;

Deinterlacing – the process of eliminating the interlacing structure of the video, which allows you to obtain a complete frame without motion artifacts. This improves the quality of the geometry of objects and the stability of their recognition;

Frame Stabilization – a video stream stabilization module that compensates for camera shake and platform movement. This provides more stable frames for analysis by the neural network;

Super-Resolution/Denoising – a stage of increasing resolution and additional noise reduction, which allows you to restore fine image details. This is critical for detecting small or distant objects;

YOLOv9 Inference Engine – the main module of the neural network that performs object detection and identification on processed frames. It uses deep learning to determine object classes and coordinates in real time;

Detection Results Visualization – a module for displaying results that superimposes bounding boxes, classes, and probabilities on images. This allows the operator or system to interpret the results of the model;

Logs & Data Storage – a data storage system that records detection results, video and system metrics. This is necessary for further analysis, training and optimization of algorithms;

Configuration Manager – a system parameter management module that allows you to configure filters, detection thresholds and model parameters. Provides flexibility in adapting the system to different conditions;

Performance Monitor – a performance monitoring tool that tracks FPS, latency and resource usage. This allows you to assess the efficiency of the system in real time;

Data Logger – a component for collecting and recording data on system operation, including video and processing results. Used for analysis and further training of models;

Model Manager – a neural network management module responsible for loading, updating and configuring YOLOv9. Provides the ability to adapt the model to new data;

UI / Dashboard – a user interface that displays video, detection results and system parameters. Provides convenient operator interaction with the system;

Processing Unit (CPU / GPU) – a computing platform that performs all stages of video processing and neural network inference. The use of GPUs or accelerators allows for real-time operation.

As a result, the main problem is that the analog 5.8 GHz video path creates information losses even before the neural network: the resolution decreases, the contours collapse, the colors change, noise, stripes, interlacing, frame instability and digitization artifacts appear. YOLOv9 can effectively process complex images, but it requires stable and representative input data; if the signal is degraded, the model receives incomplete or distorted features, which reduces the classification accuracy, reduces confidence, and increases the number of missed objects and false positives. The most correct solution is not only to “improve YOLO”, but to build a complete preprocessing pipeline: a high-quality receiver, stable digitization, noise filtering, contrast correction, deinterlacing, frame stabilization, super-resolution or denoising, as well as additional training of YOLOv9 on frames from an analog 5.8 GHz camera.

REFERENCES

1. Pultarova, T. (2026). Rise of the Autonomous Attack Drones: AI-Enabled Drones Being Developed by Ukraine and Russia are Remaking the Battlefield. IEEE Spectrum, 63(4), 40-47.

2. Yevsieiev V. Classification of Digital Twins in Collaborative Robot Modeling Problems / V. Yevsieiev, K. Luchaninov // Computer-integrated technologies, automation and robotics 2026 : Proceedings of III st All-Ukrainian Conference, May 14-15, 2026. - Kharkiv .: [electronic version], 2026. - P. 93-96

3. Industry 5.0 та колаборативна робототехніка: динамічний опис навколишнього середовища роботів-маніпуляторів з використанням мови Python: монографія / І. Ш. Невлюдов, В. В. Євсєєв, С. С. Максимова. – Харків : Видавництво Іванченка І. С., 2026. – 279 с. https://doi.org/10.30837/978-617-8332-95-2

4. Model with Neural Network Component for Adaptive Manipulator Control under Variable Load / Amer Abu-Jassar, Mohammad Hamdan, Nowfal Aweisi, Mahmoud Howaidi, V. Yevsieiev, V. Lyashenko // International Journal of Intelligent Engineering and Systems. –19(1). – 2026. – P. 855-868. https://doi.org/10.32403/10.22266/ijies2026.0131.51

5. Nevliudov , I. ., Omarov , M. ., Yevsieiev , V. ., Maksymova , S. ., & Jabrayilzade , E. . (2026). MATHEMATICAL MODELING OF TRAJECTORIES CONSTRUCTION, MOVEMENT OF THE GRIPPING DEVICE OF A COLLABORATIVE ROBOT. Advanced Information Systems, 10(1), 11–20. https://doi.org/10.20998/2522-9052.2026.1.02

6. Yevsieiev V. Digital Twin in Modeling and Control of Collaborative Robots: Analysis, Comparison and Application Recommendations / V. Yevsieiev, S. Starikova // Computer-integrated technologies, automation and robotics 2026 : Proceedings of III st All-Ukrainian Conference, May 14-15, 2026. - Kharkiv .: [electronic version], 2026. - P. 89-92

7. Chebanchyk D. Analysis of Object Identification Methods for FPV Drones / D. Chebanchyk, V. Yevsieiv // Manufacturing & Mechatronic Systems 2025 : Thesises of Reports of IX-st International Conference, October 25-26, 2025. - Kharkiv, 2025. - P. 30-33.

8. Development of a program for processing 3d models of objects in a collaborative robot workspace using an HD camera / V. Yevsieiev, S. Maksymova, Ahmad Alkhalaileh, D. Gurin // Acumen: International Journal of Multidisciplinary Research. – 2025. – Vol. 2(1). – 194-210.

9. Невлюдов І. Ш. Проектування мобільних роботів на базі одноплатних комп'ютерів (Raspberry Pi и мови Python 3.6 : підручник / І. Ш. Невлюдов, А. О. Андрусевич, В. В. Євсєєв – Харків, 2020. - 257 с.

10. 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.

11. 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.

12. Невлюдов І. Ш. BEAM робототехніка : навч. посіб. / І. Ш. Невлюдов, В. В. Євсєєв, С. С. Максимова ; Харків. нац. ун-т радіоелектроніки, кафедра комп’ютерно-інтегрованих технологій, автоматизації та робототехніки (КІТАР). – Кривий Ріг : Видавець Чернявський Д. О., 2024. – 276 с. – ISBN 978-617-8045-79-1

13. Chala, O., Yevsieiev, V., Maksymova, S., & Abu-Jassar, A. (2025). Using the Human Face Recognition Method Based on the MobileNetV2 Neural Network in Authentication Systems. Multidisciplinary Journal of Science and Technology, 5(3), 882-895.

Scientific adviser: *Владислав Євсєєв, д.т.н., професор, професор Department of Computer Integrated Technologies, Automation, Robotics and Safety Engineering, Kharkiv National University of Radio Electronics.*