Proceedings of CONF-MLA 2025 Symposium: Intelligent Systems and Automation: AI Models, IoT, and Robotic Algorithms
DOI: 10.54254/2755-2721/2025.LD27858

Optimization of YOLOv8 for UAV-Based Object Detection: A Literature Review

Maoheng Ma

Department of Electrical and Electronic Engineering, The Hong Kong Polytechnic University, Hong Kong, China
mamaoheng123@gmail.com

Abstract. YOLOv8 offers high accuracy and real-time performance, making it suitable for UAV applications. However, challenges such as limited onboard computing power, varying target sizes, and complex environments persist. This review systematically explores recent advancements in optimising YOLOv8 for drone applications, while also studying and evaluating its performance under extreme conditions such as low light and occlusion. The results indicate that current optimisations primarily focus on three directions. First, improving small object detection accuracy through multi-scale feature fusion, attention mechanisms (such as hybrid attention modules), and data augmentation (such as adaptive anchor allocation). Second, achieving model lightweighting and edge deployment through backbone replacement (such as MobileNetV3), model quantisation, and embedded deployment optimisation. Additionally, robustness in complex environments is improved through techniques such as image preprocessing (e.g., CLAHE, gamma correction), adversarial training, and multimodal fusion..This work aims to guide future studies toward automated optimization, edge-cloud integration, and multi-task learning for intelligent UAV vision systems.

Keywords: YOLOv8, Unmanned Aerial Vehicle (UAV), Small Object Detection, Model Lightweighting, Edge Deployment

1. Introduction

With the rapid development of UAV technology and artificial intelligence, vision-based object detection has become a core function in UAV systems [1]. The YOLO series, particularly YOLOv8, is widely adopted for its real-time performance and accuracy on resource-constrained platforms [2]. However, deploying YOLOv8 on UAVs faces challenges such as limited computing power, variable object scales, and complex environmental conditions [3-4]. To address these issues, recent research focuses on improving small-object detection via multi-scale feature fusion and attention mechanisms, and achieving model lightweighting for efficient edge deployment. Some studies also enhance robustness under adverse conditions through preprocessing and adversarial training. Therefore, this review systematically examines current optimization strategies for YOLOv8 in UAV applications, identifies research gaps, and discusses future directions such as automated model design, edge-cloud collaboration, and multi-task learning. This paper aims to systematically review

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2025.LD27858

the optimization path of YOLOv8 in unmanned aerial vehicle (UAV) object detection, with the intention of comprehensively discussing the future development direction based on existing research, and providing theoretical support and practical references for researchers in related fields. From a research perspective, this review can serve as a foundational reference for further exploration into transformer-based detection architectures, self-supervised learning paradigms, and automated model design techniques such as Neural Architecture Search (NAS). In terms of application, future studies could extend the optimization strategies discussed here toward multi-task learning frameworks, enabling UAVs to perform not only object detection but also semantic segmentation, pose estimation, and scene understanding simultaneously. By bridging the gap between theoretical insights and real-world deployment challenges, this line of work holds strong potential for contributing to the advancement of intelligent and autonomous UAV vision systems.

2. Overview of YOLOv8 and UAV target detection

2.1. A brief overview of YOLO series algorithm development

Since YOLOv1 was first proposed in 2016, the YOLO series algorithms have been iteratively updated, and have gradually become one of the representative frameworks in the field of single-stage target detection [5]. YOLOv2 introduces an anchor frame mechanism and a high-resolution classifier to enhance the recognition capability of small targets. YOLOv3 significantly improves the model accuracy through multiscale prediction, and YOLOv4 introduces the CSPDarknet53 backbone network with PANet feature pyramid structure to further enhance the performance [6]. YOLOv5 is more flexible in structure design, supports multiple size models (n/s/m/l/x), and optimizes the inference speed, which is widely used in industrial deployment scenarios [5]. And the latest generation, YOLOv8 has made several architectural improvements on the basis of YOLO series, including anchorless detection head, C2f module, etc., which realizes higher accuracy and faster inference efficiency, and becomes one of the most representative real-time target detection solutions at present [7].

2.2. Characteristics of YOLOv8 and UAV image data

YOLOv8 adopts an advanced backbone network and neck structure, which integrates the advantages of Transformer and CNN to effectively improve the image feature extraction capability [8]. The core improvements include anchor-free head, c2f module, multi-scale feature fusion mechanism, and adaptive anchor frames selection mechanism. Among them, the anchor-free head simplifies the model design and improves the adaptability to targets of different scales. The C2f Module replaces the C3 module in YOLOv5, which enhances the feature expression capability and reduces the amount of computation at the same time [7]. Multi-scale feature fusion mechanism improves the detection of small targets, while the adaptive anchor frames selection mechanism enhances the generalization of the model. In addition, YOLOv8 has good scalability and supports a variety of visual tasks, such as target detection, image segmentation, attitude estimation, etc., which is suitable for complex and changing UAV application scenarios [9].

The images collected by UAV platforms are characterised by high viewing angle and small target, complex background, dramatic changes in illumination, and high dynamics, which put forward higher requirements for target detection algorithms. For example, small pixel ratio and high recognition difficulty by the target in being far away from the camera in aerial images, background interference caused by diverse environments such as cities, forests, and farmlands, uneven

DOI: 10.54254/2755-2721/2025.LD27858

illumination or shadow occlusion caused by changes in flight altitude and angle, and image jitter and exacerbated by the motion state of UAVs [10]. These make traditional target detection methods difficult to be applied directly, and it is urgent to optimize the algorithm for UAV scenes.

2.3. Typical application cases of YOLOv8 in UAV scenarios

YOLOv8 has been widely used in many UAV-related fields due to its high efficiency and accuracy. In agricultural monitoring, YOLOv8 is used to enhance the intelligent level of agriculture, including crop pest and disease identification, livestock tracking, weed detection in the fields, etc. In security patrol, it can identify and track suspicious individuals and vehicles in real-time to ensure public safety. In disaster rescue, it can quickly identify people and obstacles in the disaster-stricken area to assist in decision-making. In power inspection, it can automatically identify transmission line defects, foreign object suspension, and other issues to improve operation and maintenance efficiency. In summary, YOLOv8 shows strong potential in UAV target detection, but it still needs to be further optimized to meet the actual deployment requirements in the face of challenges such as resource constraints, difficulties in identifying small targets, and the impact of complex environments.

3. Optimization direction one: enhance the UAV small object detection precision

3.1. Problems

Due to the uniqueness of small features compared to UAV aerial images based on their high altitude and long-distance shooting as the UAV passes through the complex background, the difficult situations for small object detection increase. Because of the tiny area of the identified small target relative to the overall area of the ground-level images, its pixel ratio with the whole image pixels is extremely low, and the low ratio renders traditional object detection methods ineffective [11]. In addition, the platform does not remain stock while flying, and frequently, there are unplanned changes during the flight, which causes problems like occlusions, illumination changes, image jitter, etc. These issues further worsen the blurriness and uncertainty of small items [12].

3.2. Current improvement strategies

3.2.1. Multi-scale feature fusion

One of the ways to solve the problem of small object detection is the use of multi-scale feature inference. The YOLOv8 paradigm is an object detection model that uses PAFPN (Pyramid Anchors Feature Pyramid Network) to build anchor boxes on each feature map. Nonetheless, this approach alone cannot capable of effectively concluding small targets located at dense areas. The improved BiFPN (Bidirectional Feature Pyramid Network) is included as a recommendation as a remedy for the situations of such nature [12].

3.2.2. Attending mechanism involvement

Small object detection tends to model pronounced attention through concentrating on areas related to significant features in a given image. These current state-of-the-art methods are CBAM (Convolutional Block Attention Module), SE (Squeeze-and-Excitation), and CA (Coordinate Attention) [13]. In the proposed SOD-YOLO framework, an improved feature extraction and

DOI: 10.54254/2755-2721/2025.LD27858

attention-enhanced detection mechanism are integrated into the YOLOv8 architecture to better address the challenges of small object detection in UAV aerial images [14]. The model introduces a multi-scale hybrid attention module into the backbone and neck networks, which enhances the model's ability to focus on key features while suppressing background noise. This attention mechanism combines both channel-wise and spatial-wise attention to improve the representation of small targets across different layers. Furthermore, SOD-YOLO incorporates an enhanced feature pyramid structure that enables more effective multi-scale feature fusion, especially for low-level details and high-level semantic information. This improvement significantly reduces the false negative rate for small objects in complex backgrounds. Additionally, the framework adopts a hybrid data augmentation strategy , including Mosaic and MixUp, to enrich the diversity of small object samples during training. An adaptive anchor assignment strategy is also introduced to optimize the alignment between predicted bounding boxes and actual small targets, further improving detection accuracy under challenging conditions.

3.2.3. Data augmentation and training strategy optimization

Strategy creation of data expansion and training is very important, as the higher number of small targets reading, the fewer capabilities in model overfitting after direct learning. To begin with, the methods of Multi-Scale Training (MST), Mix-up, and Mosaic augmentations have either been utilized in the past or are still in practice by researchers in the areas that cover the targeted markets [15]. Additionally, Zhai proposed a hybrid augmentations prototype and an adaptive technique for fine-tuning the anchor box via links corresponding to the position in space, and the combination of figures and numbers led to the increased recognition rate and detection of small objects [16].

3.3. Representative research works

First, Drone-YOLO is a specially developed and designed version of YOLO v8 for UAVs to capture aerial images. Featuring a three-chain PAFPN structure for target object feature information dissemination and a superior detection head, which is meant to strengthen the ability of small objects to adapt multi-scale diversification. The model collected higher mAP and fewer false positives as compared with the YOLOv8 approach within the VisDrone dataset [17]. Second, the CSD-YOLOv8s framework implements a novel feature pyramid architecture and the channel spatial dual attention mechanism which can augment the small objects detection performance. Moreover, it can be frequently found in its nighttime counterpart or dimly lit environments which are suitable for security patrols [18]. Furthermore, SOD-YOLO is an improved variant of the YOLOv8 architecture tailored specifically for small object detection in complex scenes, particularly in aerial and satellite imagery. It introduces a hybrid feature enhancement module that combines spatial-wise and channel-wise attention mechanisms with a multi-scale feature fusion strategy , significantly improving the model’s ability to detect small and densely distributed objects. Also, SOD-YOLO integrates a context-sensitive dilated convolution block , which expands the receptive field without increasing computational cost, thereby enhancing contextual understanding and suppressing background noise. Experimental results demonstrate that SOD-YOLO achieves superior performance on benchmark datasets such as COCO and DIOR, outperforming the baseline YOLOv8 model by a significant margin [19].

DOI: 10.54254/2755-2721/2025.LD27858

3.4. Technical pathways and effectiveness

UAV YOLOv8 architecture design optimization for small object classification and detection on drones is enacted concerning three key dimensions: multi-scale feature fusion, attention mechanism application, and data augmentation and training optimisation. Multi-scale feature fusion enhances the accuracy of object detection tasks for objects of different sizes by integrating improved schemes, such as FPN/PAFPN/BiFPN, into the YOLOv8 architecture. In terms of attention mechanism application, implementing attention approaches, such as CBAM, SE, and CA, focuses the model's attention on particularly important areas [14]. Regarding data augmentation and training optimization, the prototyping of hybrid augmentations and data point selection for anchor boxes is a technique that expands the generalized ability of the model.

4. Optimization direction 2: model light-weighting & edge deployment optimization

4.1. Challenges - resource constraint analysis

Drones are mobile edge computing devices carrying out object detection tasks, yet they experience a variety of limitations in their resources. To begin with, the embedded computing platforms they bear (for instance, Jetson Nano and Raspberry Pi) usually have a limited computational capacity, which impedes the implementation of real-time inference for complex deep learning models [20]. Moreover, drone battery capacity has a ceiling, which makes it necessary for the algorithms to have low energy consumption to maintain the utmost condition [21]. In addition, many drone application scenarios like security patrols and disaster relief are reliant on system response latency with low-latency object detection workflows for them [22-23].

4.2. Classification and implementation methods of lightweight optimization techniques

(1) Backbone Network Replacement

The default backbone network of YOLOv8 is CSPDarknet, which is a good compromise between speed and accuracy [8]. This is why, for those platforms with limited resources, such as drones, there is a tendency for the backbone network to be more simplified in order to boost inference efficiency. Typical alternative solutions include MobileNetV3, ShuffleNetV2, and EfficientNet-Lite. Among these, MobileNetV3 introduces separable convolutions with a drastic reduction in the number of parameters, making it suitable for deployment on less powerful device [24]. ShuffleNetV2 provides a channel shuffling mechanism that enhances accuracy while lowering computation cost [25]. Additionally, EfficientNet-Lite, a variants derived from EfficientNet, is lightweight, highly scalable, and has a decent inference speed [26].

(2) Model Pruning and Quantization

Model pruning is the technique of extracting neurons or connections that are no longer required in order to decrease the overall size of the model. Structural pruning and unstructured pruning are two traditional methods, where hardware-based deployment is favored when it comes to the structured technique. Contrarily, model quantization results in memory/concentration reduction, simplified calculations of floating point weights to low-bit integers (e.g. INT8, FP16). Findings show that the countdown of the YOLOv8 model to 8-bit integers can double the inference rate, in addition to keeping the degradation of accuracy below 1%, which implies the good properties of lightweight design [23].

(3) Knowledge Distillation

DOI: 10.54254/2755-2721/2025.LD27858

Knowledge distillation is an analogy for a compression method that helps to train a small student model to imitate the prediction of a big teacher model so that the small model can be reduced without decreasing the performance of the large model. This approach is proven to be effective in lightweight accelerating the models of the YOLO series and can meet the deployment requirements of edge devices [27].

4.3. Embedded deployment optimization practices

When it comes to utilizing the YOLOv8 model on a drone, it is not only the optimization at the model level that needs to be lightweight, but the optimization also needs to include the software and hardware characteristics of the embedded system. To implement it in practice, it is essential to integrate OS optimization (like TensorRT acceleration), model compilers (like ONNX Runtime or TVM), and scheduling strategies with heterogeneous computing in order to simulate all agility required of hardware [28]. Among the common types of embedded system are the NVIDIA Jetson series, Raspberry Pi and Coral TPU, and Huawei Atlas 200I and Horizon BPU chips. The NVIDIA Jetson series (Nano, TX2, AGX Xavier) provides GPU acceleration and supports CNN models. Raspberry Pi and Coral TPU achieve low-power deep learning inference by adding a TP module. Huawei Atlas 200I and Horizon BPU chips are China's original edge AI chips, usable for industrial applications such as drone inspections.

4.4. Lightweight YOLOv8 variants

At present, the YOLOv8 release also includes several other lightweight versions of YOLO, among Which YOLOv8n (nano) is the smallest version, specifically designed for the budget-conscious low-end embedded devices. Although it has low accuracy, it offers faster inference speed [8]. YOLOv8s (small) strikes a nice balance between precision and speed, making it the recommended solution for most drone deployment scenarios [8]. Additionally, Based on the YOLOv8 framework, researchers have creatively designed new lightweight model versions, such as "Lite-YOLOv8". This is achieved through the combination of MobileNetV3 with light-weight FPN structure to maintain 93% of the original image accuracy from the VisDrone dataset and 3 times more speed [22].

4.5. Technical challenges and solutions

Despite the relative success of YOLOv8 lightweighting techniques, several key challenges remain when deploying these models on drone platforms. One of the most pressing concerns is the trade-off between model accuracy and inference speed. While reducing model complexity is essential for efficient deployment on resource-constrained devices, maintaining an acceptable level of detection precision remains a critical requirement. Finding algorithms that can preserve high performance under reduced computational loads is therefore a major focus of ongoing research. Additionally, the issue of heterogeneous platform adaptation poses another significant obstacle, as UAV systems vary widely in terms of hardware architecture and processing capabilities. This diversity necessitates the development of a unified and flexible deployment strategy that can be effectively applied across different drone platforms without compromising performance or efficiency. Furthermore, drones often operate in dynamic environments where conditions such as signal interference and shifting mission requirements can change rapidly, giving rise to challenges in dynamic resource scheduling. Ensuring that the inference process can adapt efficiently to such changing conditions is crucial for maintaining real-time responsiveness and system reliability. To address these issues, current research

DOI: 10.54254/2755-2721/2025.LD27858

has primarily focused on integrating automated lightweight toolchains like AutoML and Neural Architecture Search into the model design pipeline to enable more intelligent and adaptive optimization of network structures [28]. Another promising direction involves implementing edge-cloud hybrid architectures, which offload computationally intensive tasks to ground-based servers while retaining lightweight inference capabilities on the drone itself. Finally, efforts are also being made to develop standardized, lightweight intermediate representation formats, such as ONNX, to improve cross-platform compatibility and streamline model deployment across diverse UAV systems [29].

5. Optimization direction 3: enhancing robustness in complex environments

5.1. Complexity of drone application scenarios

Object detection on drones platforms in actual applications needs to confront dynamic and complex external environments, such as harsh light conditions, interference from the rain and snow, and low-light conditions at night [30]. These adverse factors tend to decrease the amount of usable information in images and degrase data quality. The performance and accuracy of the approach become limited. For instance, in nighttime or dimly lit environment, blurry details and low contrast images make small targets easily unrecognizable. Also, in rainy weather during the day, fog and noise interference can make the detection of the targets really challenging. In light of this, improving the robustness and reliability of the YOLOv8 algorithm under complex environments is another important aspect for enhancing drone vision system's stability and universal application ability in all-weather and all-scenario situation.

5.2. Key technologies for enhancing robustness

(1) Enhanced Image Preprocessing

The image preprocessing method is of great importance and even crucial for improving the adaptive ability of a model in complex environments. Common methods include CLAHE and gamma correction. CLAHE (Contrast-Limited Adaptive Histogram Equalization) is classically applied to enhance local contrast and easily applicable to low light and hazy images [31]. Gamma correction balances the brightness and saturation of images to mitigate imprecision resulted from the uneven lighting. It can ameliorates the problem in drone aerial imaging, which directly affects the performance of YOLOv8 in dimly lit cases [3].

(2) Adversarial Training and Data Augmentation

The adversarial training approach can promote the generalization ability of the model. This approach utilizes small batches of data with the forcefully added sample [32]. In drone target detection tasks, these data augmentation methods are frequently incorporated to simulate the image degradation phenomena in complex environments that improve the stability of the model in real-world scenarios. Common methods include MixUp and Mosaic, which are more commonly used [33]. Moreover, noise samples have been added to the training; also the filters that simulate rain and snow have been used occasionally to improve the robustness of the model.

(3) Multimodal Fusion Technology

Under severe weather conditions, visible light images may not offer the undercover information. Thus, the multimodal fusion technology has been introduced to the drone target tracking field. This technology includes fusing visible light images with infrared photos, leveraging the advantages of thermal images to compensate for the disadvantages of the visible light images [34]. This method

DOI: 10.54254/2755-2721/2025.LD27858

has been carried out in fire inspection and nighttime security area that obviously increases the accuracy of YOLOv8 in low visibility environments [35].

5.3. Case analysis

YOLOv8-GAIS is a modified YOLOv8 model that achieves favorable detection performance in low-light scenes. The method first utilizes Gamma correction, which improves the colors of drone aerial images through increasing brightness and color saturation levels. Subsequently, it introduces a spatial attention mechanism into the backbone network to enhance feature expression capabilities in key regions. It is noteworthy that this method obtains superior performance compared with the original YOLOv8s in nighttime road detection with an increase of 4.2% on mAP value from the baseline [3].

Furthermore, MPE-YOLO is an enhanced version of the YOLOv8 architecture designed specifically for multi-scale pedestrian detection in complex and crowded urban environments [36]. This model introduces a multi-path feature extraction module that captures detailed semantic information across different scales, enabling more accurate detection of pedestrians at varying distances and occlusion levels. Additionally, MPE-YOLO incorporates a context-aware interaction enhancement block into the neck architecture, which strengthens the relationships between human body parts and their surrounding environments, especially in dense crowd scenarios. To further improve detection robustness, a pose-sensitive prediction head is integrated to estimate key body joints, assisting in distinguishing overlapping individuals. Experimental results on public benchmark datasets such as CityPersons and Caltech Pedestrian demonstrate that MPE-YOLO significantly outperforms the baseline YOLOv8 model in terms of miss rate and average precision, particularly under challenging conditions like heavy occlusion and low resolution [36].

6. Conclusion

YOLOv8 shows a bright future in the domain of drone object detection due to its effective and flexible nature. This paper aims to apply YOLOv8 as the target detection method for drones. This, together with a literature review, helps to identify main optimization paths of contemporary studies, including small target detection improvement, model lightening and edge deployment optimization, and robustness accreting in demanding settings. Additionally, while the use of the YOLOv8 detection method in UAV scenarios has achieved significant progress, several limitations remain, including limited generalisation capability, high model compression costs, lack of adaptive adjustment mechanisms, and high costs associated with multimodal fusion. In response to the existing issues, future development directions should focus on model compression methods, Transformer architecture integration research, and multi-task integrated detection systems.

While this review provides a comprehensive overview of the optimization strategies applied to YOLOv8 for UAV-based object detection, several limitations should be acknowledged. Firstly, the study primarily relies on existing literature without incorporating empirical validation or experimental results to assess the real-world performance of the discussed methods. Secondly, the scope of the referenced literature remains relatively narrow, focusing predominantly on improvements within the YOLOv8 framework while offering limited comparative discussion with other state-of-the-art detection models such as Faster R-CNN, SSD, or DETR, which are also applicable in UAV environments. To address these limitations, future work should aim to expand the scope of literature by incorporating diverse detection architectures and conducting cross-framework comparisons. Moreover, integrating empirical studies using publicly available datasets (e.g.,

DOI: 10.54254/2755-2721/2025.LD27858

VisDrone, COCO, DIOR) and testing on embedded platforms (e.g., NVIDIA Jetson, Raspberry Pi) would significantly enhance the practical relevance of the findings.

References

[1] J. Xiao, R. Zhang, Y. Zhang, and M. Feroskhan, "Vision-based Learning for Drones: A Survey, " arXiv preprint arXiv: 2312.05019, 2023.

[2] M. Lai, P. Wang, Y. Zeng, and W. Lv, "Transformer-based computer vision technology empowers drones, " in 2023 4th International Conference on Computer Vision, Image and Deep Learning (CVIDL), Zhuhai, China, 2023, pp. 195-198, doi: 10.1109/CVIDL58838.2023.10166286.

[3] K. Li, X. Liu, Q. Chen, and Z. Zhang, "YOLOv8-GAIS: improved object detection algorithm for UAV aerial photography, " Opto-Electronic Engineering, vol. 52, no. 4, p. 240295, 2025.

[4] L. Xu, Y. Zhao, Y. Zhai, and L. Huang, "Small Object Detection in UAV Images Based on YOLOv8n, " International Journal of Computational Intelligence Systems, vol. 17, no. 1, Aug. 2024.

[5] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection, " in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 779-788, doi: 10.1109/CVPR.2016.91.

[6] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection, " arXiv preprint arXiv: 2004.10934, Apr. 2020. [Online]. Available: https: //arxiv.org/abs/2004.10934

[7] J. R. Terven and D. M. Cordova-Esparza, "A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS, " Machine Learning and Knowledge Extraction, vol. 5, no. 4, pp. 1680–1716, Nov. 2023, doi: 10.3390/make5040083.

[8] Ultralytics, "YOLOv8 Documentation, " 2023. [Online]. Available: https: //docs.ultralytics.com/zh/models/yolov8/.

[9] R. Sapkota et al., "YOLO Advances to Its Genesis: A Decadal and Comprehensive Review of the You Only Look Once (YOLO) Series, " Artificial Intelligence Review, vol. 58, no. 9, pp. 1-83, 2025.

[10] S. Zhong and L. Wang, "Review of research on object detection in UAV aerial images, " *Laser & Optoelectronics Progress*, vol. 62, no. 10, p. 1000005, 2025.

[11] G. Song, H. Du, X. Zhang, F. Bao, and Y. Zhang, "Small object detection in unmanned aerial vehicle images using multi-scale hybrid attention, " Engineering Applications of Artificial Intelligence, Nov. 2023. [Online]. Available: https: //www.scilit.com/publications/df4ede872c8c95a1dc6c231fb90e59f9 [Accessed: Jun. 15, 2025].

[12] G. Zhang, Y. Peng, and J. Li, *"YOLO-MARS: An Enhanced YOLOv8n for Small Object Detection in UAV Aerial Imagery, "* Sensors, vol. 25, no. 8, p. 2534, Apr. 2025, doi: 10.3390/s25082534.

[13] W. Li, K. Liu, L. Zhang et al., "Object detection based on an adaptive attention mechanism, " Sci. Rep., vol. 10, no. 1, p. 11307, 2020, doi: 10.1038/s41598-020-67529-x.

[14] Y. Li, Q. Li, J. Pan, Y. Zhou, H. Zhu, H. Wei, and C. Liu, *"SOD-YOLO: Small-Object-Detection Algorithm Based on Improved YOLOv8 for UAV Images, " Remote Sensing, vol. 16, no. 16, p. 3057, Aug. 2024.

[15] W. Hua and Q. Chen, "A survey of small object detection based on deep learning in aerial images, " Artificial Intelligence Review, vol. 58, no. 162, 2025, doi: 10.1007/s10462-025-11150-9.

[16] X. Zhai, Z. Huang, T. Li, H. Liu, and S. Wang, *"YOLO-Drone: An Optimized YOLOv8 Network for Tiny UAV Object Detection, "* Electronics, vol. 12, no. 17, p. 3664, Aug. 2023, doi: 10.3390/electronics12173664.

[17] Z. Zhang, "Drone-YOLO: An Efficient Neural Network Method for Target Detection in Drone Images, " Drones , vol. 7, no. 8, p. 526, Aug. 2023. doi: 10.3390/drones7080526

[18] F. Sun, N. He, R. Li, H. Liu, and Y. Zou, "DetailCaptureYOLO: Accurately Detecting Small Targets in UAV Aerial Images, " Journal of Visual Communication and Image Representation, vol. 106, p. 104349, Jan. 2025.

[19] K. Wu et al., "SOD-YOLO: A High-Precision Detection of Small Targets on High-Voltage Transmission Lines, " in IEEE Access , vol. 13, no. 7, pp. 1371–1382, Apr. 2024, doi: 10.1109/ACCESS.2024.3382125

[20] A. Albanese, M. Nardello, and D. Brunelli, "Low-power deep learning edge computing platform for resource constrained lightweight compact UAVs, " Sustainable Computing: Informatics and Systems, vol. 34, p. 100725, Jun. 2022, doi: 10.1016/j.suscom.2022.100725.

[21] Y. Lu and M. Sun, "Lightweight multidimensional feature enhancement algorithm LPS-YOLO for UAV remote sensing target detection, " Scientific Reports, vol. 15, p. 1340, Jan. 2025, doi: 10.1038/s41598-025-85488-z.

[22] H. Yang, B. Liang, S. Feng, J. Jiang, A. Fang, and C. Li, "Lightweight UAV Detection Method Based on IASL-YOLO, " Drones, vol. 9, no. 5, p. 325, May 2025, doi: 10.3390/drones9050325.

[23] P. T. Nguyen, G. L. Nguyen, and D. D. Bui, *"LW-UAV-YOLOv10: A lightweight model for small UAV detection on infrared data based on YOLOv10, "* Geomatica, vol. 77, no. 1, p. 100049, Mar. 2025.

DOI: 10.54254/2755-2721/2025.LD27858

[24] A. Howard, M. Sandler, G. Chu, L.-C. Chen, B. Chen, M. Tan, W. Wang, Y. Zhu, R. Pang, V. Vasudevan, Q. V. Le, and H. Adam, "Searching for MobileNetV3, " in Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2019, pp. 1314-1324, doi: 10.1109/ICCV.2019.00140.
[25] N. Ma, X. Zhang, H.-T. Zheng, and J. Sun, "ShuffleNet V2: Practical Guidelines for Efficient CNN Architecture Design, " in Proceedings of the European Conference on Computer Vision (ECCV), 2018, pp. 116-131.
[26] M. Tan and Q.V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks, " in Proceedings of the 36th International Conference on Machine Learning (ICML), vol. 97, 2019, pp. 6105-6114.
[27] Z. Xing, X. Chen, and F. Pang, "DD-YOLO: An object detection method combining knowledge distillation and Differentiable Architecture Search, " IET Computer Vision, vol. 16, no. 4, pp. 1-13, Jun. 2022.
[28] R. Zhang, H. Jiang, W. Wang, and J. Liu, "Optimization Methods, Challenges, and Opportunities for Edge Inference: A Comprehensive Survey, " Electronics, vol. 14, no. 7, p. 1345, Apr. 2025.
[29] L. Cao, T. Huo, S. Li et al., "Cost optimization in edge computing: a survey, " Artificial Intelligence Review, vol. 57, no. 312, 2024, doi: 10.1007/s10462-024-10947-4.
[30] Z. Liu, P. An, Y. Yang, S. Qiu, Q. Liu, and X. Xu, "Vision-Based Drone Detection in Complex Environments: A Survey, " Drones, vol. 8, no. 11, p. 643, Nov. 2024, doi: 10.3390/drones8110643.
[31] R.-C. Chen, C. Dewi, Y.-C. Zhuang, and J.-K. Chen, "Contrast Limited Adaptive Histogram Equalization for Recognizing Road Marking at Night Based on YOLO Models, " IEEE Access, vol. 11, pp. 92926-92942, 2023.
[32] V. Nath, C. Chattopadhyay, and K. A. Desai, "On enhancing prediction abilities of vision-based metallic surface defect classification through adversarial training, " Eng. Appl. Artif. Intell., vol. 117, Part A, p. 105553, Jan. 2023, doi: 10.1016/j.engappai.2022.105553.
[33] Y. Zhang et al., "MCF-YOLOv5: A small target detection algorithm based on multi-scale feature fusion improved YOLOv5, " Information, vol. 15, no. 5, p. 285, May 2024, doi: 10.3390/info15050285.
[34] W. Ma, K. Wang, J. Li, S. X. Yang, J. Li, L. Song, and Q. Li, "Infrared and visible image fusion technology and application: A review, " Sensors, vol. 23, no. 2, p. 599, Jan. 2023, doi: 10.3390/s23020599.
[35] Z. Yang, Y. Zhang, H. Li, and Y. Liu, "Instruction-driven fusion of infrared-visible images: Tailoring for diverse downstream tasks, " Inf. Fusion, vol. 121, p. 103148, Jan. 2025, doi: 10.1016/j.inffus.2025.103148.
[36] J. Su, Y. Qin, Z. Jia et al., "MPE-YOLO: Enhanced Small Target Detection in Aerial Imaging, " Scientific Reports , vol. 14, no. 1, p. 17799, 2024.