International Journal of Safety and Security Engineering
Vol. 16, No. 7, July, 2026, pp. 1409-1422
Journal homepage: http://iieta.org/journals/ijsse

Enhancing Drone Surveillance Reliability Through Structured Hybrid Data Reinforcement under Environmental Variability

Astika Ayuningtyas[1,2*], Imam Riadi[1], Anton Yudhana[1]

[1] Doctoral Program in Informatics, Universitas Ahmad Dahlan, Yogyakarta 55191, Indonesia
[2] Department of Informatics, Adisutjipto Institute of Aerospace Technology, Yogyakarta 55198, Indonesia

Corresponding Author Email: 2437083006@webmail.uad.ac.id

https://doi.org/10.18280/ijsse.160701

Received: 21 May 2026
Revised: 6 July 2026
Accepted: 13 July 2026
Available online: 31 July 2026

Keywords:
drone surveillance, surveillance systems, operational reliability, hybrid data reinforcement, YOLOv5

ABSTRACT

The increasing use of unmanned aerial vehicles (UAVs) has created significant challenges for aerial surveillance systems operating under heterogeneous environmental conditions, including illumination variation, motion blur, background complexity, and scale diversity. This study proposes a structured hybrid data reinforcement framework for robust drone detection by integrating proportional hybrid dataset construction, illumination harmonisation, Adaptive Multi-Context Augmentation (AMCA), aeromodelling-based synthetic reinforcement, and Convolutional Block Attention Module (CBAM)-enhanced YOLOv5 feature extraction. The proposed framework was evaluated using the DAB-Sky hybrid dataset, comprising 39,717 images collected from public, private, and synthetic aerial sources to improve environmental representativeness and reduce domain imbalance. All experiments were conducted under deterministic benchmark settings to ensure reproducibility and fair comparison. Experimental evaluation demonstrated that the proposed Aero-YOLOv5 achieved 96.30% precision, 93.70% recall, 96.30% mAP@0.5, and 71.00% mAP@0.5:0.95. Convergence analysis confirmed stable optimisation behaviour, with a validation loss of 0.048. The runtime evaluation achieved an average throughput of 9.27 FPS, demonstrating the feasibility of continuous aerial monitoring with the experimental hardware configuration, though additional optimisation is still necessary for safety-critical, real-time deployment. These findings demonstrate that systematic hybrid dataset reinforcement, combined with attention-guided feature learning, improves surveillance robustness, localisation consistency, and detection reliability across heterogeneous environmental conditions without requiring substantial modifications to the underlying detection architecture.

1. INTRODUCTION

The rapid evolution of uncrewed aerial vehicle (unmanned aerial vehicles (UAVs)/Drone) technology has greatly expanded its use in surveillance, logistics, infrastructure inspection, agriculture, and military operations. Despite these advantages, the increasing number of drones flying in controlled and restricted airspace has created significant operational safety and security challenges for modern aerial surveillance systems [1-4]. Unauthorised drone activities can have serious implications for airport operations, critical infrastructure security, and public safety, and can introduce weaknesses in intelligent surveillance environments. This was exemplified in 2018 by the closure of Gatwick Airport in the United Kingdom due to unlawful drone activity, which led to delays for hundreds of aircraft [5, 6]. More recent reports indicate that airspace violation incidents continue to rise worldwide, underscoring the growing need for reliable.

surveillance systems that maintain robust detection performance across heterogeneous environments [7-9]. Accordingly, improving the reliability and operational robustness of aerial surveillance systems has become an important research priority, particularly in safety-critical environments where missed detections may lead to operational disruption and security risks.

In practical aerial surveillance, drone detection remains challenging due to illumination variation, atmospheric disturbance, motion blur, scale diversity, and complex backgrounds, which often reduce localisation accuracy and detection consistency. These limitations are particularly critical in safety-sensitive applications such as airport surveillance, border security, critical infrastructure protection, and urban airspace monitoring, where missed detections may compromise situational awareness [10]. Consequently, improving surveillance robustness under heterogeneous environmental conditions has become a fundamental requirement for intelligent drone monitoring systems. These challenges have motivated the widespread adoption of deep learning-based object detection methods that learn robust visual representations from heterogeneous aerial imagery.

Recent advances in deep learning have improved aerial object detection by enabling models to learn robust visual representations directly from large-scale image datasets. Convolutional Neural Networks (CNNs) and single-stage object detectors, particularly the You Only Look Once (YOLO) family, have become dominant approaches because they balance detection accuracy and computational efficiency, making them well suited for real-time aerial surveillance applications [11-13]. Attention mechanisms such as the Convolutional Block Attention Module (CBAM) have also been incorporated into YOLO-based detectors to enhance feature representation by emphasising informative spatial regions and channel-wise responses. These attention-enhanced architectures have demonstrated improved localisation performance, especially for small aerial targets under cluttered backgrounds and varying illumination conditions [14-16].

Beyond architectural improvements, recent research has increasingly emphasised the role of data-centric AI, in which the quality, diversity, and representativeness of training data are considered equally important for improving detection performance. Data augmentation and synthetic data generation have therefore emerged as effective strategies for enriching aerial datasets and increasing model robustness under diverse environmental conditions. Conventional augmentation techniques, including rotation, scaling, flipping, and photometric transformations, have demonstrated their ability to improve generalisation by introducing moderate geometric and illumination variations. Likewise, synthetic data generation has enabled the simulation of challenging operational scenarios that are difficult, costly, or unsafe to acquire through field observations alone [15, 17-22]. These developments suggest that surveillance reliability depends not only on detector architecture but also on systematic dataset engineering. However, existing studies remain limited by reliance on public or laboratory-controlled datasets, isolated augmentation strategies, and a predominant focus on architectural optimisation rather than on integrated data-centric reinforcement. Consequently, the combined effects of hybrid dataset engineering, illumination harmonisation, adaptive augmentation, and deterministic experimental protocols remain insufficiently explored, motivating the development of a unified surveillance-oriented framework for robust drone detection.

To address these limitations, this study proposes a structured hybrid data-reinforcement framework for robustness in drone surveillance across heterogeneous environmental conditions. The proposed framework provides four principal contributions. First, a proportional hybrid dataset construction method is introduced to improve environmental representativeness and reduce domain imbalance by integrating diverse image sources with complementary characteristics. Second, illumination harmonisation, Adaptive Multi-Context Augmentation (AMCA), and aeromodelling-based synthetic reinforcement are integrated into a unified data-centric reinforcement pipeline. Third, CBAM-enhanced YOLOv5 is employed to improve feature discrimination and localisation consistency for small aerial targets under complex surveillance scenarios. Finally, a deterministic experimental protocol is established to ensure reproducibility, fairness in benchmarking, and objective evaluation across different model configurations. Collectively, these contributions provide a reproducible, surveillance-oriented framework that enhances operational robustness and supports the development of reliable intelligent aerial surveillance systems capable of operating under heterogeneous environmental conditions with improved adaptability and generalisation capability.

2. LITERATURE REVIEW

2.1 Drone surveillance systems and operational challenges

The rapid proliferation of UAVs has substantially increased demand for intelligent aerial surveillance systems to support airport security, border monitoring, infrastructure inspection, disaster response, and restricted airspace protection [23-25]. In safety-critical environments, surveillance systems must maintain reliable target detection and continuous situational awareness despite heterogeneous operational conditions. However, practical drone surveillance remains challenging because UAVs often appear as small targets against complex backgrounds and are affected by illumination variation, atmospheric disturbances, motion blur, occlusion, and scale diversity [26, 27]. These environmental factors frequently reduce detection reliability, increase false alarms, and elevate the risk of missed detections, thereby degrading situational awareness during continuous surveillance operations. Consequently, recent research has increasingly focused on developing robust detection frameworks that maintain surveillance reliability and stable operational performance across diverse environmental conditions.

2.2 Deep learning-based drone detection

Recent advances in deep learning have substantially improved aerial object detection by enhancing feature learning and enabling real-time inference [12, 28-35]. Among existing detectors, the YOLO family has become one of the most widely adopted approaches because it provides an effective balance between computational efficiency and detection accuracy, making it suitable for real-time aerial surveillance [36, 37]. Successive YOLO architectures, including YOLOv3, YOLOv4, YOLOv5, and YOLOv7, have progressively improved feature aggregation and localisation performance for small aerial objects [38, 39]. Nevertheless, most existing studies have been evaluated using public benchmarks or laboratory-controlled datasets with relatively homogeneous environmental characteristics. Consequently, the robustness of these models under operational surveillance conditions involving dynamic illumination, cluttered backgrounds, and motion degradation remains insufficiently investigated.

2.3 Data-centric reinforcement and hybrid dataset engineering

Beyond architectural improvements, recent research has increasingly emphasised data-centric AI, in which the quality and representativeness of training data are considered equally important to improving detection performance. Public UAV datasets often offer limited environmental diversity and may therefore not adequately represent operational surveillance scenarios involving varying weather, lighting, and flight dynamics [22]. To overcome this limitation, data augmentation and synthetic data generation have become widely adopted strategies to increase dataset diversity and improve model generalisation [40]. Conventional augmentation methods introduce variability but do not effectively capture the complex interactions between environmental changes and aerial motion. Synthetic data has shown considerable potential for simulating challenging surveillance conditions that are difficult to acquire through real-world data collection. However, most existing studies employ these techniques independently, without systematically integrating illumination harmonisation, adaptive augmentation, and synthetic reinforcement into a unified data-centric framework. Consequently, their combined contribution to surveillance robustness and operational reliability remains insufficiently explored.

2.4 Attention mechanisms for robust aerial surveillance

Attention mechanisms have become an important component of modern object detectors, improving feature discrimination and localisation accuracy through adaptive feature weighting [41-44]. Among these approaches, the CBAM sequentially applies channel and spatial attention to emphasise informative features while suppressing background interference [15, 22]. Previous studies have demonstrated that CBAM enhances small-object detection and improves localisation performance in complex visual environments [15, 45-47]. However, attention mechanisms alone cannot fully compensate for the lack of environmental diversity in the training data. Their effectiveness remains strongly dependent on dataset quality, representativeness, and the ability of the training samples to capture operational surveillance variability. Consequently, combining attention-enhanced feature learning with structured hybrid dataset engineering offers a promising strategy to improve surveillance robustness across heterogeneous environmental conditions

2.5 Comparative analysis of previous studies

Previous drone detection studies have primarily focused on detector architecture, feature extraction, or data reinforcement as separate approaches. Table 1 summarises these representative studies and their limitations.

In contrast, the proposed framework integrates hybrid dataset engineering, adaptive augmentation, synthetic reinforcement, and attention-guided feature learning into a reproducible, surveillance-oriented methodology to improve operational robustness across heterogeneous environmental conditions.

Table 1. Comparison of representative approaches for aerial surveillance and drone detection [19, 21, 48, 49]

Representative Approach	Main Contribution	Remaining Limitation
YOLOv4-based drone detection	Developed a YOLOv4-based framework for real-time drone detection under controlled aerial environments.	Limited environmental variability; no attention mechanism or structured dataset reinforcement.
Improved YOLOv5s for SAR aircraft detection	Improved YOLOv5s for aircraft detection in SAR imagery through architectural refinement.	Focused on SAR imagery without hybrid dataset engineering or surveillance-oriented environmental adaptation.
CBAM-enhanced YOLOv8	Enhanced feature representation using CBAM within a YOLOv8 detector.	Evaluated on industrial defect detection rather than aerial surveillance applications.
Synthetic data augmentation	Reviewed recent synthetic data augmentation techniques for improving model robustness and dataset diversity.	Did not propose an integrated framework combining synthetic reinforcement with attention-guided surveillance systems.
Proposed Aero-YOLOv5	Integrates proportional hybrid dataset construction, illumination harmonization, AMCA, aeromodelling-based synthetic reinforcement, CBAM-enhanced YOLOv5, and deterministic experimental validation for surveillance reliability.	-

Note: Adaptive Multi-Context Augmentation (AMCA); Convolutional Block Attention Module (CBAM).

2.6 Research gap and study contribution

Existing drone detection studies exhibit several important limitations, including limited evaluation under real-world aerial surveillance conditions, insufficient integration of environmental harmonisation and adaptive reinforcement strategies, and a predominant emphasis on detector architecture optimisation rather than structured data-centric approaches [15, 17-22].

Moreover, the combined effects of proportional hybrid dataset engineering, illumination harmonisation, AMCA, aeromodelling-based synthetic reinforcement, attention-guided feature learning, and deterministic experimental validation have not been comprehensively investigated within a unified surveillance-oriented framework. To address these gaps, this study proposes a structured hybrid data-reinforcement framework that integrates proportional hybrid dataset construction, illumination harmonisation, AMCA, aeromodelling-based synthetic reinforcement, and CBAM-enhanced YOLOv5 feature extraction under deterministic benchmark settings to improve surveillance reliability, operational robustness, and detection consistency in heterogeneous aerial environments. The methodological framework, developed from the identified research gaps, is presented in the following section.

3. METHODOLOGY

The objective of this study is to develop a systematic hybrid data-reinforcement framework to improve the reliability of drone surveillance systems operating under heterogeneous environmental conditions. As illustrated in Figure 1, the proposed methodology was designed to satisfy four fundamental requirements of engineering research: reproducibility, architectural clarity, experimental validity, and benchmark fairness. All methodological procedures, including dataset preparation, preprocessing, model configuration, experimental settings, and evaluation protocols, are explicitly specified to enable independent replication and objective verification. Dataset partitioning was performed using stratified sampling before any augmentation to preserve class distribution. In contrast, all augmentation and synthetic reinforcement procedures were applied exclusively to the training partition to prevent information leakage into the validation and test sets. The experimental protocol was designed to isolate the contribution of each methodological component by controlling architectural and computational variables across each benchmark comparison, while intentionally varying specific components to meet the experimental objectives. The overall workflow comprises six stages: dataset preparation, preprocessing and illumination harmonisation, model construction, deterministic experimental configuration, performance evaluation, and benchmark validation, as shown in Figure 1.

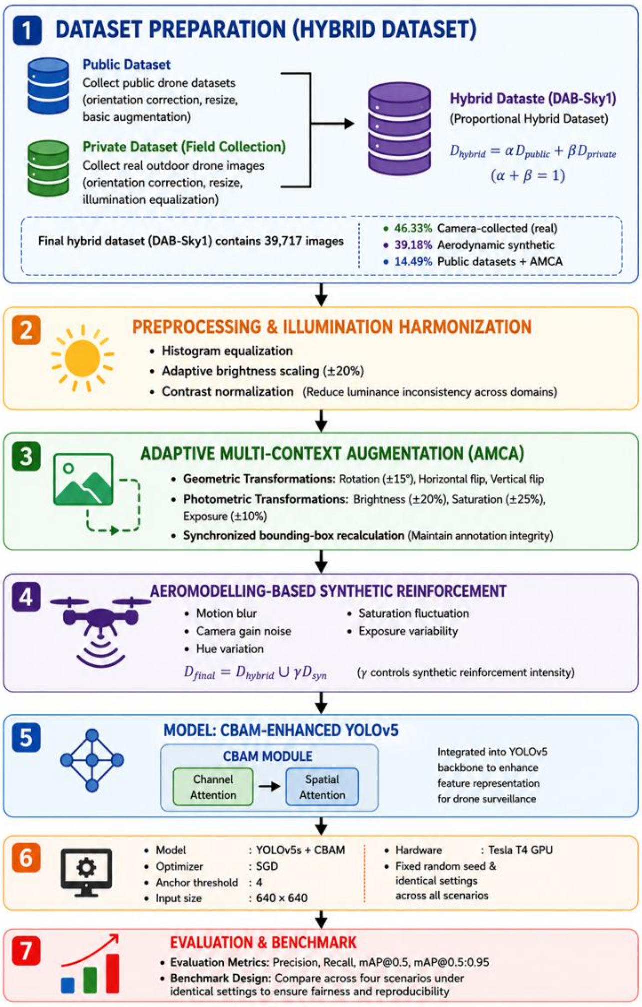


Figure 1. Structured hybrid reinforcement framework for robust drone surveillance under environmental variability

3.1 Dataset preparation

The first stage of the proposed framework establishes the DAB-Sky dataset, which serves as the foundation for all experiments. The dataset was constructed by integrating publicly available UAV images with a locally acquired aerial surveillance dataset to improve environmental diversity and operational representativeness. Figure 1 illustrates the overall dataset preparation workflow, while the following subsections describe each stage in detail.

(1) Public dataset collection

Public UAV datasets from previous object detection studies [40, 50] were used to establish the baseline aerial surveillance dataset. These datasets were selected because they contain diverse drone appearances, viewing angles, and environmental conditions. To ensure consistency across heterogeneous data sources, all images were automatically oriented and resized to 640 × 640 pixels before further processing.

(2) Private dataset acquisition

To improve representativeness under real-world aerial surveillance conditions, a private dataset was collected through outdoor image acquisition [51-53]. Data collection was designed to capture realistic operational variability, including illumination changes, atmospheric effects, motion dynamics, object-scale variation, and diverse sky backgrounds. Drone trajectories, viewing distances, flight orientations, and ambient lighting conditions were intentionally varied to represent practical surveillance scenarios.

(3) Proportional hybrid dataset construction

Rather than directly merging heterogeneous datasets, a proportional hybridisation strategy was employed to reduce domain imbalance while preserving environmental diversity. The hybrid dataset is defined as shown in Eq. (1).

$$D_{hybrid} = \alpha D_{public} + \beta D_{private} \tag{1}$$

where, D_{hybrid} denotes the resulting hybrid dataset, D_{public} and $D_{private}$ represent the public and private. The proportional weights satisfy, as shown in Eq. (2).

$$\alpha + \beta = 1 \tag{2}$$

thereby maintaining balanced dataset composition and preventing dominance from a single environmental domain.

(4) Dataset composition

The DAB-Sky dataset contains 39,717 images collected from three complementary sources: camera-collected aerial images, aerodynamically generated synthetic images, and publicly available UAV datasets enhanced using AMCA.

This hybrid composition was designed to increase environmental diversity, reduce domain bias, and improve model generalisation under heterogeneous surveillance conditions. Table 2 summarises the distribution of object classes within the DAB-Sky dataset.

Drone images constitute the largest proportion of the dataset to ensure adequate representation of the primary surveillance target. Bird images were intentionally included because of their visual similarity to drones, thereby reducing false-positive detections. Meanwhile, the sky, sun, and tree classes provide realistic environmental contexts commonly encountered in aerial surveillance, thereby improving background diversity and feature discrimination. The composition of data sources is summarised in Table 3.

The integration of real, synthetic, and AMCA-enhanced images provides complementary environmental characteristics that improve surveillance-oriented feature learning while reducing domain imbalance during model training.

Table 2. Distribution of object classes in the DAB-Sky dataset

Object Class	Number of Images	Percentage (%)
Drone	17.902	45.08
Bird	6.812	17.15
Sky	9.664	24.33
Sun	3.646	9.18
Tree	1.693	4.26
Total	39.717	100.00

Table 3. Composition of the DAB-Sky hybrid dataset

Data Source	Percentage (%)
Camera-Collected Data	46.33
Aerodynamic Synthetic Data	39.18
Public datasets + AMCA-enhanced images	14.49
Total	100.00

(5) Dataset partitioning

To ensure reproducible and unbiased performance evaluation, the DAB-Sky dataset was partitioned into independent training, validation, and test subsets using a stratified sampling strategy that preserved both object-class distribution and environmental diversity across all partitions, as shown in Table 4.

Table 4. Dataset partitioning strategy

Object Class	Number of Images	Percentage (%)
Training Set	35.984	89
Validation Set	2.939	7
Test Set	794	4
Total	39.717	100.00

The training partition was intentionally allocated the largest proportion of the dataset to maximise exposure to diverse drone poses, illumination conditions, and environmental variations during model optimisation. The validation partition was used for hyperparameter tuning and convergence monitoring, whereas the test partition was reserved exclusively for final performance evaluation.

Although the test partition represents only 4% of the complete dataset, stratified sampling ensured that all object classes and environmental conditions remained proportionally represented, providing an unbiased assessment while maximising the amount of data available for model training. Dataset partitioning was completed before any augmentation or synthetic reinforcement was applied. Subsequently, AMCA and aeromodelling-based synthetic reinforcement were applied exclusively to the training partition, whereas the validation and test partitions remained unchanged across all experiments. This protocol prevented information leakage and ensured fair benchmark comparison across all evaluated model configurations. As presented in Table 4, 89% of the dataset was allocated for training to maximise exposure to rare drone poses, illumination variations, and synthetic perturbations.

In comparison, 7% and 4% were reserved for validation and independent testing, respectively. Although the test subset comprises only 794 images, it was generated using stratified sampling to preserve the original class proportions and environmental diversity of the complete DAB-Sky dataset. Consequently, all object categories (drone, bird, sky, sun, and tree) remained adequately represented in the independent test partition, providing a reliable basis for unbiased performance evaluation while preventing data leakage between the training and test partitions.

3.2 Preprocessing and illumination harmonisation

To improve photometric consistency across heterogeneous aerial surveillance conditions, illumination harmonisation was applied before geometric augmentation. As described in Section 3.1.5, this preprocessing step was performed exclusively on the training partition after dataset splitting to prevent information leakage for each input image, as shown in Eq. (3).

$$\mathrm{I} \in D_{hybrid} \tag{3}$$

where, I denote an input image and D_{hybrid} represents the proportionally constructed hybrid dataset; photometric normalisation was applied to reduce illumination inconsistency while preserving discriminative visual information.

The harmonisation procedure consisted of three sequential operations: Global histogram equalisation, Adaptive brightness scaling (±20%), and Contrast normalisation. Histogram equalisation redistributed image intensity to improve global contrast, adaptive brightness scaling simulated moderate illumination variability commonly encountered during aerial surveillance, and contrast normalisation enhanced feature consistency across different acquisition conditions. The harmonised images served as input to AMCA, ensuring that subsequent geometric transformations were applied to photometrically normalised data.

3.3 Adaptive Multi-Context Augmentation

To improve model robustness under heterogeneous aerial surveillance conditions, AMCA was developed to generate synchronised geometric and photometric perturbations while preserving the integrity of annotations. Consistent with the dataset partitioning strategy described in Section 3.1.5, AMCA was applied exclusively to the training partition after dataset splitting. Unlike conventional augmentation methods that apply transformations independently, AMCA applies coordinated transformations to images and annotations within predefined parameter ranges. Each training sample is represented as an image–annotation pair (I, B), where, I denotes the input image and B represents the corresponding bounding-box annotations, including the object centre coordinates, width, height, and class label. Image transformations and annotation updates were performed synchronously to maintain spatial consistency throughout the augmentation process.

Geometric transformations consisted of rotation (±15°), horizontal flipping, and vertical flipping. Because these operations modify the object's spatial orientation, bounding-box coordinates were automatically recalculated using geometric transformation matrices, followed by boundary clipping, removal of degenerate bounding boxes, and minimum-area verification to preserve annotation consistency. Photometric transformations consisted of brightness adjustment (±20%), exposure variation (−10%), and saturation variation (+25%), thereby simulating realistic illumination variability encountered in aerial surveillance environments. Unlike geometric transformations, these operations modify only the pixel intensity and colour distribution without altering object geometry; therefore, the original bounding-box annotations remain unchanged. Representative AMCA operations are illustrated in Figure 2. As shown in Figure 2(a), geometric transformations require automatic bounding-box recalculation to preserve annotation consistency after spatial modification. In contrast, Figure 2(b) demonstrates that photometric transformations alter only the image appearance while preserving the original bounding-box annotations, as object geometry is unaffected.

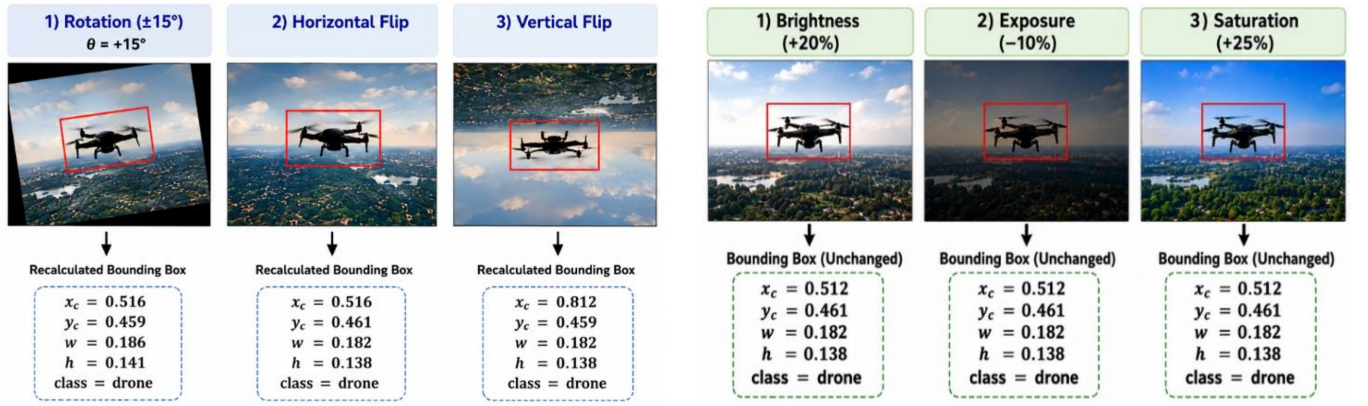
3.4 Aeromodelling-based synthetic reinforcement

To improve model robustness under heterogeneous aerial surveillance conditions, aeromodelling-based synthetic reinforcement was employed to generate realistic operational perturbations that are difficult to capture through field data acquisition. Controlled environmental and motion perturbations were applied to the training data to produce a synthetic dataset, denoted by D_{SYN}. The generation process was governed by a predefined transformation parameter set, θ_{SYN}, which controls illumination variation, viewpoint changes, motion blur, atmospheric effects, and geometric perturbations. Consistent with the dataset-partitioning strategy described in Section 3.1, synthetic reinforcement was applied exclusively to the training partition after dataset splitting, while the validation and test partitions remained unchanged to prevent information leakage and ensure unbiased performance evaluation.

Five surveillance-oriented perturbation models were incorporated: Motion blur, simulating drone and camera movement; Camera gain noise, modelling sensor noise and electronic gain variation; Hue variation, representing atmospheric colour shifts; Saturation fluctuation, simulating weather- and illumination-induced colour variation; Exposure variability, modelling over- and under-exposure caused by changing illumination conditions. The final reinforced training dataset is defined by Eq. (4).

$$D_{final} = D_{hybrid} \cup \gamma\, D_{SYN} \quad (4)$$

where, D_{final} denotes the reinforced training dataset, D_{hybrid} represents the proportionally constructed hybrid dataset, D_{SYN} denotes the synthetic sample set, and $\gamma \in [0,1]$ controls the reinforcement intensity. Representative synthetic perturbations are illustrated in Figure 3. Each perturbation model employs predefined parameter ranges to simulate realistic surveillance conditions while preserving object identity and annotation consistency, thereby improving environmental diversity and feature generalisation during training.



(a) Geometric transformation (b) Photometric transformation

Figure 2. Representative Adaptive Multi-Context Augmentation (AMCA) operations. (a) Geometric transformations with automatic bounding-box recalculation to preserve annotation consistency, and (b) photometric transformations with unchanged bounding-box annotations, as object geometry remains unaffected

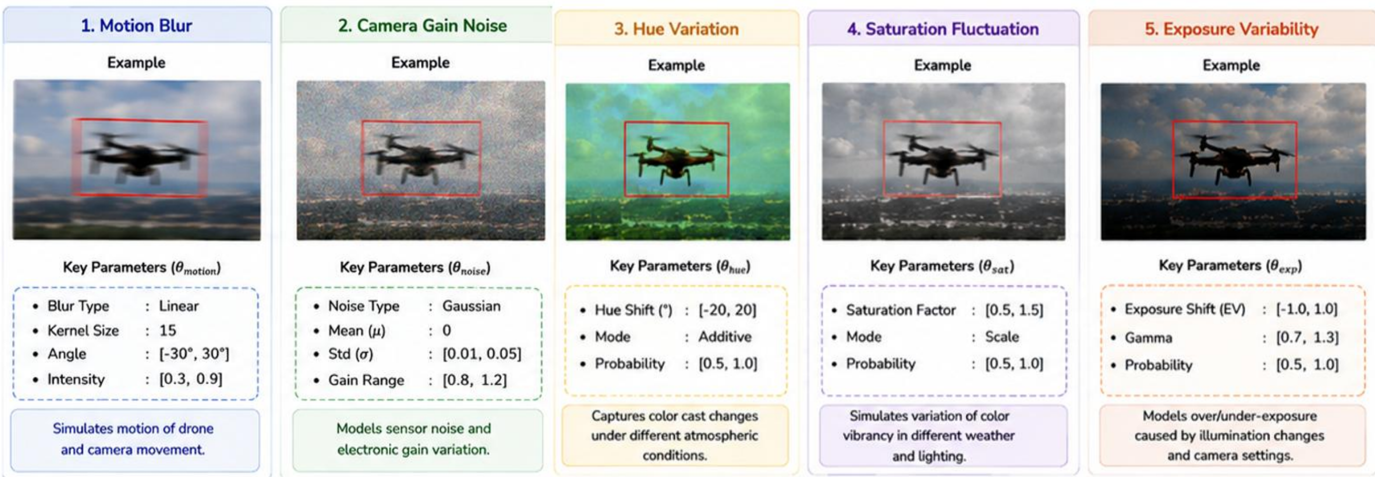


Figure 3. Representative aeromodelling-based synthetic reinforcement is applied during training, together with the corresponding transformation parameters used to simulate realistic aerial surveillance conditions

3.5 Convolutional Block Attention Module-enhanced YOLOv5 architecture

To evaluate the contribution of attention-guided feature learning, the CBAM was incorporated into the selection of YOLOv5 configurations within the proposed framework. The baseline YOLOv5 architecture was retained, while CBAM was integrated into the backbone feature extraction network without modifying the overall detection pipeline. CBAM sequentially applies the channel attention followed by spatial attention to improve feature representation. The channel attention mechanism emphasises informative feature channels associated with drone characteristics, whereas the spatial attention mechanism highlights discriminative object regions while suppressing background interference. This sequential attention process improves feature discrimination for small aerial targets observed under heterogeneous surveillance conditions. Because CBAM was introduced as a lightweight attention module, the overall computational structure of YOLOv5 remained unchanged, thereby preserving inference efficiency while enhancing surveillance-oriented feature representation. The effectiveness of CBAM was subsequently evaluated through benchmark experiments comparing it against the baseline YOLOv5 architecture under identical conditions.

3.6 Experimental setup and reproducibility protocol

All experiments were conducted under deterministic conditions to ensure reproducibility, objective evaluation, and fair benchmarking. The experimental protocol was designed to isolate the contribution of individual methodological components while maintaining identical computational conditions wherever applicable. Across all benchmark configurations, the YOLOv5s backbone, pretrained weights, hardware platform, dataset partition, image resolution, and evaluation metrics were kept constant. In contrast, the incorporation of CBAM and the optimisation strategy (SGD, AdamW, or AdamW with cosine learning rate scheduling) were intentionally varied in selected experiments to evaluate their individual contributions to detection performance.

(1) Hardware Configuration: All experiments were performed on an NVIDIA Tesla T4 GPU with CUDA acceleration and a single CPU core. Maintaining an identical hardware platform ensured computational consistency across all benchmark configurations.
(2) Software Environment: The implementation was based on the Ultralytics YOLOv5 framework, using Python 3.12.12 and the pretrained yolov5s model.pt weights.

The input image resolution, anchor configuration, training epochs, and evaluation metrics were kept identical throughout all experiments, whereas the optimiser was varied only in the designated benchmark scenarios.

(3) Deterministic Training Protocol: To ensure experimental reproducibility, all benchmark configurations employed fixed random seed initialisation, identical dataset partitioning, consistent preprocessing procedures, and standardised hyperparameter settings whenever applicable. Only the methodological components under investigation, namely dataset reinforcement, CBAM integration, and optimisation strategy, were intentionally modified according to the benchmark design. This controlled protocol enabled performance differences to be attributed to the evaluated methodological factors rather than to uncontrolled experimental variation.

3.7 Evaluation metrics

Detection performance was evaluated using four widely adopted object detection metrics: Precision, Recall, mean Average Precision at an Intersection-over-Union threshold of 0.5 (mAP@0.5), and mean Average Precision averaged over IoU thresholds from 0.5 to 0.95 (mAP@0.5:0.95). Precision measures the proportion of correctly identified drone detections among all predicted detections, whereas Recall quantifies the proportion of detected drone targets relative to all ground-truth objects. The mAP@0.5 metric evaluates detection accuracy using a fixed IoU threshold of 0.5, while mAP@0.5:0.95 provides a more stringent assessment by averaging localisation performance across multiple IoU thresholds. Together, these complementary metrics provide a comprehensive evaluation of detection accuracy, localisation quality, and operational robustness under heterogeneous aerial surveillance conditions, enabling fair comparison across all benchmark configurations.

3.8 Benchmark design and fair experimental comparison

To systematically evaluate the contribution of each methodological component, six benchmark configurations were designed to represent the progressive development of the proposed surveillance framework. Each configuration introduces a specific combination of hybrid dataset engineering, illumination harmonisation, AMCA, aeromodelling-based synthetic reinforcement, CBAM integration, and optimisation strategy, enabling incremental assessment of their individual and combined effects on drone detection performance. To ensure benchmark fairness, several experimental factors were maintained constant across all benchmark configurations, including the YOLOv5s backbone, pretrained weights, input image resolution, dataset partition, hardware platform, evaluation metrics, and deterministic training protocol. In contrast, CBAM integration and optimiser selection (SGD, AdamW, and AdamW with cosine learning-rate scheduling) were intentionally varied across the corresponding benchmark configurations to quantify their individual contributions. At the same time, dataset reinforcement was progressively introduced in accordance with the experimental design. This controlled benchmark design enabled performance differences to be attributed to the evaluated methodological components rather than to uncontrolled variations in computational settings. Consequently, the proposed framework provides an objective and reproducible basis for comparing the influence of data-centric reinforcement, attention-guided feature learning, and training optimisation on surveillance performance under heterogeneous aerial environmental conditions.

4. RESULTS AND DISCUSSION

This section presents the experimental evaluation of the proposed structured hybrid data reinforcement framework using the DAB-Sky dataset under deterministic benchmark settings. The analysis examines the effects of progressive dataset engineering, CBAM-enhanced feature learning, and training optimisation on detection accuracy, localisation performance, convergence behaviour, and operational robustness. The results are discussed through benchmark comparisons, convergence analyses, ablation studies, comparisons with representative previous studies, and real-time operational evaluations to demonstrate the effectiveness of the proposed framework for intelligent aerial surveillance under heterogeneous environmental conditions.

4.1 Detection performance and benchmark analysis

To evaluate the effectiveness of the proposed framework, six benchmark configurations were developed by progressively introducing hybrid dataset engineering, illumination harmonisation, AMCA, aeromodelling-based synthetic reinforcement, CBAM integration, and optimisation strategies. The comparative results are summarised in Table 5. As shown in Table 5, the baseline YOLOv5 achieved 91.84% precision, 88.58% recall, 92.01% mAP@0.5, and 49.38% mAP@0.5:0.95. Although the baseline detector provided satisfactory detection accuracy, the relatively low mAP@0.5:0.95 indicates limited localisation consistency under stricter Intersection-over-Union (IoU) thresholds. Incorporating additional public aerial data substantially improved localisation performance. In contrast, direct integration of the hybrid dataset alone did not consistently improve all evaluation metrics, indicating that increasing dataset diversity without appropriate environmental harmonisation may introduce domain inconsistencies.

The highest overall performance was obtained by the proposed Aero-YOLOv5, which combines structured hybrid dataset engineering, illumination harmonisation, AMCA, aeromodelling-based synthetic reinforcement, CBAM-enhanced feature learning, and SGD optimisation. The proposed model achieved 96.30% precision, 93.70% recall, 96.30% mAP@0.5, and 71.00% mAP@0.5:0.95, demonstrating consistent improvements in both detection accuracy and localisation performance. These results suggest that the observed performance gains arise from the combined contribution of balanced dataset engineering, attention-guided feature learning, and appropriate optimisation rather than from any individual component alone. Representative detection outputs obtained under identical experimental conditions are presented in Figure 4. Visual comparison shows progressive improvement in detection confidence, localisation consistency, and robustness across different environmental conditions, particularly under variations in illumination, background complexity, and viewing distance. The proposed Aero-YOLOv5 consistently produced tighter bounding-box localisation and higher confidence scores than the baseline configuration, indicating improved discrimination between drone targets and visually similar background objects.

Table 5. Benchmark comparison of model configurations

Configuration	Public Data	Synthetic Data	Illumination Harmonization	AMCA	CBAM	Optimizer	Precision (%)	Recall (%)	mAP@0.5 (%)	mAP@0.5:0.95 (%)
Baseline YOLOv5	×	×	×	×	×	SGD	91.84	88.58	92.01	49.38
YOLOv5 + Public Dataset	√	×	×	×	×	SGD	88.59	88.59	93.68	73.00
YOLOv5 + Hybrid Dataset	√	√	×	×	×	SGD	91.15	83.19	86.86	68.70
YOLOv5 + Hybrid Dataset + Multi-Scale	√	√	√	√	√	AdamW + Cosine LR	95.90	84.20	95.10	67.30
YOLOv5 + Hybrid Dataset + CBAM	√	√	√	√	√	AdamW	95.50	90.50	95.40	67.50
YOLOv5 + Hybrid Dataset + CBAM (Proposed Aero-YOLOv5)	√	√	√	√	√	SGD	96.30	93.70	96.30	71.00

Note: Adaptive Multi-Context Augmentation (AMCA); Convolutional Block Attention Module (CBAM).

(a) (b) (c) (d)

Figure 4. Representative drone detection results obtained under identical experimental conditions. (a) Baseline YOLOv5; (b) YOLOv5 with public dataset reinforcement; (c) YOLOv5 with hybrid dataset engineering; and (d) the proposed Aero
Note: YOLOv5 integrates structured dataset reinforcement, CBAM-enhanced feature learning, and SGD optimization

4.2 Comparison with existing studies

To provide broader context for the proposed framework, Aero-YOLOv5 was compared with several representative YOLOv5-based drone detection studies reported in the recent literature. Because the compared studies employed different datasets, sensing modalities, and experimental protocols, the comparison should be regarded as indicative rather than a strictly controlled benchmark. The objective is therefore to position the proposed framework relative to existing approaches rather than to establish a definitive performance ranking. As summarised in Table 6, previous studies have reported competitive detection performance with conventional YOLOv5 architectures across various aerial surveillance datasets. The proposed Aero-YOLOv5 achieved 96.30% precision, 93.70% recall, 96.30% mAP@0.5 and 71.00% mAP@0.5:0.95 on the DAB-Sky dataset, indicating competitive performance while maintaining robust localisation under heterogeneous surveillance conditions. The observed results suggest that structured hybrid dataset engineering, illumination harmonisation, AMCA, aeromodelling-based synthetic reinforcement, and CBAM-enhanced feature learning contribute positively to feature representation and localisation consistency.

Although direct numerical comparison should be interpreted with caution because of differences in datasets and evaluation settings, the proposed framework demonstrates that systematic data-centric reinforcement can effectively complement attention-guided feature learning for robust drone detection.

4.3 Convergence stability and optimization analysis

All benchmark configurations were trained under deterministic experimental settings using identical hardware, dataset partitions, random seed initialisation, and training protocols to ensure reproducible convergence analysis. Figure 5 illustrates the evolution of training loss, validation loss, and mAP@0.5 across the four-representative dataset-reinforcement scenarios. Across all configurations, the training and validation losses decreased smoothly, with no noticeable divergence or unstable oscillations, while mAP@0.5 increased steadily throughout training, indicating stable optimisation behaviour. The public baseline achieved a final validation loss of 0.085 with 92.01% mAP@0.5. Incorporating illumination harmonisation and AMCA reduced the validation loss to 0.061 and increased mAP@0.5 to 93.68%, suggesting improved feature consistency under heterogeneous environmental conditions. In contrast, the Initial Hybrid Dataset Integration configuration produced a higher validation loss (0.072) and a lower mAP@0.5 (86.86%), indicating that increased dataset diversity alone is insufficient without appropriate environmental balancing.

Table 6. Indicative comparison with representative YOLOv5-based drone detection studies

Ref.	Year	Dataset	Method	Precision (%)	Recall (%)	mAP@0.5 (%)	mAP@0.5:0.95 (%)
[10]	2023	Drone Detection Dataset	YOLOv5	91.80	87.30	90.40	N/A
[52]	2025	Drone Dataset	YOLOv5 (50 epochs)	91.80	88.60	92.00	49,4
[54]	2023	UAV Thermal Dataset	YOLOv5s	86.20	80.10	85.20	56.90
Proposed	2026	DAB-Sky	YOLOv5+CBAM+ AMCA+Synthetic Reinforcement	96.30	93.70	96.30	71.00

Note: Convolutional Block Attention Module (CBAM); unmanned aerial vehicles (UAVs).

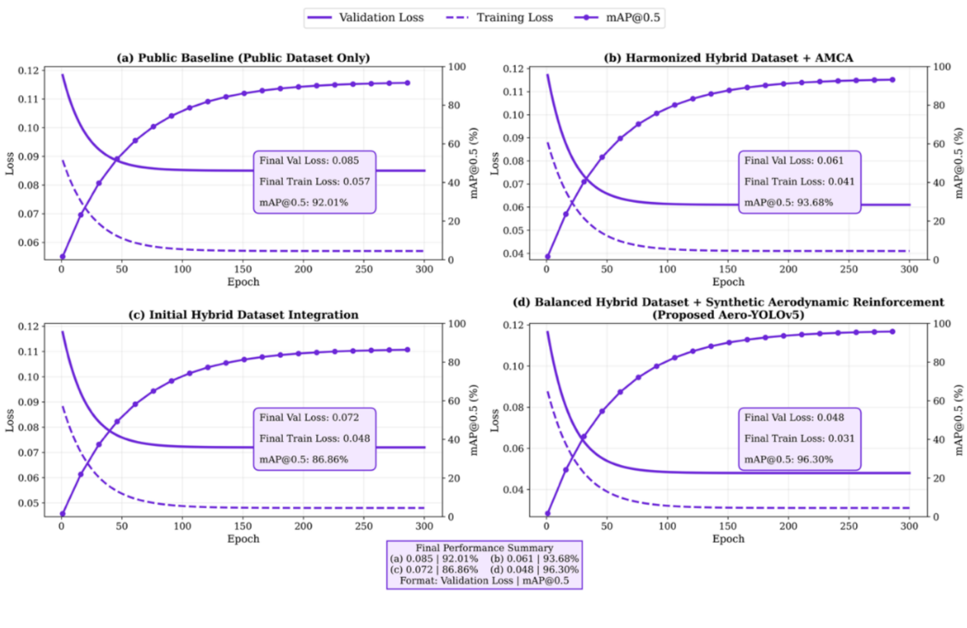


Figure 5. Convergence analysis of representative benchmark configurations showing training loss, validation loss, and mAP@0.5 progression under deterministic experimental settings: (a) public baseline; (b) harmonised hybrid dataset with AMCA; (c) initial hybrid dataset integration; and (d) proposed Aero-YOLOv5 with balanced hybrid dataset engineering and aeromodelling-based synthetic reinforcement

The proposed Aero-YOLOv5 exhibited the most favourable convergence characteristics, achieving the lowest validation loss (0.048), lowest training loss (0.031), and highest mAP@0.5 (96.30%) among the evaluated configurations.

As illustrated in Figure 5(d), the training process converged smoothly, with consistent improvements in detection accuracy. These observations suggest that combining structured hybrid dataset engineering, illumination harmonisation, AMCA, aeromodelling-based synthetic reinforcement, and CBAM-enhanced feature learning contributes to stable optimisation and improved feature generalisation under heterogeneous aerial surveillance conditions.

4.4 Ablation study and component contribution analysis

An ablation analysis was conducted to evaluate the contribution of each component in the proposed framework. As shown in Table 5, illumination harmonisation, AMCA, CBAM-enhanced feature learning, and appropriate optimisation progressively improved detection performance, whereas hybrid dataset integration alone did not consistently enhance accuracy, highlighting the importance of balanced dataset engineering. Overall, the results indicate that the superior performance of Aero-YOLOv5 arises from the synergistic integration of structured dataset reinforcement, attention-guided feature learning, and optimisation rather than from any individual component.

4.5 Scientific and operational implications

The findings of this study provide important implications for both data-centric AI and intelligent aerial surveillance. From a scientific perspective, the results suggest that structured hybrid dataset engineering, illumination harmonisation, AMCA, and aeromodelling-based synthetic reinforcement complement attention-guided feature learning by improving environmental representativeness and feature generalisation [55-59]. The ablation analysis further indicates that increasing dataset diversity alone is insufficient; balanced environmental representation and appropriate adaptation strategies are equally important for achieving stable optimisation and robust detection performance.

From an operational perspective, the proposed Aero-YOLOv5 demonstrated reliable detection accuracy and stable convergence under heterogeneous surveillance conditions. Although the measured processing speed supports continuous aerial monitoring under the experimental hardware configuration, additional optimisation and edge-device validation are required before deployment in safety-critical real-time surveillance systems. Nevertheless, the proposed framework shows potential for applications such as airport perimeter surveillance, border monitoring, infrastructure protection, and low-altitude airspace observation, where improved localisation consistency and reduced false detections can support more reliable operational decision-making [60, 61].

4.6 Real-time operational evaluation

The operational evaluation assessed the computational performance of Aero-YOLOv5 during continuous video-stream inference under the experimental hardware configuration. As illustrated in Figure 6, the proposed framework successfully detected drone targets while reporting confidence score, inference latency, processing speed, and image resolution for each frame.

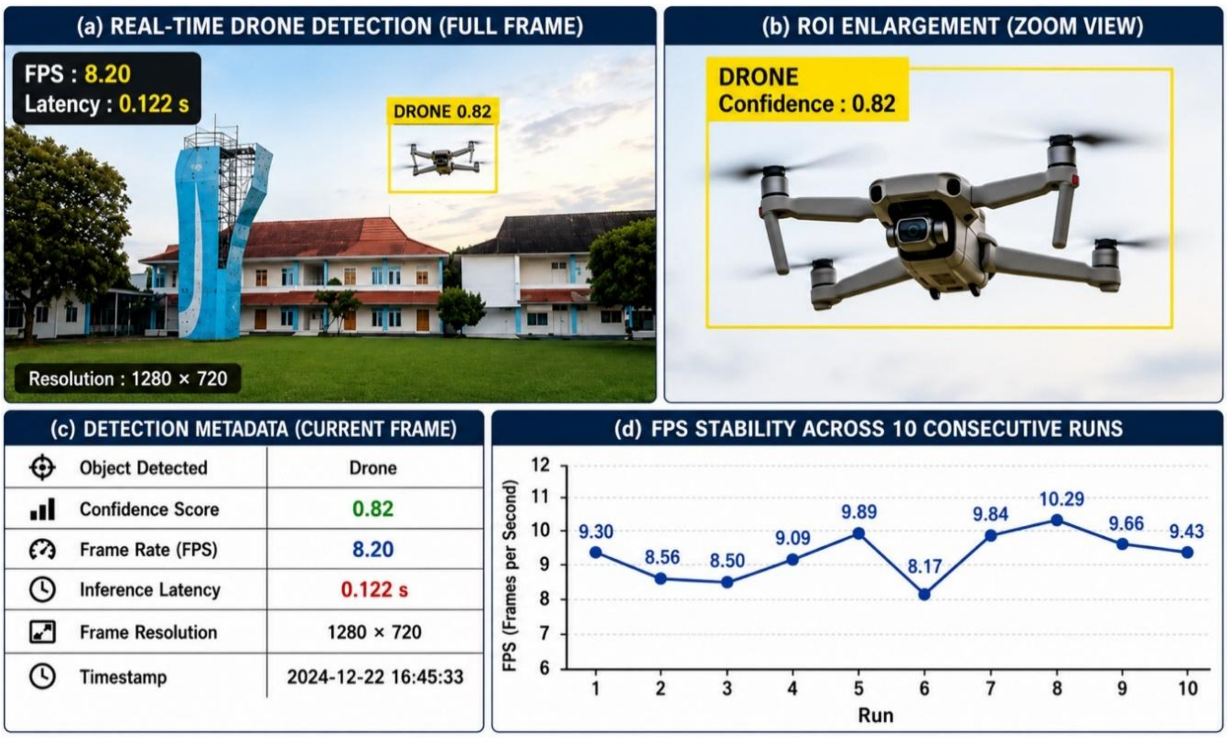


Figure 6. Real-time operational evaluation of Aero-YOLOv5 under the experimental hardware configuration: (a) full-frame drone detection; (b) enlarged ROI; (c) representative inference metadata including confidence score, FPS, latency, and image resolution; and (d) example of continuous operational monitoring

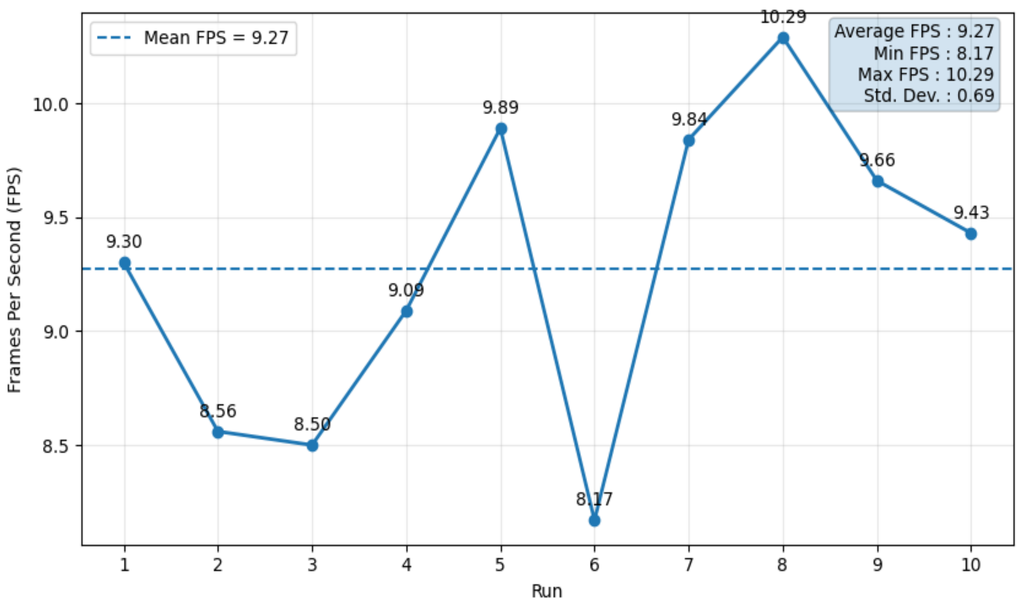


Figure 7. The processing speed stability of Aero-YOLOv5 was measured over ten consecutive operational runs under the experimental hardware configuration

The enlarged region of interest (ROI) in Figure 6(b) demonstrates accurate localisation of a relatively small drone target, achieving a representative confidence score of 0.82 under outdoor surveillance conditions. The corresponding inference latency was approximately 0.122 s, indicating stable inference throughout the evaluation. The computational stability of the proposed framework is further illustrated in Figure 7, which summarises the frame processing speed across 10 consecutive operational runs. The average throughput was 9.27 FPS, ranging from 8.17 FPS to 10.29 FPS, with a standard deviation of 0.69 FPS, indicating consistent computational behaviour with minimal performance fluctuations during continuous operation. These results suggest that Aero-YOLOv5 can support continuous aerial monitoring under the experimental conditions evaluated. However, the achieved throughput may not yet meet the stricter real-time requirements of safety-critical airspace surveillance systems, and further optimisation, along with embedded edge-device validation, will be required before operational deployment. Nevertheless, the proposed framework demonstrates reliable target localisation and stable inference performance, supporting its potential application in airport perimeter surveillance, infrastructure protection, border monitoring, and other intelligent aerial observation tasks.

4. CONCLUSIONS

This study proposed a structured hybrid data-reinforcement framework for robust drone detection across heterogeneous aerial surveillance conditions. The proposed Aero-YOLOv5 integrates proportional hybrid dataset construction, illumination harmonisation, AMCA, aeromodelling-based synthetic reinforcement, and CBAM-enhanced feature learning within a unified surveillance-oriented pipeline. Using the DAB-Sky dataset comprising 39,717 images from real, public, and synthetic sources, the proposed framework achieved 96.30% precision, 93.70% recall, 96.30% mAP@0.5, and 71.00% mAP@0.5:0.95 under deterministic benchmark settings. Convergence analysis indicated stable optimisation with the lowest validation loss (0.048), while operational evaluation achieved an average throughput of 9.27 FPS, supporting continuous aerial monitoring under the evaluated experimental hardware configuration.

The results suggest that the performance improvements arise from the synergistic interaction of structured hybrid dataset engineering, illumination harmonisation, AMCA, aeromodelling-based synthetic reinforcement, attention-guided feature learning, and appropriate optimisation, rather than from architectural modification alone. An important finding of this study is that increasing dataset diversity alone does not necessarily improve detection performance; balanced environmental representation and controlled dataset reinforcement are equally important for achieving robust generalisation. Overall, the proposed framework supports the growing paradigm of data-centric AI for aerial surveillance. Future work will focus on transformer-based detection architectures, multimodal sensor fusion, embedded edge-device optimisation, and adverse-weather evaluation to further improve operational scalability and practical deployment readiness.

REFERENCES

[1] Srivastav, A., Shandilya, S., Ortiz-Rodriguez, F. (2026). Introduction to drones, UAVs, and their applications. In AI, Cybersecurity and Data Science for Drones and Unmanned Aerial Vehicles. https://doi.org/10.1201/9781003481584-1

[2] Gupta, C., Pandit, S., Das, R.K., Nayak, S.B. (2026). Unmanned aerial vehicle: Technological development and its applications. In Remote sensing of the North Eastern Himalayan Ecosystem. Springer, pp. 723-756. https://doi.org/10.1007/978-981-95-3221-6_24

[3] Singh, M.K., Singha, N.S., Prabhakar, V.D. (2026). Advanced AI-enabled UAV swarms for real-time air traffic surveillance. In Machine Learning Based Air Traffic Surveillance System Using Image Processing, Emerald Publishing Limited. https://doi.org/10.1108/978-1-80592-062-520251006

[4] Singh, R., Kumar, S. (2025). A comprehensive insights

into drones: History, classification, architecture, navigation, applications, challenges, and future trends. arXiv preprint arXiv:2501.10066. https://doi.org/10.48550/arXiv.2501.10066
[5] Shvetsova, S.V., Shvetsov, A.V. (2021). Ensuring safety and security in employing drones at airports. Journal of Transportation Security, 14: 41-53. https://doi.org/10.1007/s12198-020-00225-z
[6] Kotkova, B. (2022). Airport defense systems against drones attacks. In 2022 26th International Conference on Circuits, Systems, Communications and Computers (CSCC), Crete, Greece, pp. 85-90. https://doi.org/10.1109/cscc55931.2022.00025
[7] Howard, R.O. (2023). FAA unmanned aircraft systems (UAS) sighting reports: A preliminary survey. In AIAA AVIATION 2023 Forum, San Diego, CA, USA, p. 4099. https://doi.org/10.2514/6.2023-4099
[8] Mistry, S., Degadwala, S. (2024). Improved multi-type vehicle recognition with a customized YOLO. In 2024 4th International Conference on Pervasive Computing and Social Networking (ICPCSN), Salem, India, pp. 361-365. https://doi.org/10.1109/icpcsn62568.2024.00063
[9] Weber, L. (2023). International Civil Aviation Organization (ICAO). In Handwörterbuch Internationale Organisationen, Springer, pp. 211-213. https://doi.org/10.1007/978-3-322-86673-8_66
[10] Aydin, B., Singha, S. (2023). Drone detection using YOLOv5. Eng, 4(1): 416-433. https://doi.org/10.3390/eng4010025
[11] Putri Silmina, E., Sunardi, S., Yudhana, A. (2025). Comparative analysis of YOLO deep learning model for image-based beef freshness detection. JITK (Jurnal Ilmu Pengetahuan dan Teknologi Komputer), 11(1): 250-265. https://doi.org/10.33480/jitk.v11i1.6784
[12] Yudhana, A., Reski, J.M. (2025). Classification of crystallization images of pharmaceutical raw materials using convolutional neural network algorithm. International Journal of Advances in Data and Information Systems, 6(3): 705-715. https://doi.org/10.59395/ijadis.v6i3.1440
[13] Kozlenko, M., Demiral, E., Yudhana, A. (2026). Demodulation of chaotic signals using convolutional neural network. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLVIII-4/W19-2025: 79-84. https://doi.org/10.5194/isprs-archives-xlviii-4-w19-2025-79-2026
[14] Riadi, I., Yudha, R.P., Fadlil, A. (2025). Comparison method of convolutional neural network and support vector machine for facial expression recognition. Journal of Innovation Information Technology and Application (JINITA), 7(2): 223-233. https://doi.org/10.30595/jinita.v7i2.2836
[15] Hussaini, H., Bano, S., Elyan, E., Moreno-Garcia, C.F. (2025). Modified CBAM: Sub-block pooling for improved channel and spatial attention. In Annual Conference on Medical Image Understanding and Analysis, Leeds, UK, pp. 116-130. https://doi.org/10.1007/978-3-031-98691-8_9
[16] Zakariah, M., Alnuaim, A. (2024). Recognizing human activities with the use of convolutional block attention module. Egyptian Informatics Journal, 27: 100536. https://doi.org/10.1016/j.eij.2024.100536
[17] Agustien, L., Rohman, T., Hujairi, A.W. (2021). Real-time deteksi masker berbasis deep learning menggunakan algoritma CNN YOLOv3. Jurnal Teknologi Informasi dan Terapan, 8(2): 129-137. https://doi.org/10.25047/jtit.v8i2.246
[18] Muwardi, R., Nugroho, I.P., Salamah, K.S., Yunita, M., Rahmatullah, R., Chung, G.J. (2024). Optimization of YOLOv4-tiny algorithm for vehicle detection and vehicle count detection embedded system. Jurnal Ilmiah Teknik Elektro Komputer dan Informatika, 10(3): 639-648. https://doi.org/10.26555/jiteki.v10i3.29693
[19] Singha, S., Aydin, B. (2021). Automated drone detection using YOLOv4. Drones, 5(3): 95. https://doi.org/10.3390/drones5030095
[20] Lenhard, T.R., Weinmann, A., Franke, K., Koch, T. (2025). SynDroneVision: A synthetic dataset for image-based drone detection. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, Tucson, AZ, USA, pp. 7637-7647. https://doi.org/10.1109/wacv61041.2025.00742
[21] Mumuni, A., Mumuni, F., Gerrar, N.K. (2024). A survey of synthetic data augmentation methods in machine vision. Machine Intelligence Research, 21(5): 831-869. https://doi.org/10.1007/s11633-022-1411-7
[22] Alkhalisy, M.A.E., Abid, S.H. (2023). Students behavior detection based on improved YOLOv5 algorithm combining with CBAM attention mechanism. International Journal of Intelligent Engineering and Systems, 16(5): 473-487. https://doi.org/10.22266/ijies2023.1031.41
[23] Agarwal, A., Mohanta, C., Mehta, S.N. (2023). Drone technologies: State-of-the-art, challenges, and future scope. In Drone Technology: Future Trends and Practical Applications, pp. 1-19. https://doi.org/10.1002/9781394168002.ch1
[24] Molinari, S., Tomasello, F., Capasso, P.J., Dallau, A. (2024). Sustainable urban drone operations: FF2020 view. In Journal of Physics: Conference Series. IOP Publishing, Salerno, Italy, p. 012061. https://doi.org/10.1088/1742-6596/2716/1/012061
[25] Emimi, M., Khaleel, M., Alkrash, A. (2023). The current opportunities and challenges in drone technology. International Journal of Electrical Engineering and Sustainability, pp. 74-89. https://doi.org/10.65998/ijees.v1i3.47
[26] Wu, P., He, X., Dai, W., Zhou, J., Shang, Y., Fan, Y., Hu, T. (2025). A review on research and application of AI-based image analysis in the field of computer vision. IEEE Access, 13: 76684-76702. https://doi.org/10.1109/access.2025.3565300
[27] Ma, J., Jiang, X., Fan, A., Jiang, J., Yan, J. (2021). Image matching from handcrafted to deep features: A survey. International Journal of Computer Vision, 129(1): 23-79. https://doi.org/10.1007/s11263-020-01359-2
[28] Umar, R., Riadi, I., Miladiah, M. (2018). Sistem identifikasi keaslian uang kertas rupiah menggunakan metode K-means clustering. Techno. COM, 17(2): 179-185. https://doi.org/10.33633/tc.v17i2.1681
[29] Rahman, A., Salim, M., Riadi, I. (2024). Klasifikasi citra spesies bunga di Indonesia berbasis convolutional neural network menggunakan teknik transfer learning. Jurnal Software Engineering and Computational Intelligence, 2(2): 92-100. https://doi.org/10.36982/jseci.v2i02.4942
[30] Dzulqarnain, M.F., Fadlil, A., Riadi, I. (2024). Improving the accuracy of batik classification using deep

convolutional auto encoder. Compiler, 13(2): 123-130. https://doi.org/10.28989/compiler.v13i2.2649

[31] Faroek, D.A., Umar, R., Riadi, I. (2022). Classification based on machine learning methods for identification of image matching achievements. Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi), 6(2): 198-206. https://doi.org/10.29207/resti.v6i2.3826

[32] Subandi, R., Yudhana, A. (2024). Pneumonia medical image classification using convolution neural network model AlexNet & GoogleNet. International Journal of Computing and Digital Systems, 16(1): 1675-1684. https://doi.org/10.12785/ijcds/1601124

[33] Peryanto, A., Yudhana, A., Umar, R. (2022). Convolutional neural network and support vector machine in classification of flower images. Khazanah Informatika: Jurnal Ilmu Komputer dan Informatika, 8(1): 1-7. https://doi.org/10.23917/khif.v8i1.15531

[34] Fahmi, M., Yudhana, A., Sunardi, S. (2023). Pemilahan sampah menggunakan model klasifikasi support vector machine gabungan dengan convolutional neural network. JURIKOM (Jurnal Riset Komputer), 10(1): 76-81.

[35] Muis, A., Sunardi, S., Yudhana, A. (2023). Comparison analysis of brain image classification based on thresholding segmentation with convolutional neural network. Journal of Applied Engineering and Technological Science (JAETS), 4(2): 664-673. https://doi.org/10.37385/jaets.v4i2.1583

[36] Zhang, J., He, X., Yuan, X.D. (2024). Human figure detection in Han portrait stone images via enhanced YOLO-v5. Heritage Science, 12(1): 119. https://doi.org/10.1186/s40494-024-01232-2

[37] Zheng, S., Wang, R., Zheng, S., Wang, L., Jiang, H. (2024). Adaptive density guided network with CNN and Transformer for underwater fish counting. Journal of King Saud University-Computer and Information Sciences, 36(6): 102088. https://doi.org/10.1016/j.jksuci.2024.102088

[38] Anupriya, Kumar, R. (2024). Attention-based improved YOLOv5 for effective waste detection and classification system. In 2024 International Conference on Electrical Electronics and Computing Technologies (ICEECT), Greater Noida, India, pp. 1-6. https://doi.org/10.1109/ICEECT61758.2024.10739282

[39] Zhang, Q., Han, H., Li, F., Du, Y. (2023). Attention-aware based differentiable architecture search network for waste recognition. In 2023 IEEE 13th International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER), Qinhuangdao, China, pp. 408-412. https://doi.org/10.1109/CYBER59472.2023.10256542

[40] Wang, M., Yang, W., Wang, L., Chen, D., Wei, F., KeZiErBieKe, H., Liao, Y. (2023). FE-YOLOv5: Feature enhancement network based on YOLOv5 for small object detection. Journal of Visual Communication and Image Representation, 90: 103752. https://doi.org/10.1016/j.jvcir.2023.103752

[41] Cui, C., Liu, L., Qiao, R. (2024). A cutting-edge video anomaly detection method using image quality assessment and attention mechanism-based deep learning. Alexandria Engineering Journal, 108: 476-485. https://doi.org/10.1016/j.aej.2024.07.103

[42] Wang, Z., Zhou, W., Li, Y. (2024). GFN: A garbage classification fusion network incorporating multiple attention mechanisms. Electronics, 14(1): 75. https://doi.org/10.3390/electronics14010075

[43] He, Y. (2023). Waste sorting using attention mechanism based on deep learning. In 2023 International Conference on Applied Physics and Computing (ICAPC), Ottawa, ON, Canada, pp. 497-501. https://doi.org/10.1109/ICAPC61546.2023.00097

[44] Li, R., Wu, Y. (2022). Improved YOLO v5 wheat ear detection algorithm based on attention mechanism. Electronics, 11(11): 1673. https://doi.org/10.3390/electronics11111673

[45] Wang, F., Song, C. (2025). YOLO-ARM: An enhanced YOLOv7 framework with adaptive attention receptive module for high-precision robotic vision object detection. Alexandria Engineering Journal, 129: 1326-1339. https://doi.org/10.1016/j.aej.2025.09.001

[46] Zhang, X., Zhao, Y., Smith, J. S., Wang, Y., et al. (2026). Ghost-attention you only look once (GA-YOLO): Enhancing small object detection for traffic monitoring. Computers, Materials, & Continua, 87(2): 1. https://doi.org/10.32604/cmc.2026.075415

[47] Dewi, C., Chernovita, H.P., Santoso, M.V., et al. (2026). Improving early fire and smoke detection with YOLOv12N via histogram-based contrast enhancement. International Journal of Safety and Security Engineering, 16(2): 431-445. https://doi.org/10.18280/ijsse.160217

[48] Huang, M., Yan, W., Dai, W., Wang, J. (2023). EST-YOLOv5s: SAR image aircraft target detection model based on improved YOLOv5s. IEEE Access, 11: 113027-113041 https://doi.org/10.1109/ACCESS.2023.3323575

[49] Sun, A., Mao, Z. (2026). YOLO-CFAEW: A YOLOv8-based detection network with CBAM attention, RFAConv module, and multi-head structure for engine casting defect detection. Journal of Intelligent Manufacturing. https://doi.org/10.1007/s10845-026-02801-x

[50] Arezoomandan, S., Klohoker, J., Han, D.K. (2025). Data augmentation pipeline for enhanced UAV surveillance. In International Conference on Pattern Recognition, Kolkata, India, pp. 366-380. https://doi.org/10.1007/978-3-031-78172-8_24

[51] Ayuningtyas, A., Riadi, I., Yudhana, A. (2025). DAB-Sky multitemporal drone dataset. Mendeley Data, V1. https://doi.org/10.17632/6s449dpwv6.1

[52] Ayuningtyas, A., Riadi, I., Yudhana, A. (2025). A comparative evaluation of drone detection models on aerial imagery across varying training epochs. JUITA: Jurnal Informatika, pp. 277-286. https://doi.org/10.30595/juita.v13i3.26618

[53] Ayuningtyas, A., Riadi, I., Yudhana, A. (2025). Comparison of drone and helicopter image classification accuracy using naïve Bayes based on mean red-green-blue (RGB) values and first-order statistics. International Journal of Informatics and Computation, 7(2): 427-439. https://doi.org/10.35842/ijicom.v7i2.144

[54] Soeleman, M.A., Supriyanto, C., Purwanto. (2023). Deep learning model for unmanned aerial vehicle-based object detection on thermal images. Revue d'Intelligence Artificielle, 37(6): 1441-1447. https://doi.org/10.18280/ria.370608

[55] Zha, D., Bhat, Z.P., Lai, K.H., Yang, F., Jiang, Z., Zhong, S., Hu, X. (2025). Data-centric artificial intelligence: A survey. ACM Computing Surveys, 57(5): 1-42. https://doi.org/10.1145/3711118

[56] Jakubik, J., Vössing, M., Kühl, N., Walk, J., Satzger, G. (2024). Data-centric artificial intelligence. Business & Information Systems Engineering, 66(4): 507-515. https://doi.org/10.1007/s12599-024-00857-8

[57] Bhatt, N., Bhatt, N., Prajapati, P., Sorathiya, V., Alshathri, S., El-Shafai, W. (2024). A data-centric approach to improve performance of deep learning models. Scientific Reports, 14(1): 22329. https://doi.org/10.1038/s41598-024-73643-x

[58] Patel, J., Patel, D., Raval, A. (2025). Artificial intelligence for predictive modeling in CRISPR/Cas9 gene editing: A survey of methods and design strategies. Journal of Gene Medicine, 27(12): e70061. https://doi.org/10.1002/jgm.70061

[59] Gao, T., Huang, H., Liu, Y. (2025). Machine learning-driven nanoscale synthesis for electrocatalytic performance: From data-driven methodologies to closed-loop optimization. Advanced Materials, pp. e08263. https://doi.org/10.1002/adma.202508263

[60] Mustafa, M.A.S. (2025). Predictive reliability-driven optimization of spare parts management in aircraft fleets using AI, IoT, and digital twin technologies. Journal of Engineering Management and Systems Engineering, 4(3), 218-236. https://doi.org/10.56578/jemse040305

[61] Singh, P. (2023). Systematic review of data-centric approaches in artificial intelligence and machine learning. Data Science and Management, 6(3): 144-157. https://doi.org/10.1016/j.dsm.2023.06.001