Journal of Electronic Research and Application, 2025, Volume 9, Issue 3
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

An Improved Lightweight Pest Detection Method Based on YOLOv8

Liling Zhang*, **Xueqian Du**

College of Computer Science and Engineering, Chongqing Three Gorges University, Chongqing 404100, China

**Author to whom correspondence should be addressed.*

Abstract: This study systematically addresses the limitations of traditional pest detection methods and proposes an optimized version of the YOLOv8 object detection model. By integrating the GhostConv convolution module and the C3Ghost module, the Polarized Self-Attention (PSA) mechanism is incorporated to enhance the model's capacity for extracting pest features. Experimental results demonstrate that the improved YOLOv8 + Ghost + PSA model achieves outstanding performance in critical metrics such as precision, recall, and mean Average Precision (mAP), with a computational cost of only 5.3 GFLOPs, making it highly suitable for deployment in resource-constrained agricultural environments.

Keywords: Agricultural pest detection; YOLOv8; Lightweight; PSA attention mechanism

Online publication: 5 June, 2025

1. Introduction

Currently, agriculture faces diverse pest threats, which pose significant risks to crop yield and quality. Within the agricultural system, establishing an effective pest monitoring and forecasting mechanism is crucial for ensuring the yield and quality of agricultural products [1]. Traditional methods for identifying crop pests primarily rely on on-site inspections, measurements, and statistics, followed by judgments based on conventional experience. These methods are heavily reliant on human labor and exhibit certain limitations, often leading to issues such as missed or incorrect inspections.

With the rapid advancement of deep learning, research into agricultural pest detection using deep learning has achieved notable results. The application of YOLO series algorithms in the agricultural domain has also yielded significant outcomes [2]. The YOLOv8 algorithm inherits the advantages of the YOLO series, including fast speed and high precision, while introducing novel network architectures and optimization techniques to further enhance detection accuracy and efficiency. Despite its strengths, YOLOv8 still exhibits certain shortcomings. Current experiments have introduced innovative improvements by incorporating the Ghost convolution module from the GhostNet network into YOLOv8. Compared with models such as YOLOv8, the YOLOv8 + Ghost model demonstrates enhanced performance in several key indicators, including precision, mean Average Precision, recall, and computational cost.

To better adapt to complex agricultural environments and further improve the model's ability to identify pests, this study integrates the PSA attention mechanism into the YOLOv8 + Ghost model, aiming to achieve superior performance enhancements.

2. Optimization of the YOLOv8 model

2.1. YOLOv8 + Ghost model

The YOLO algorithm was first introduced by Redmon *et al*. in 2015 as an innovative single-stage object detection framework [3]. Its core concept involves transforming the object detection task into a regression problem, utilizing a convolutional neural network to directly infer on images for real-time object detection. Among the YOLO series, YOLOv8 (https://docs.ultralytics.com), developed by Ultralytics, represents an advanced object detection model. It incorporates numerous improvements and innovations over YOLOv5, significantly enhancing detection accuracy and speed, making it particularly suitable for lightweight real-time detection of agricultural pests.

During the training of conventional convolutional neural networks, a substantial amount of redundant and repetitive features are often generated, leading to unstable training, reduced efficiency, and resource wastage. The GhostConv module [4], proposed by Huawei Noah's Ark Laboratory in 2020 as a core component of the GhostNet network, efficiently generates feature maps, reducing computational complexity and parameter counts while maintaining performance. It is particularly suited for scenarios requiring efficient computation and low latency.

To enhance the YOLOv8 model in terms of computational resource consumption, detection accuracy, and response speed, the Conv and C2f modules in the Backbone and Neck networks of the YOLOv8n model are replaced with GhostConv and C3Ghost modules. The optimized structure of the YOLOv8 + Ghost model is illustrated in Figure 1. This design streamlines the feature extraction process, minimizes redundant calculations, and substantially reduces the number of parameters and computational volume while preserving the model's high performance.

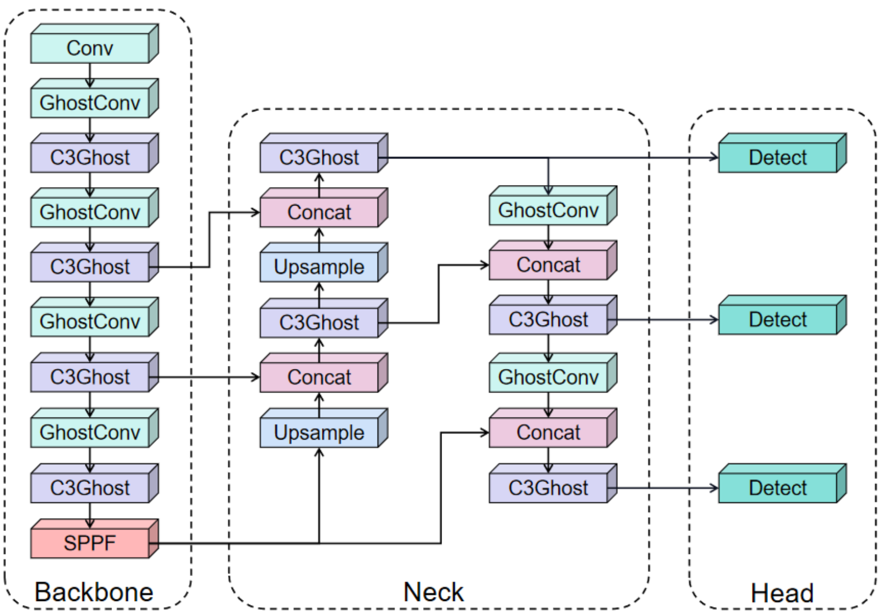


Figure 1. YOLOv8 + Ghost model

2.2. PSA attention mechanism

Polarized Self-Attention (PSA) is an advanced self-attention mechanism designed to enhance a model's ability to process high-resolution input and output features through polarization operations and efficient feature decomposition, while simultaneously reducing computational complexity. The core design of the PSA mechanism incorporates polarized filtering and high dynamic range (HDR) enhancement, integrating channel and spatial dual attention mechanisms. This makes it particularly suitable for fine-grained tasks such as image segmentation and super-resolution.

The PSA mechanism integrates two key designs:

(1) Polarized filtering: By folding specific dimensions of the input tensor, this design maintains high internal resolution in channel and spatial attention calculations while reducing computational complexity. This approach is analogous to the filtering effect of an optical lens on light, enhancing or weakening features to improve contrast.

(2) High Dynamic Range (HDR) enhancement: The dynamic range of the attention mechanism is expanded through a combination of Softmax and Sigmoid functions, enabling direct adaptation to the output distribution of fine-grained regression tasks. Specifically, the Softmax function is used for channel attention to force competition among channels, highlighting the most discriminative ones. The Sigmoid function is employed for spatial attention, allowing each position to activate independently and preserving the diversity of local details. HDR enhancement enables the model to capture both global channel dependencies and local spatial details simultaneously, overcoming the limitations of a single activation function.

The PSA module can combine channel attention and spatial attention in either parallel or serial configurations. In this study, a parallel layout structure is adopted.

2.3. YOLOv8 + Ghost + PSA model

This study further integrates the PSA attention mechanism into the YOLOv8 + Ghost model to optimize its feature extraction capabilities. The improvement primarily focuses on the Neck network, aiming to enhance the model's ability to capture spatial information and thereby improve detection performance. The improved model structure is depicted in **Figure 2**.

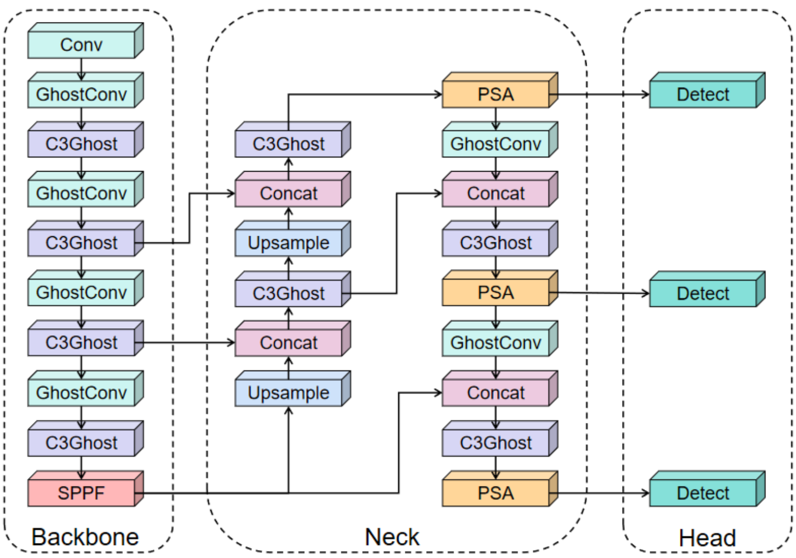


Figure 2. YOLOv8 + Ghost + PSA model

The enhanced YOLOv8 + Ghost + PSA model retains the lightweight design of YOLOv8 + Ghost while further optimizing the feature fusion process through the introduction of the PSA attention mechanism. The overall architecture of the model is divided into three parts. The Backbone network continues to employ the lightweight design of YOLOv8 + Ghost, utilizing GhostConv and C3Ghost modules for feature extraction to reduce computational complexity and parameter count. The Head network remains unchanged, responsible for outputting the final detection results.

In the feature fusion process, the Neck part integrates multi-scale features from the Backbone through upsampling and feature concatenation operations. Following these operations, the PSA module performs self-attention calculations on the fused feature maps. Specifically, the PSA module is added after each upsampling and concatenation operation, ensuring that the advantages of the self-attention mechanism are fully utilized at every stage of feature fusion.

Through these enhancements, the improved YOLOv8 + Ghost + PSA model optimizes the feature fusion capabilities while maintaining a lightweight design, thereby improving the model's detection performance. This design is especially well-suited for agricultural pest detection scenarios requiring high precision and low computational costs, effectively enhancing detection accuracy and efficiency.

3. Experiments

3.1. Experimental configuration

The experiments were conducted on a Windows 10 64-bit operating system with the following hardware configuration: An Intel Core i7-10700KF central processing unit (CPU) and a single NVIDIA A5000 24GB graphics processing unit (GPU), supporting CUDA version v11.0.

3.2. Dataset design and evaluation criteria

The dataset used in this study is based on the large-scale agricultural pest dataset IP102 released at CVPR 2019. Additionally, publicly available pest image resources were collected from the Internet. Ultimately, 10 major pests commonly found in farmlands were selected, and a total of 634 images were collected. Through systematic data augmentation techniques, including image rotation and brightness adjustment, the original dataset was expanded to 1904 images. The dataset was then divided into training and testing sets in an 8:2 ratio to ensure an even distribution of pest images across various categories.

Considering the application of this study in agricultural environments, where high precision and real-time performance are critical, Precision (P), Recall (R), mean Average Precision (mAP), and computation volume (GFLOPs) were chosen as evaluation metrics.

3.3. Experimental results and analysis

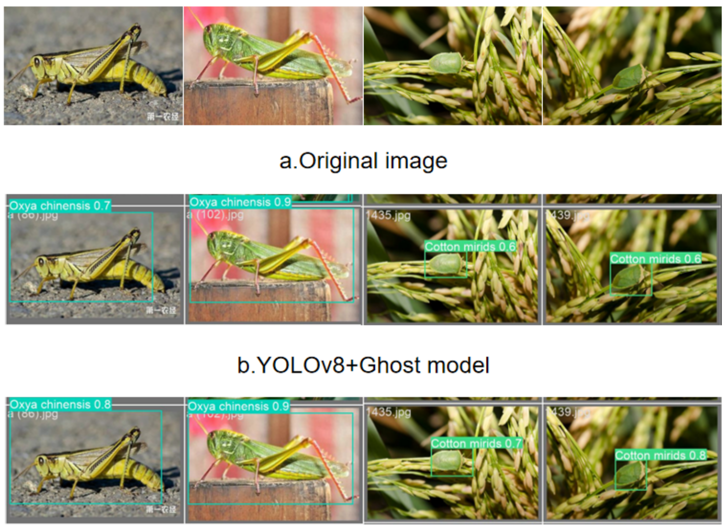
This study is based on the YOLOv8 + Ghost model and further improves it. To verify that the added PSA attention mechanism improves the model's performance, comparison experiments were conducted using pre-trained weights to evaluate the performance of the improved model. The relevant data of the trained models and the test results are shown in **Table 1**.

Table 1. Comparison results of models with PSA module

Models	P (%)	R (%)	mAP (%)	GFLOPs
YOLOv8 + Ghost + PSA	92.8	98.6	92.9	5.3
YOLOv8 + Ghost	91.4	83.0	90.8	5.0

Based on the experimental results, the following conclusions can be drawn: Adding the PSA attention mechanism to the YOLOv8 + Ghost model significantly improves the model's performance. Specifically, the precision of the YOLOv8 + Ghost + PSA model increased from 91.4% to 92.8%, the recall increased from 83.0% to 98.6%, and the mAP increased from 90.8% to 92.9%. This indicates that the PSA attention mechanism can better focus on target features, thereby improving the model's detection and classification accuracy.

The agricultural pest detection results are shown in **Figure 3**, with the original image at the top, the detection results of the YOLOv8 + Ghost model in the middle, and the detection results of the YOLOv8 + Ghost + PSA model at the bottom. A comparison also shows that the YOLOv8 + Ghost model has a slightly higher accuracy for pest detection.



a.Original image

b.YOLOv8+Ghost model

c.YOLOv8+Ghost+PSA model

Figure 3. Agricultural pest detection results

4. Conclusion

Based on the YOLOv8 + Ghost model, this study further optimized the Neck network part by introducing the PSA mechanism. Through the dual attention mechanism of channel and space, the feature fusion process was optimized. This improvement significantly enhanced the model's ability to capture spatial information, laying a solid foundation for improving recognition accuracy. Experimental results show that the improved YOLOv8 + Ghost + PSA model outperforms other comparative models in key indicators such as precision (92.8%), recall (98.6%), and mAP (92.9%). Meanwhile, the model's computation volume is only 5.3 GFLOPs, achieving a good balance between accuracy and efficiency. This optimized design has significant research and application value in resource-constrained agricultural environments and provides an efficient and accurate solution for agricultural pest detection.

Disclosure statement

The authors declare no conflict of interest.

References

[1] Cheng X, Zhang Y, Chen Y, et al., 2017, Pest Identification via Deep Residual Learning in Complex Background. Computers and Electronics in Agriculture, 141: 351–356.

[2] Yue G, Liu Y, Niu T, et al., 2024, GLU-YOLOv8: An Improved Pest and Disease Target Detection Algorithm Based on YOLOv8. Forests, 15(9): 1486.

[3] Redmon J, Divvala S, Girshick R, et al., 2016, You Only Look Once: Unified, Real-time Object Detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 779–788.

[4] Han K, Wang Y, Tian Q, et al., 2020, GhostNet: More Features from Cheap Operations. Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 1577–1586.