International Journal of Science and Engineering Applications
Volume 15-Issue 07, 88 – 94, 2026, ISSN:- 2319 - 7560
DOI: 10.7753/IJSEA1507.1015

# A Lightweight YOLOv8n UAV Small Object Detection Method Based on Detection Layer Reconstruction

Wang Xinyu
School of Geophysics and Petroleum Resources, Yangtze University, Wuhan, China

**Abstract:** Aiming at the problems of small object scale, complex background, dense target distribution and redundant detection branches in Unmanned Aerial Vehicle (UAV) aerial images, this paper proposes a lightweight UAV small object detection method based on YOLOv8n detection layer reconstruction. Instead of stacking complex enhancement modules, the proposed method adjusts the detection structure according to the matching relationship between target scale distribution and detection layers. A high-resolution P2 detection layer is introduced to strengthen shallow detail representation for small objects, while the P5 large-object detection layer is removed to reduce redundant parameters and computation. Thus, a lightweight P2-P3-P4 detection structure named YOLOv8-P2-woP5 is constructed. Experiments on the VisDrone2019 dataset show that the proposed model reduces parameters from 3.01M to 1.38M while maintaining an mAP50 of 0.332. Additional GhostConv experiments indicate that excessive compression may weaken small-object feature representation and reduce accuracy. The results demonstrate that the proposed model achieves a better balance between detection performance and model complexity.

**Keywords:** YOLOv8n; Unmanned Aerial Vehicle (UAV) image; small object detection; detection layer reconstruction; lightweight model; VisDrone2019

## 1. Introduction

With the development of UAV platforms and deep learning, UAV image object detection has been widely applied in traffic monitoring, urban management, public safety and disaster rescue**Error! Unknown switch argument.**. UAVs can efficiently collect large-scale ground information, but aerial images often contain small objects, dense distributions, occlusion and complex backgrounds. After repeated downsampling in deep networks, edge, texture and location details of small objects are easily weakened, resulting in missed detections and inaccurate localization.

YOLOv8n is a lightweight member of the YOLO series and is suitable for edge deployment because of its small parameter scale and fast inference speed**Error! Unknown switch argument.**. However, the original P3, P4 and P5 detection structure may be insufficient for datasets such as VisDrone2019, where small and medium objects dominate. The P3 layer may still lack enough spatial resolution for extremely small objects, while the P5 layer mainly serves large objects and may introduce redundant computation.

Based on this observation, this paper reconstructs the detection layers of YOLOv8n by introducing a P2 small-object detection layer and removing the P5 large-object detection layer. The goal of this study is not to simply pursue higher accuracy, but to reduce model parameters while keeping detection performance stable, thereby achieving a balance between UAV small object detection and lightweight deployment.

The main contributions are summarized as follows:

(1) a YOLOv8n lightweight detection method based on detection layer reconstruction is proposed;

(2) a P2-P3-P4 detection structure is constructed to enhance small-object representation and reduce redundant large-object branches;

(3)ablation experiments verify that YOLOv8-P2-woP5 significantly reduces parameters while maintaining mAP50, and further lightweight experiments show that excessive compression may damage small-object detection performance.

## 2. Related Work

### 2.1 Object Detection and UAV Small Object Detection

Object detection aims to identify object categories and locate their bounding boxes in images. Existing methods are commonly divided into two-stage and one-stage detectors. Two-stage methods, such as Faster R-CNN, first generate candidate regions and then perform classification and regression **Error! Unknown switch argument.**. They usually achieve high accuracy but have complex structures and slower inference. One-stage methods, such as SSD, RetinaNet and YOLO, directly predict categories and boxes on feature maps and are more suitable for real-time detection **Error! Unknown switch argument.**. In addition, feature pyramid representation, path aggregation and efficient feature fusion have become important strategies for multi-scale object detection.

YOLO-based detectors are widely used because of their concise structure, fast speed and easy deployment . YOLOv8n is a lightweight version suitable for UAV and edge-device scenarios **Error! Unknown switch argument.**. Nevertheless, UAV images differ from common natural scenes because objects such as pedestrians, vehicles, bicycles and motorcycles usually occupy only a small number of pixels. Dense object distribution, occlusion and complex backgrounds further increase detection difficulty **Error! Unknown switch argument.**. Therefore, the detection layer design should be adjusted according to the target scale distribution of UAV datasets **Error! Unknown switch argument.**. Recent improved YOLOv8 and SOD-YOLO studies have also focused on UAV or remote sensing small-object detection through feature enhancement, lightweight design and detection structure optimization .

### 2.2 Lightweight Object Detection Networks

As object detection is increasingly deployed on UAVs, mobile terminals and edge devices, lightweight model design has become important. Common methods include pruning, quantization, knowledge distillation, lightweight convolution, attention mechanisms and detection layer pruning **Error! Unknown switch argument.**. Lightweight convolutions such as depthwise separable convolution, group convolution and GhostConv can reduce computation, but excessive compression may weaken shallow edge, texture and location information, which is important for small objects.Attention modules can enhance key features, but they often introduce additional parameters and their effectiveness depends on the dataset and network structure **Error! Unknown switch argument.**. Adaptive feature fusion methods can also improve multi-scale representation, but they may increase structural complexity. For UAV small object detection, lightweight design should not only minimize parameters, but should also retain sufficient detail representation. Therefore, this paper adopts a detection layer reconstruction strategy by changing the original P3-P4-P5 structure to P2-P3-P4, so as to reduce redundant branches while preserving small-object representation ability.

# 3. Methodology

## 3.1 Overall Structure of the Improved YOLOv8n Model

YOLOv8n mainly consists of a Backbone, Neck and Head. The Backbone extracts image features, the Neck fuses multi-scale features, and the Head performs classification and bounding box regression. The original YOLOv8n generally uses P3, P4 and P5 detection layers for small, medium and large objects, respectively. This structure is effective for general object detection, but it is not fully matched with the scale distribution of VisDrone2019, where small and medium objects are dominant.

To address this issue, a P2 detection layer is added to exploit higher-resolution shallow features, while the P5 detection layer is removed to reduce the redundant large-object branch. The resulting structure changes the detection layers from P3-P4-P5 to P2-P3-P4, which is consistent with the small-object-oriented feature enhancement idea used in recent UAV detection studies **Error! Unknown switch argument.**. In this design, P2 focuses on extremely small and small objects, P3 handles small and part of medium objects, and P4 mainly detects medium objects. The proposed model is named YOLOv8-P2-woP5.

## 3.2 P2 Detection Head and P5 Layer Removal

Small objects in UAV images have low pixel proportions, and their details can be easily lost after multiple downsampling operations. Therefore, the P2 detection head is introduced to perform prediction on higher-resolution feature maps, which improves the representation of distant pedestrians, small vehicles and non-motorized vehicles, and helps reduce missed detections and localization errors.

However, directly adding P2 while keeping P3, P4 and P5 increases computation. To control model complexity, the P5 detection layer, which mainly targets large objects, is removed. Since large objects are relatively limited in VisDrone2019, removing P5 can reduce redundant parameters while retaining the benefit of P2 for small objects **Error! Unknown switch argument.**. In addition, Ghost-LP-YOLOv8 is designed as a supplementary experiment to analyze the effect of lightweight convolution **Error! Unknown switch argument.**.Although GhostConv further reduces parameters and computation, the experimental results show a clear decline in mAP50, indicating that enough shallow detail representation should be preserved in UAV small object detection.

# 4. Experiment & Result Analysis

## 4.1 Dataset Description and Experimental Setup

The VisDrone2019 dataset is used to verify the proposed method . This dataset is collected by UAV platforms and contains multiple object categories such as pedestrians, vehicles, bicycles and motorcycles. It has the characteristics of small object scale, dense distribution, complex background and frequent occlusion, which makes it suitable for evaluating UAV small object detection algorithms. A representative UAV road scene from the dataset is shown in Figure 1.

Figure 1. Urban road scene with pedestrians and vehicles.

The experiments were conducted using the YOLOv8 framework. YOLOv8n was selected as the baseline model, and all comparison models were trained and evaluated on the VisDrone2019 dataset under the same experimental settings. To ensure fair comparison, only the network structure was changed in different experimental groups, while the dataset, input size, training strategy and evaluation metrics were kept consistent. The main experimental settings are summarized in Table 1.

**Table 1. Main experimental settings.**

| Group | Model | Main modification | Purpose |
|---|---|---|---|
| A | YOLOv8n | Original model | Baseline |
| B | YOLOv8-P2 | Add P2 detection layer | Enhance small-object representation |
| C | YOLOv8-P2-woP5 | Add P2 and remove P5 | Proposed main model |
| D | YOLOv8-P2-woP5-Sandwich | Add Sandwich structure | Structural comparison |
| E | YOLOv8-P2-woP5-Sandwich-CBAM | Add CBAM attention module | Attention comparison |
| F | Ghost-LP-YOLOv8 | Replace part of convolution with GhostConv | Further lightweight comparison |

## 4.2 Training Settings and Evaluation Metrics

To ensure comparability, all models are trained and validated with basically the same data configuration, input size, training epochs, batch size and device settings. Only the model structure is changed. The evaluation metrics include Precision, Recall, mAP50, mAP50-95, Params and GFLOPs. mAP50 reflects the average precision at an IoU threshold of 0.5, while mAP50-95 evaluates model performance under multiple IoU thresholds. Params and GFLOPs are used to measure model storage scale and computational cost. The experiments focus on whether adding P2 improves small-object detection, whether removing P5 can reduce model size while maintaining accuracy, and whether further lightweight compression affects small-object representation.

## 4.3 Result Comparison

Six groups of experiments are conducted on VisDrone2019 to verify the effectiveness of detection layer reconstruction. The results are shown in Table 2.

International Journal of Science and Engineering Applications
Volume 15-Issue 07, 88 – 94, 2026, ISSN:- 2319 - 7560
DOI: 10.7753/IJSEA1507.1015

**Table 2. Comparison results of different models.**

| Group | Model | Params | GFLOPs | P | R | mAP50 | mAP50-95 |
|---|---|---|---|---|---|---|---|
| A | YOLOv8n | 3.01M | 8.1 | 0.425 | 0.335 | 0.332 | 0.195 |
| B | YOLOv8-P2 | 2.92M | 12.2 | 0.449 | 0.330 | 0.334 | 0.195 |
| C | YOLOv8-P2-woP5 | 1.38M | 10.9 | 0.437 | 0.329 | 0.332 | 0.194 |
| D | YOLOv8-P2-woP5-Sandwich | 1.46M | 11.4 | 0.423 | 0.335 | 0.329 | 0.191 |
| E | YOLOv8-P2-woP5-Sandwich-CBAM | 1.54M | 11.5 | 0.423 | 0.339 | 0.332 | 0.194 |
| F | Ghost-LP-YOLOv8 | 1.35M | 10.6 | 0.430 | 0.337 | 0.315 | 0.182 |

The training curves of YOLOv8-P2-woP5 are shown in Figure 2.

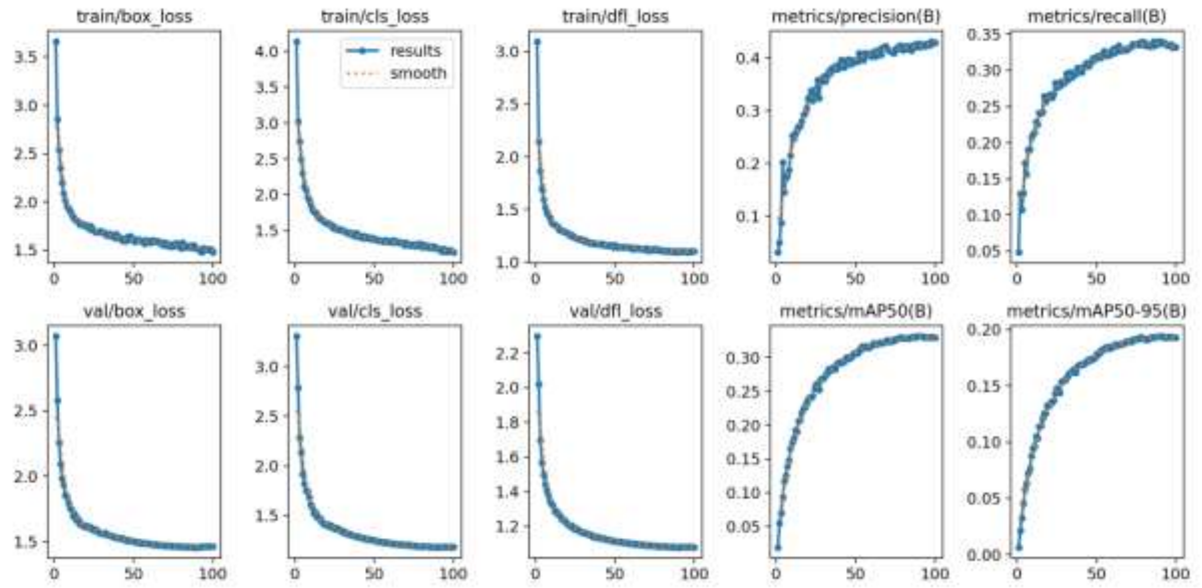


Figure 2. Training curves of YOLOv8-P2-woP5.

The baseline YOLOv8n contains 3.01M parameters and achieves an mAP50 of 0.332. After adding the P2 detection layer, YOLOv8-P2 improves mAP50 to 0.334, indicating that high-resolution features are helpful for small-object detection. However, GFLOPs increase from 8.1 to 12.2, showing that simply adding P2 brings obvious computational overhead. The training curves show that box loss, cls loss and dfl loss decrease steadily, while Precision, Recall, mAP50 and mAP50-95 gradually stabilize.

The precision-recall curve of YOLOv8-P2-woP5 is shown in Figure 3.

International Journal of Science and Engineering Applications
Volume 15-Issue 07, 88 – 94, 2026, ISSN:- 2319 - 7560
DOI: 10.7753/IJSEA1507.1015

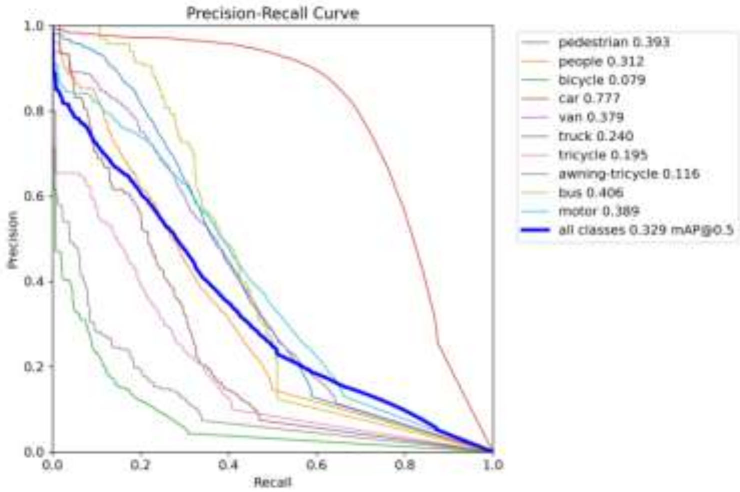


Figure 3. Precision-recall curve of YOLOv8-P2-woP5.

As shown in Figure 3, the precision-recall curve indicates category-level differences. YOLOv8-P2-woP5 removes the P5 layer while keeping P2. Its parameters decrease from 3.01M to 1.38M, while mAP50 remains 0.332 and mAP50-95 only decreases from 0.195 to 0.194. This indicates that the P5 large-object layer is partly redundant in VisDrone2019. Vehicle classes perform relatively well, while small or easily occluded categories such as bicycles and tricycles are still difficult.
The recall-confidence curve of YOLOv8-P2-woP5 is shown in Figure 4.

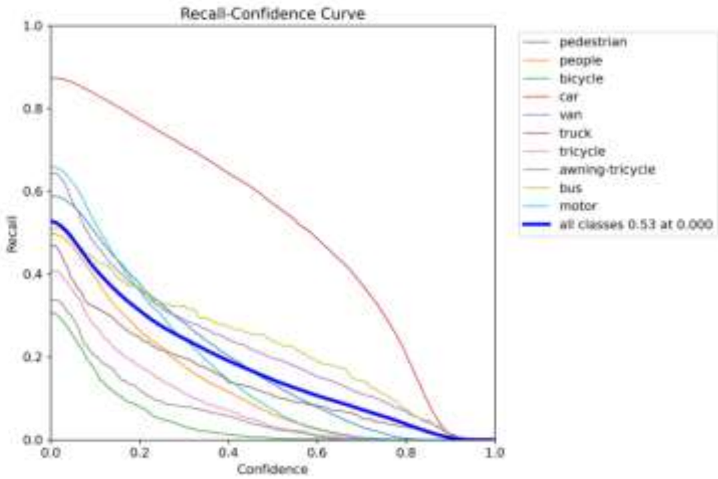


Figure 4. Recall-confidence curve of YOLOv8-P2-woP5.

As shown in Figure 4, the recall-confidence curve further reflects the instability of difficult small-object categories. The Sandwich and CBAM variants do not clearly outperform the proposed model. In particular, the CBAM variant obtains the same mAP50 as YOLOv8-P2-woP5 but introduces more parameters and computation **Error! Unknown switch argument.**. Ghost-LP-YOLOv8 further reduces parameters to 1.35M and GFLOPs to 10.6, but mAP50 drops to 0.315 and mAP50-95 drops to 0.182. This result suggests that excessive lightweight compression can weaken small-object detail representation **Error! Unknown switch argument.**. Overall, YOLOv8-P2-woP5 achieves a better balance among accuracy, parameter scale and computational cost.

## 4.4 Analysis of Detection Layer Reconstruction and Lightweight Effect

The experimental results show that the P2 detection layer improves the perception of small objects, but its high-resolution feature map also increases computation. Therefore, adding P2 alone is not the optimal solution **Error! Unknown switch argument.**. Removing P5 after adding P2 is more consistent with the target scale distribution of VisDrone2019, because this dataset contains many small and medium objects but relatively few large objects.
Compared with the original YOLOv8n, YOLOv8-P2-woP5 greatly reduces parameters while maintaining mAP50. Compared with Sandwich and CBAM extensions, the proposed structure is simpler, lighter and more stable **Error! Unknown switch argument.**. The GhostConv experiment further confirms that lightweight design is not equivalent to continuously reducing parameters. Small object detection still relies on shallow edges, textures and spatial details. Thus, a reasonable lightweight detector should consider both target scale distribution and feature representation capacity.

## 5. Conclusion

This paper proposes a lightweight YOLOv8n UAV small object detection method based on detection layer reconstruction. A P2 small-object detection layer is introduced to enhance shallow high-resolution feature representation, and the P5 large-object detection layer is removed to reduce redundant branches **Error! Unknown switch argument.**. The final YOLOv8-P2-woP5 model constructs a P2-P3-P4 detection structure that is more consistent with the scale distribution of UAV images.

Experimental results show that adding P2 alone slightly improves mAP50 from 0.332 to 0.334 but significantly increases GFLOPs. After removing P5, YOLOv8-P2-woP5 reduces parameters from 3.01M to 1.38M while maintaining an mAP50 of 0.332. Compared with Sandwich, CBAM and GhostConv extensions, the proposed model provides more stable overall performance **Error! Unknown switch argument.**. In particular, GhostConv further reduces model size but causes an mAP50 drop to 0.315, which indicates that small object detection should not pursue model compression blindly.

The proposed model can reduce model scale while basically maintaining detection accuracy, showing certain value for lightweight UAV deployment. However, the accuracy improvement over the original YOLOv8n is not obvious, and the generalization ability needs to be verified on more UAV datasets and real scenarios. Future work will optimize shallow feature fusion, supplement FPS and edge-device deployment experiments, and conduct more detailed category-level error analysis for extremely small objects.

## 6. References

[1] Zhu P, Wen L, Bian X, Ling H, Hu Q. Vision meets drones: a challenge. arXiv:1804.07437 [Preprint]. 2018. doi:10.48550/arXiv.1804.07437.

[2] Redmon J, Divvala S, Girshick R, Farhadi A. You only look once: unified, real-time object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2016. p. 779-788. doi:10.1109/CVPR.2016.91.

[3] Ren S, He K, Girshick R, Sun J. Faster R-CNN: towards real-time object detection with region proposal networks. IEEE Trans Pattern Anal Mach Intell. 2017;39(6):1137-1149. doi:10.1109/TPAMI.2016.2577031.

[4] Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu CY, et al. SSD: single shot MultiBox detector. In: Computer Vision - ECCV 2016. Cham: Springer; 2016. p. 21-37. doi:10.1007/978-3-319-46448-0_2.

[5] Li Y, Fan Q, Huang H, Han Z, Gu Q. A modified YOLOv8 detection network for UAV aerial image recognition. Drones. 2023;7(5):304. doi:10.3390/drones7050304.

[6] Wang G, Chen Y, An P, Hong H, Hu J, Huang T. UAV-YOLOv8: a small-object-detection model based on improved YOLOv8 for UAV aerial photography scenarios. Sensors. 2023;23(16):7190. doi:10.3390/s23167190.

[7] Yi H, Liu B, Zhao B, Liu E. Small object detection algorithm based on improved YOLOv8 for remote sensing. IEEE J Sel Top Appl Earth Obs Remote Sens. 2024;17:1734-1747. doi:10.1109/JSTARS.2023.3339235.

[8] Wang P, Zhao J. SOD-YOLO: enhancing YOLO-based detection of small objects in UAV imagery. arXiv:2507.12727 [Preprint]. 2025. doi:10.48550/arXiv.2507.12727.

[9] Liu S, Huang D, Wang Y. Learning spatial fusion for single-shot object detection. arXiv:1911.09516 [Preprint]. 2019. doi:10.48550/arXiv.1911.09516.

[10] He Z, Cao L. SOD-YOLO: small object detection network for UAV aerial images. IEEJ Trans Electr Electron Eng. 2025;20(3):431-439. doi:10.1002/tee.24195.

[11] Lin TY, Dollár P, Girshick R, He K, Hariharan B, Belongie S. Feature pyramid networks for object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2017. p. 936-944. doi:10.1109/CVPR.2017.106.

[12] Han K, Wang Y, Tian Q, Guo J, Xu C, Xu C. GhostNet: more features from cheap operations. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition; 2020. p. 1580-1589. doi:10.1109/CVPR42600.2020.00165.

[13] Woo S, Park J, Lee JY, Kweon IS. CBAM: convolutional block attention module. In: Computer Vision - ECCV 2018. Cham: Springer; 2018. p. 3-19. doi:10.1007/978-3-030-01234-2_1.

[14] Sandler M, Howard A, Zhu M, Zhmoginov A, Chen LC. MobileNetV2: inverted residuals and linear bottlenecks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2018. p. 4510-4520. doi:10.1109/CVPR.2018.00474.

[15] Zhang X, Zhou X, Lin M, Sun J. ShuffleNet: an extremely efficient convolutional neural network for mobile devices. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition; 2018. p. 6848-6856. doi:10.1109/CVPR.2018.00716.

[16] Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, et al. MobileNets: efficient convolutional neural networks for mobile vision applications. arXiv:1704.04861 [Preprint]. 2017. doi:10.48550/arXiv.1704.04861.

[17] Liu S, Qi L, Qin H, Shi J, Jia J. Path aggregation network for instance segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition; 2018. p. 8759-8768. doi:10.1109/CVPR.2018.00913.

International Journal of Science and Engineering Applications
Volume 15-Issue 07, 88 – 94, 2026, ISSN:- 2319 - 7560
DOI: 10.7753/IJSEA1507.1015

[18] Tan M, Pang R, Le QV. EfficientDet: scalable and efficient object detection. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition; 2020. p. 10781-10790. doi:10.1109/CVPR42600.2020.01079.