



DOI

10.66073/ReFocus-Conf/
0910275



QR-kod
maqolaning DOI
sahifasini
ochadi

SAHIFALAR

1326–1328

TIL

Ingliz

IQTIBOS UCHUN

Utegenova Aliya Timurovna,
Abdigappar Kydyrbek Seitzhapparuly.
The impact of dynamic H.265 video
compression on YOLO-based object
detection algorithms for high-speed
UAV navigation // «O'zbekiston olimi
ayollarining mamlakat taraqqiyotiga
qo'shayotgan hissasi» mavzusidagi
xalqaro ilmiy-amaliy konferensiya
materiallari (Nukus, 2026-yil 9–10-
oktyabr). – Nukus, 2026. – B. 1326–
1328. – DOI: [https://doi.org/
10.66073/ReFocus-Conf/0910275](https://doi.org/10.66073/ReFocus-Conf/0910275)



Olimo

«O'zbekiston olimi ayollarining
mamlakat taraqqiyotiga
qo'shayotgan hissasi»

XALQARO ILMIY-AMALIY
KONFERENSIYA
NUKUS, 9–10-OKTYABR 2026

THE IMPACT OF DYNAMIC H.265 VIDEO COMPRESSION ON YOLO-BASED OBJECT DETECTION ALGORITHMS FOR HIGH-SPEED UAV NAVIGATION

Utegenova Aliya Timurovna

Student

Abdigappar Kydyrbek Seitzhapparuly

Master, Senior Lecturer

Sherkhan Murtaza International Taraz University

e-mail: abdifpv@gmail.com

Abstract: This paper examines how dynamic bitrate compression within digital video transmission frameworks for micro-UAVs (specifically, the DJI O4 Unit Pro) affects computer vision performance. We analyze the degradation in object detection accuracy (mean Average Precision, mAP) of PyTorch-based YOLOv8 models caused by H.265 encoding artifacts and system latency during high-speed aerial maneuvers. Ultimately, the empirical results provide a foundation for optimizing off-board artificial intelligence architectures.

Keywords: UAV navigation, H.265 compression, YOLOv8, object detection, digital transmission artifacts, FPV drones, off-board AI.

Introduction (Literature Review)

Contemporary First-Person View (FPV) drones heavily rely on high-fidelity visual streams to navigate autonomously through intricate environments. Given the constrained onboard computational capabilities of micro-UAVs (such as the Pavo 20 Pro), processing video feeds off-board via a powerful desktop PC presents a highly efficient alternative. Nevertheless, to preserve radio link bandwidth, systems employ dynamic H.265 compression, which inherently introduces motion blur and macroblocking during rapid maneuvers. This research aims to establish the direct correlation between these digital transmission artifacts and the resulting accuracy of neural network inference.

Materials and Methods

The experimental setup is grounded in the Hardware-in-the-Loop (HITL) paradigm. The primary UAV platform utilized was the Pavo 20 Pro micro-drone, equipped with the DJI O4 Unit Pro digital video system. The visual stream, received through DJI Goggles 3, was routed directly to a PC (featuring an NVIDIA RTX GPU) using the standard USB Video Class (UVC) protocol. The software stack included Ubuntu OS, ROS 2, and a GStreamer pipeline utilizing Video4Linux2 (v4l2) for hardware-accelerated decoding. Object detection was performed using a YOLOv8 model trained within the PyTorch ecosystem. Overall video stream delay (ST_{total}) was precisely recorded using hardware-level timers and high-frequency photodiodes. Trials were conducted in a controlled indoor environment across three distinct flight profiles: stationary hover,

moderate forward flight (5 m/s), and aggressive maneuvering with rapid turns (15 m/s).

Results

The primary task for the neural network was identifying ArUco fiducial markers and standard spatial obstacles. The detection accuracy (mAP@0.5) processed from the compressed UVC live stream was benchmarked against the raw, uncompressed video recorded directly to the local camera storage (DVR).

- **Stationary Hover (0 m/s):** At a dynamic bitrate of ~20 Mbps, visual artifacts were minimal. The detection accuracy remained high, showing 98.2% on the DVR and 96.5% on the UVC stream, with an average latency of ~28 ms.

- **Moderate Flight (5 m/s):** At ~35 Mbps, edge blurring became noticeable. The mAP dropped slightly to 94.1% (DVR) versus 85.3% (UVC), while latency increased to ~32 ms.

- **Aggressive Maneuvering (15 m/s):** Operating at ~50 Mbps, severe macroblocking occurred. The accuracy divergence was significant: 88.5% on the DVR compared to a steep decline to 61.2% on the UVC feed. Latency peaked at ~45 ms.

During high-speed axial rotations (yaw/roll exceeding 200°/s), it was observed that the H.265 encoder struggles to generate new I-frames (Keyframes) rapidly enough. This limitation induces severe macroblocking in the live feed, causing the neural network to lose critical object contours. Furthermore, the spike in dynamic bitrate forces the transmission buffers to expand, substantially extending the total system latency.

Conclusions and Recommendations

To mitigate the adverse effects of dynamic compression, this study substantiates the necessity of retraining detection models on datasets intentionally augmented with H.265 compression noise, rather than relying solely on pristine images. Additionally, fine-tuning the GStreamer pipeline for hardware-level decoding (NVDEC) while capturing the UVC stream via OpenCV can effectively slash video decoding latency down to 12 ms. When designing off-board AI architectures, striking a precise balance between video fidelity, processing latency, and detection accuracy is paramount for enhancing the reliability of autonomous aerial systems.

References

1. Jocher, G., Chaurasia, A., & Qiu, J. (2023). *YOLO by Ultralytics* (Version 8.0.0) [Computer software]. Ultralytics. <https://github.com/ultralytics/ultralytics>
2. Szeliski, R. (2022). *Computer Vision: Algorithms and Applications* (2nd ed.). Springer.
3. Macenski, S., Foote, T., Gerkey, B., Lalancette, C., & Woodall, W. (2022). Robot Operating System 2: Design, architecture, and uses in the wild. *Science Robotics*, 7(66), eabm6074. <https://doi.org/10.1126/scirobotics.abm6074>
4. Hossain, S., & Lee, D. (2019). Deep Learning-Based Real-Time Multiple-Object Detection and Tracking from Aerial Imagery via a Flying Robot with GPU-Based Edge Computing. *Sensors*, 19(15), 3371. <https://doi.org/10.3390/s19153371>

5. Bradski, G. (2000). The OpenCV Library. *Dr. Dobb's Journal of Software Tools*, 25(11), 120-125.

Information about the authors

Abdigappar Kydyrbek – Senior Lecturer of the Department of Information and Communication Technologies, Sherkhan Murtaza International Taraz University, Republic of Kazakhstan, **080000, Taraz, 69B Zheltoksan st.**;

e-mail: abdifpv@gmail.com; ORCID: <https://orcid.org/0009-0003-6907-1560>.

https://t.me/Abdi_fpv

+77077550055

Utegenova Aliya – Student, Sherkhan Murtaza International Taraz University, Republic of Kazakhstan, **080000, Taraz, 69B Zheltoksan st.**

e-mail: aliyautegenovs2006@gmail.com;