



Application of YOLOv8 Improved Model Incorporating Attentional Mechanism in Weed Identification

Zhanwei Feng¹, Adisak Sangsongfa^{1*}, and Noppadol Amdee¹

¹ Faculty of Industrial Technology, Muban Chombueng Rajabhat University, 70150, Thailand

* Correspondence:adisaksan@mcr.u.ac.th

Citation:

Feng, Z.; Sangsongfa, A.; Amm-Dee, N. Application of YOLOv8 improved model incorporation attentional mechanism in weed identification. *ASEAN J. Sci. Tech. Report.* **2026**, 29(4), e260523. <https://doi.org/10.55164/ajstr.v29i4.260523>

Article history:

Received: July 26, 2025

Revised: December 23, 2025

Accepted: January 11, 2026

Available online: March 24, 2026

Publisher's Note:

This article is published and distributed under the terms of the Thaksin University.

Abstract: Accurate identification of weeds plays a pivotal role in precision agriculture, as it supports targeted herbicide application and enhances crop productivity. This paper presents an improved YOLOv8-based object detection framework that integrates a global attention mechanism to strengthen weed recognition performance. Although YOLOv8 demonstrates strong baseline accuracy and speed, it lacks sufficient global contextual modeling, which constrains its ability to capture fine-grained features in visually complex or dense field scenarios. To overcome this limitation, we incorporate a Global Attention Module (GAM) after each C2f block in the backbone, enabling the network to capture long-range dependencies and highlight semantically relevant regions. Experiments on a custom weed dataset indicate that the attention-enhanced model considerably surpasses the baseline YOLOv8 in precision, recall, and mean Average Precision (mAP), especially under challenging conditions with small-scale or overlapping targets. Furthermore, the proposed method attains a favorable trade-off between detection accuracy and computational efficiency, making it suitable for real-time smart farming applications.

Keywords: Global attention; object detection; precision agriculture; weed detection; YOLOv8

1. Introduction

Weed infestation poses a critical challenge in contemporary agriculture, significantly reducing crop yields and increasing production costs [1–3]. Accurate and timely weed identification is essential for precision management, enabling selective application of herbicides to targeted areas [4–5]. Conventional identification approaches are often inefficient, labor-intensive, and susceptible to human error, particularly in large-scale field conditions [6–7]. As a result, the adoption of automated, robust, and real-time weed-detection technologies has become an urgent requirement for smart farming practices [8–9]. Driven by the rapid progress in deep learning, the YOLO (You Only Look Once) series has gained considerable popularity for its balance of detection accuracy and speed [10–11]. In particular, YOLOv8 achieves promising results by introducing several architectural improvements, such as the C2f module and anchor-free detection heads [12–13]. However, a critical research gap persists in current agricultural object detection: standard lightweight frameworks often struggle to maintain discriminative precision when confronted with cluttered backgrounds, overlapping plant structures, and small-scale targets [14–16]. Although existing studies have attempted to incorporate attention mechanisms such as CBAM or

SE, these modules are predominantly local in nature and often fail to capture the long-range cross-dimensional dependencies required to distinguish morphologically similar species in unconstrained environments [17-18]. To address these identified limitations, we propose an improved YOLOv8-based model that innovatively integrates a Global Attention Module (GAM) into the backbone network [19-20]. The primary technical contribution of this work lies in the strategic transition from local feature processing to multi-dimensional global dependency modeling by embedding GAM layers directly after each C2f block. This architectural synergy facilitates enhanced semantic feature propagation without significantly increasing model complexity, thereby boosting performance for small, hard-to-identify weed instances [21-22]. To validate our approach, experiments were conducted using a specifically constructed dataset featuring four species: Carpetweeds, Eclipta, Goosegrass, and Prickly Sida (Figure 1).

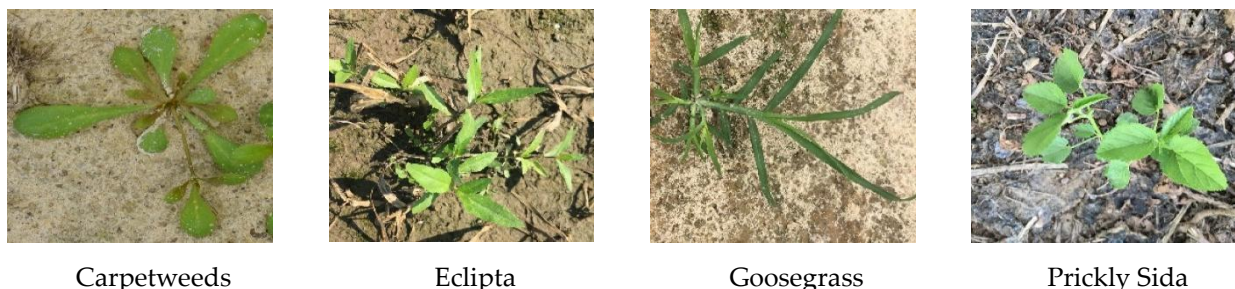


Figure 1. Representative examples of the weed types used in this study: Carpetweeds, Eclipta, Goosegrass, and Prickly Sida.

Beyond technical refinement, this study contributes to broader agricultural impacts. By preserving fine-grained morphological cues often overlooked by conventional detectors, our proposed framework bridges the gap between high-precision recognition and real-time edge deployment. Such advancement enables highly targeted herbicide application, thereby directly minimizing environmental chemical exposure and optimizing operational efficiency in precision agriculture. By incorporating global attention, this research provides a robust, field-ready solution that highlights semantically relevant weed features under complex field conditions, even when targets are partially occluded or densely overlapping. The following sections detail the proposed model's structure, the experimental setup, and a comprehensive analysis of the results.

2. The Method Described

2.1 Integration of Global Attention Module (GAM)

Although YOLOv8 offers a solid foundation for real-time object detection, its ability to capture long-range contextual relationships remains limited, particularly in complex agricultural environments with densely distributed or visually similar weeds [23-24]. To overcome this drawback, we embed a Global Attention Module (GAM) [25-26] within the YOLOv8 backbone, thereby improving the model's representational capacity by capturing global dependencies across both spatial and channel dimensions. The structural design of GAM is depicted in Figure 2. It is composed of two sequential components: a Channel Attention mechanism that emphasizes “what” information is critical, and a Spatial Attention mechanism that focuses on “where” the emphasis should be placed. Together, these attention branches refine the feature representations and yield an attention-guided output.

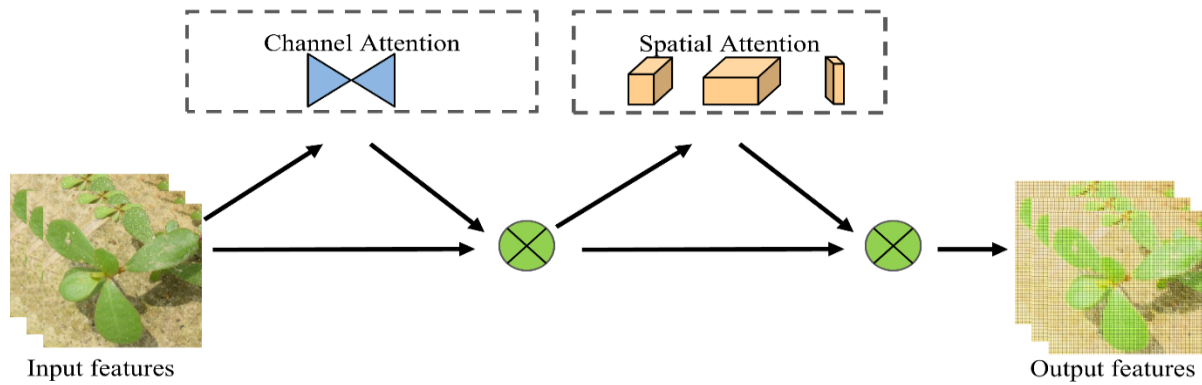


Figure 2. The Global Attention Module (GAM) comprises channel and spatial attention mechanisms. The module adaptively re-weights features to enhance relevant semantic information and suppress background noise.

Let the input feature map be denoted as $X \in \mathbb{R}^{C \times H \times W}$, where C , H , and W represent the number of channels, height, and width, respectively. To compute channel attention, global average pooling (GAP) is first applied to X , followed by two fully connected layers with ReLU activations and a sigmoid activation. This process generates the attention weights $A_c \in \mathbb{R}^C$, formulated as:

$$A_c = \sigma(W_2 \cdot \delta(W_1 \cdot \text{GAP}(X))) \quad (1)$$

where $W_1 \in \mathbb{R}^{C/r \times C}$ and $W_2 \in \mathbb{R}^{C \times C/r}$ are the weight matrices of two fully connected layers, r is the channel reduction ratio (typically set to 16), $\delta(\cdot)$ denotes the ReLU activation, and $\sigma(\cdot)$ is the sigmoid function. The resulting attention weights are then applied to the input feature map by channel-wise multiplication to produce the intermediate output X' :

$$X' = A_c \otimes X \quad (2)$$

Next, spatial attention is applied to emphasize informative regions across spatial locations. This is achieved by performing both average and max pooling along the channel axis, concatenating the results, and passing the concatenated results through a convolutional layer with a 7×7 kernel, followed by a sigmoid activation. The resulting spatial attention map $A_s \in \mathbb{R}^{1 \times H \times W}$ is computed as:

$$A_s = \sigma(\text{Conv}_{7 \times 7}([\text{AvgPool}(X'); \text{MaxPool}(X')])) \quad (3)$$

where $[\cdot; \cdot]$ indicates channel-wise concatenation, and $\text{Conv}_{7 \times 7}$ represents the convolution operation. The final output is obtained by applying spatial attention to X' :

$$X'' = A_s \otimes X' \quad (4)$$

Through this two-stage refinement, the Global Attention Module adaptively re-weights feature responses in both channel and spatial dimensions, enabling the model to capture better contextual and structural cues essential for accurate weed detection.

2.2 Proposed Method: YOLOv8 with Integrated GAM Modules

To address YOLOv8's limitation in capturing global semantic context, we introduce a novel modification by embedding the Global Attention Module (GAM) into its backbone. The schematic layout of the proposed model is presented in Figure 3. In this framework, GAM modules are strategically positioned after each C2f block within the backbone. These modules refine feature maps by successively applying spatial and channel attention mechanisms. The channel attention mechanism adjusts the relative weights of each channel based on global contextual cues, while the spatial attention mechanism emphasizes the most informative regions within the feature maps. This dual refinement enables the network to better concentrate on weed targets and suppress background noise. By integrating GAM directly into the feature-extraction hierarchy, the model progressively strengthens its semantic representation throughout the forward pass. Consequently, the network

achieves improved localization and classification accuracy, particularly for small or visually ambiguous weed instances. Notably, although introducing GAM modules incurs a modest computational cost, it retains the lightweight structure and real-time efficiency of the YOLOv8n baseline.

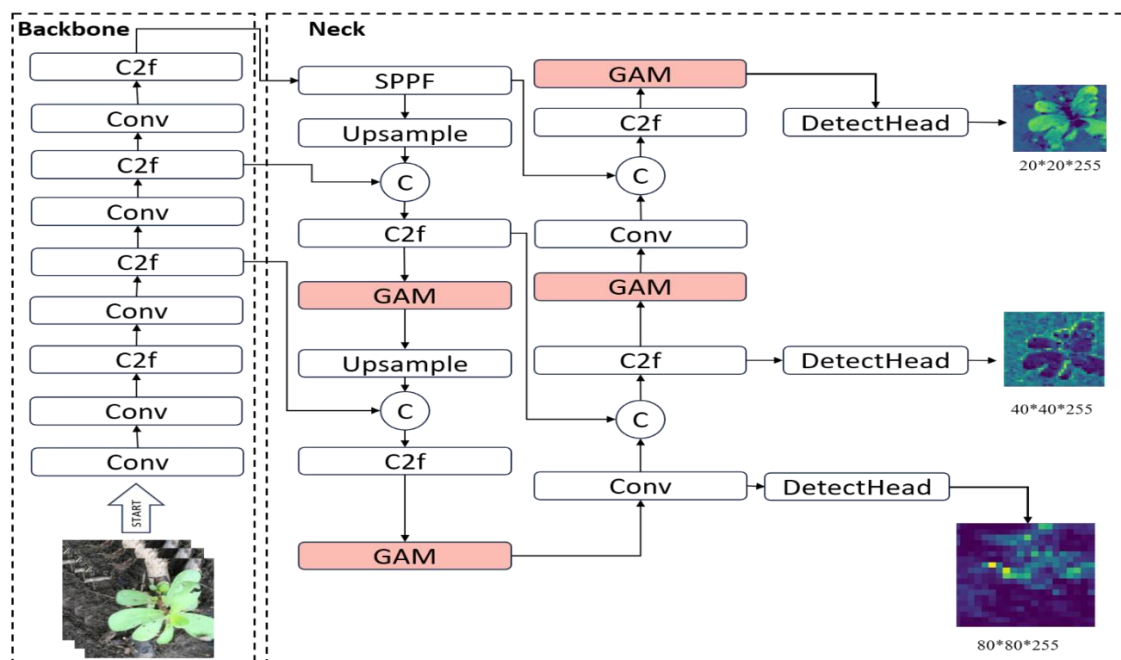


Figure 3. Architecture of the proposed GAM-enhanced YOLOv8 model. GAM modules are inserted after C2f blocks in the backbone to enrich feature representations. The output features exhibit enhanced semantic focus, especially on weed regions.

Compared to conventional attention mechanisms, such as the Convolutional Block Attention Module (CBAM), integrating GAM into our proposed model offers distinct advantages for capturing complex agricultural features. While CBAM sequentially infers attention maps along two separate dimensions (channel and spatial), it often suppresses significant cross-dimension interactions by using a simple pooling-based reduction. In contrast, the GAM employed in this study adopts a sequential channel-spatial attention structure that preserves information across multiple dimensions without drastic reduction. By utilizing a 3D-permutation-based technique and a multi-layer perceptron (MLP) within the channel submodule, and using two convolutional layers in the spatial submodule, GAM effectively mitigates the loss of fine-grained details. This characteristic is particularly vital for weed identification in dense field environments, where subtle morphological differences between crops and weeds require preserving rich spatial-channel dependencies that traditional mechanisms might overlook.

3. Results

This section presents the experimental outcomes, their interpretation, and the derived conclusions. Where appropriate, the results are organized into sub-sections to enhance clarity. The discussion focuses on how the findings align with or diverge from previous studies and their implications for the research hypotheses. Broader contextual relevance is also considered, and potential directions for future investigations are briefly outlined.

3.1 experimental design

To assess the performance of the proposed YOLOv8n + GAM framework, experiments were performed on a weed dataset containing 11,232 annotated images distributed across four categories: Carpetweeds, Eclipta, Goosegrass, and Prickly Sida. The images were primarily captured from a top-down perspective to ensure a consistent field of view for detection. To enhance the model's robustness to

environmental variability, the data collection spanned different times of day, encompassing diverse lighting conditions and weather (e.g., sunny and overcast). The dataset was partitioned into training, validation, and independent test sets at 70:20:10 (8,004 for training, 2,152 for validation, and 1,076 for final performance testing). The class distribution is detailed in Table 1, showing a relatively balanced representation to prevent model bias. To further improve generalization and prevent overfitting, offline data augmentation was applied to the training set, including horizontal flipping, random rotation ($\pm 15^\circ$), and brightness adjustment. All images were then standardized to a resolution of 640×640 pixels for consistent model input. For the training phase, we configured the Adam optimizer with a learning rate of 0.001 and processed 16-image batches across 300 epochs. To refine detection results during inference, we implemented Non-Maximum Suppression (NMS) to eliminate overlapping predictions, using a confidence cutoff of 0.25 and an IoU threshold of 0.45 for optimal box selection. The generalizability of the proposed method was subsequently verified on the test partition by measuring standard benchmarks, specifically mAP@0.5, mAP@0.5:0.95, Recall, and Precision.

Table 1. Ablation study results of GAM integration

Weed Category	Training	Validation	Test	Total Instances
Carpetweeds	2010	540	270	2820
Eclipta	1995	536	268	2799
Goosegrass	2004	541	271	2816
Prickly Sida	1995	535	267	2797
Total	8004	2152	1076	11232

3.2 Ablation Study

To examine the impact of the Global Attention Module (GAM), an ablation study was performed by comparing the baseline YOLOv8n model with the proposed YOLOv8n + GAM variant. Both models were trained under the same experimental conditions to ensure a fair comparison. A summary of the results is provided in Table 2.

Table 2. Ablation study results of GAM integration.

Method	GFLOPS	mAP@0.5	FPS
YOLOv8n (baseline)	8.2	0.86	60
YOLOv8n + GAM(our)	8.2	0.90	70

According to Table 2, the YOLOv8n + GAM model achieves a 4% improvement in mAP@0.5, increasing from 0.86 to 0.90. Notably, this accuracy gain comes without any increase in computational complexity, as both models share the same GFLOPs value of 8.2. Moreover, the inference speed of the GAM-enhanced model improves slightly, reaching 70 FPS compared to 60 FPS for the baseline.

Although attention mechanisms often entail additional computational overhead, the observed acceleration can be attributed to more discriminative feature representations produced by GAM. By effectively suppressing background noise and emphasizing informative spatial regions, the model reduces the number of redundant low-confidence predictions during inference. This refinement alleviates the computational burden of the Non-Maximum Suppression stage, which is commonly a bottleneck in object detection pipelines. Furthermore, the parallel-friendly design of the GAM module enables efficient execution on modern GPUs, such as the NVIDIA RTX 4090, thereby offsetting the nominal increase in internal operations and improving real-time detection performance. Overall, these experimental findings verify the efficacy of incorporating GAM into the YOLOv8n framework. The enhanced model effectively addresses challenges such as dense weed distribution, small-scale targets, and complex agricultural backgrounds—scenarios where conventional detectors often struggle. Moreover, the improvement in both accuracy and speed underscores the practicality of the proposed method for deployment in real-time, resource-constrained agricultural environments. This makes YOLOv8n + GAM a compelling solution for intelligent weed monitoring and precision farming applications.

Overall, these experimental findings verify the efficacy of incorporating GAM into the YOLOv8n framework. The enhanced model effectively addresses challenges such as dense weed distribution, small-scale

targets, and complex agricultural backgrounds—scenarios where conventional detectors often struggle. Moreover, the improvement in both accuracy and speed underscores the practicality of the proposed method for deployment in real-time, resource-constrained agricultural environments. This makes YOLOv8n + GAM a compelling solution for intelligent weed monitoring and precision farming applications.

3.3 Comparative Evaluation

To provide additional verification of the proposed approach, we conducted a comparative study between YOLOv8n + GAM and several widely adopted lightweight object detection models, including Faster R-CNN, SSD, YOLOv8n + CBAM, and an experimental version of YOLOv11. All models were trained and evaluated on the identical weed dataset under uniform conditions to guarantee fairness. A summary of the results is reported in Table 3.

According to Table 3, the YOLOv8n + GAM model reaches the highest mAP@0.5 (0.90) among all compared methods, surpassing both conventional and modern lightweight detectors. Notably, this improvement is achieved with a computational demand of 8.2 GFLOPs, which is considerably lower than that of heavier architectures such as Faster R-CNN, which achieves only 0.70 mAP@0.5 at the cost of 170 GFLOPs and 12 FPS. Although SSD offers a more balanced trade-off between accuracy and speed (35 GFLOPs and 36 FPS), its detection accuracy (0.76) is lower than that of the GAM-augmented YOLOv8n. Furthermore, when compared with the YOLOv8n + CBAM model, a commonly adopted attention-based baseline, the GAM-integrated model achieves an additional 2% improvement in mAP@0.5. This gain can be attributed to the GAM module's ability to model richer cross-dimensional feature interactions via a 3D permutation mechanism, whereas CBAM's sequential channel and spatial attention may limit global dependency modeling. Importantly, this accuracy enhancement is achieved without increasing computational complexity, as the model maintains a low cost of 8.2 GFLOPs while achieving a higher inference speed of 70 FPS, exceeding the 65 FPS of the CBAM-based variant. These results collectively confirm the effectiveness and efficiency of the proposed GAM-enhanced YOLOv8n for real-time weed detection in complex agricultural environments.

Table 3. Comparative Evaluation results of GAM integration.

Method	GFLOPS	mAP@0.5	FPS
YOLOv8n (baseline)	8.2	0.86	60
Faster-rcnn[27]	170	0.70	12
SSD[28]	35	0.76	36
YOLOv11[29]	12.5	0.80	52
YOLOv8n + CBAM[30]	8.3	0.88	65
YOLOv8n + GAM(our)	8.2	0.90	70

The experimental YOLOv11 model, although more efficient than Faster R-CNN and SSD in both GFLOPs and speed, still falls short in both accuracy (0.80 mAP@0.5) and FPS (52) compared to YOLOv8n + GAM. In contrast, our method achieves the best performance in both accuracy and inference speed (70 FPS), highlighting the effectiveness of the integrated Global Attention Module in improving feature discrimination without compromising real-time performance. These results strongly suggest that including GAM enables the detector to more effectively capture spatial and semantic context, particularly for small or overlapping weeds commonly found in real-world agricultural environments. By enhancing inter-channel and spatial interactions within feature maps, the attention mechanism enables the model to focus more precisely on weed regions while suppressing irrelevant background noise. The high mAP of 0.90 achieved by YOLOv8n+GAM is attributed to the Global Attention Module's ability to generate more distinct and non-redundant feature maps. This reduces the density of low-quality candidate proposals, allowing the NMS process to accurately select the most representative bounding box for each weed instance, thereby significantly reducing false positives. In conclusion, YOLOv8n + GAM offers a superior balance of accuracy, computational efficiency, and inference speed, making it well-suited for real-time deployment in precision agriculture applications such as autonomous weeding, field surveillance, and intelligent spraying systems. To intuitively demonstrate the

practical effectiveness of the proposed YOLOv8n + GAM model, we present several qualitative detection results on the validation set, as shown in Figure 4.



Figure 4. Qualitative detection results using the proposed YOLOv8n + GAM model. The model shows high confidence and accurate bounding boxes across different weed species and visual conditions.

From the visualizations, it is evident that the model successfully identifies and localizes various weed instances, including Carpetweeds, Eclipta, Goosegrass, and Prickly Sida, with high confidence scores and tight bounding boxes. The results remain stable even under challenging scenarios such as complex textures, partial occlusions, or intra-class shape variability. The visual clarity and localization precision in these examples further support the quantitative results reported earlier. Specifically, the use of the Global Attention Module enhances the model's ability to suppress background interference while focusing on the most informative regions of the image, resulting in more robust detection. These results affirm the model's suitability for deployment in real-world agricultural environments, where weed species often exhibit inconsistent appearances and grow under noisy, uncontrolled conditions. The proposed framework not only improves detection accuracy but also supports real-time, image-based weed monitoring solutions in precision farming systems.

4. Discussion of Results

Experimental validation on a custom weed dataset revealed that the YOLOv8n + GAM configuration consistently surpasses the baseline across multiple evaluation metrics. The enhanced model demonstrated marked improvements in precision, recall, and mAP, particularly when processing images containing small-scale weeds, overlapping instances, or dense spatial arrangements. Notably, these performance gains were achieved without introducing additional computational overhead, thereby maintaining the model's viability for deployment on resource-constrained edge devices commonly used in agricultural settings. The results confirm that attention-based mechanisms can effectively complement lightweight detection architectures, providing a practical pathway toward more accurate field-level weed-monitoring systems. Despite the promising outcomes, several limitations warrant acknowledgment. The current evaluation was conducted on a dataset collected under relatively controlled lighting and weather conditions, which may not fully represent the variability encountered across different growing seasons or geographic regions. Additionally, while GAM enhances global context modeling, the framework does not explicitly address challenges posed by extreme occlusion or severe background interference, scenarios that occasionally arise in densely planted fields. Furthermore, the study focused exclusively on RGB imagery; incorporating additional spectral bands (such as near-infrared) could potentially yield further performance improvements, but remains unexplored in this work. Several avenues exist for extending this research. Adaptive attention mechanisms that dynamically

adjust their receptive fields based on input characteristics could be investigated to improve robustness across diverse environmental conditions. Beyond this, integrating multi-modal data sources—combining RGB with thermal or hyperspectral imagery—may enable more comprehensive weed characterization and discrimination between weed species and crop plants. Expanding the evaluation to encompass multiple crop types and weed species under varied field conditions would further strengthen the generalizability of the proposed approach. Finally, exploring model compression techniques and knowledge distillation strategies could reduce computational requirements, facilitating deployment on low-power agricultural robots and drones. This research demonstrates that incorporating attention mechanisms into real-time object detectors represents a viable strategy for advancing precision agriculture applications. The proposed framework establishes a foundation for developing intelligent, autonomous crop management systems that support sustainable farming practices by enabling efficient, accurate weed identification.

5. Conclusions

This study introduced an enhanced weed detection framework by integrating a Global Attention Module (GAM) into the YOLOv8 architecture. The baseline YOLOv8, despite its robust real-time detection capabilities, exhibits limited ability to model global contextual information, which restricts its effectiveness in complex agricultural scenes characterized by visual clutter and irregular weed distribution patterns. By strategically inserting GAM blocks after each C2f module within the backbone network, our approach enables the model to establish long-range feature dependencies and selectively emphasize semantically important weed regions.

6. Acknowledgements

The authors gratefully acknowledge the insightful discussions and support of colleagues and peers throughout this work. We also thank those who assisted with data collection and experimental procedures. Finally, we appreciate the constructive feedback provided by the anonymous reviewers, which greatly enhanced the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

- [1] Paul, A.; Ghosh, S.; Mandal, A.; Dutta, S.; Banerjee, S.; Das, S. Smart Solutions for Capsicum Harvesting: Unleashing the Power of YOLO for Detection, Segmentation, Growth Stage Classification, Counting, and Real-Time Mobile Identification. *Comput. Electron. Agric.* **2024**, *219*, 108832. <https://doi.org/10.1016/j.compag.2024.108832>
- [2] Badgajar, C. M.; Poulouse, A.; Gan, H. Agricultural Object Detection with You Only Look Once (YOLO) Algorithm: A Bibliometric and Systematic Literature Review. *Comput. Electron. Agric.* **2024**, *223*, 109090. <https://doi.org/10.1016/j.compag.2024.109090>
- [3] Gerhards, R.; Andújar, D.; Hamouz, P.; Peteinatos, G. G.; Christensen, S.; Fernandez-Quintanilla, C. Advances in Site-Specific Weed Management in Agriculture—A Review. *Weed Res.* **2022**, *62*, 123–133. <https://doi.org/10.1111/wre.12526>
- [4] Singh, P.; Zhao, B.; Shi, Y. Computer Vision for Site-Specific Weed Management in Precision Agriculture: A Review. *Agriculture* **2025**, *15*(21), 2296. <https://doi.org/10.3390/agriculture15212296>
- [5] Guo, Z.; Cai, D.; Bai, J.; Xu, T.; Yu, F. Intelligent Rice Field Weed Control in Precision Agriculture: From Weed Recognition to Variable Rate Spraying. *Agronomy* **2024**, *14*, 1702. <https://doi.org/10.3390/agronomy14081702>
- [6] Hasan, A. S. M.; Sohel, F.; Diepeveen, D.; Laga, H.; Jones, M. G. K. A Survey of Deep Learning Techniques for Weed Detection from Images. *Comput. Electron. Agric.* **2021**, *184*, 106067. <https://doi.org/10.1016/j.compag.2021.106067>

- [7] Wang, A.; Zhang, W.; Wei, X. A Review on Weed Detection Using Ground-Based Machine Vision and Image Processing Techniques. *Comput. Electron. Agric.* **2019**, *158*, 226–240. <https://doi.org/10.1016/j.compag.2019.02.005>
- [8] Ji, X.; Yue, Z.; Yang, H.; Zhang, Z. Infrared Image Classification and Detection Algorithm for Power Equipment Based on Improved YOLOv10. *IEEE Access* **2024**, *12*, 184976–184988. <https://doi.org/10.1109/ACCESS.2024.3514103>
- [9] Lippi, M.; Bonucci, N.; Carpio, R. F.; Contarini, M.; Speranza, S.; Gasparri, A. A YOLO-Based Pest Detection System for Precision Agriculture. In *Proc. 29th Mediterr. Conf. Control Autom. (MED)*; **2021**; pp 342–347. <https://doi.org/10.1109/MED51440.2021.9480344>
- [10] Tang, Z.; Chen, Z.; Qi, F.; Zhang, L.; Chen, S. Pest-YOLO: Deep Image Mining and Multi-Feature Fusion for Real-Time Agriculture Pest Detection. In *Proc. IEEE Int. Conf. Data Mining (ICDM)*; **2021**; pp 1348–1353. <https://doi.org/10.1109/ICDM51629.2021.00169>
- [11] Pulipalupula, M.; Patlola, S.; Nayaki, M.; Yadlapati, M.; Das, J.; Reddy, B. R. S. Object Detection Using You Only Look Once (YOLO) Algorithm in Convolution Neural Network (CNN). In *Proc. IEEE 8th Int. Conf. Convergence in Technology (I2CT)*; **2023**; pp 1–4. <https://doi.org/10.1109/I2CT57861.2023.10126213>
- [12] Li, M.; Zhang, Z.; Lei, L.; Wang, X.; Guo, X. Agricultural Greenhouses Detection in High-Resolution Satellite Images Based on Convolutional Neural Networks: Comparison of Faster R-CNN, YOLOv3 and SSD. *Sensors* **2020**, *20*, 4938. <https://doi.org/10.3390/s20174938>
- [13] Liu, Y.; Zeng, F.; Diao, H.; Zhu, J.; Ji, D.; Liao, X.; Zhao, Z. YOLOv8 Model for Weed Detection in Wheat Fields Based on a Visual Converter and Multi-Scale Feature Fusion. *Sensors* **2024**, *24*, 4379. <https://doi.org/10.3390/s24134379>
- [14] Kumar, Y.; Kumar, P. Comparative Study of YOLOv8 and YOLO-NAS for Agriculture Application. In *Proc. 11th Int. Conf. Signal Process. Integr. Networks (SPIN)*; **2024**; pp 72–77. <https://doi.org/10.1109/SPIN60856.2024.10511673>
- [15] Liu, Q.; Lv, J.; Zhang, C. MAE-YOLOv8-Based Small Object Detection of Green Crisp Plum in Real Complex Orchard Environments. *Comput. Electron. Agric.* **2024**, *226*, 109458. <https://doi.org/10.1016/j.compag.2024.109458>
- [16] Huang, J.; Xia, X.; Diao, Z.; Li, X.; Zhao, S.; Zhang, J.; Zhang, B.; Li, G. A Lightweight Model for Weed Detection Based on the Improved YOLOv8s Network in Maize Fields. *Agronomy* **2024**, *14*(12), 3062. <https://doi.org/10.3390/agronomy14123062>
- [17] Wang, X.; Girshick, R.; Gupta, A.; He, K. Non-Local Neural Networks. In *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*; **2018**; pp 7794–7803. <https://doi.org/10.1109/CVPR.2018.00813>
- [18] Cao, Y.; Xu, J.; Lin, S.; Wei, F.; Hu, H. GCNet: Non-Local Networks Meet Squeeze-Excitation Networks and Beyond. In *Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshops (ICCVW)*; **2019**; pp 1971–1980. <https://doi.org/10.1109/ICCVW.2019.00246>
- [19] Ji, X.; Yue, Z.; Yang, H.; Li, M.; Han, H. SCAD-DETR: An Infrared Image Detection Method for Switchgear Equipment. *IEEE Internet Things J.* **2025**, in press. <https://doi.org/10.1109/JIOT.2025.3571499>
- [20] Liu, J.; Yu, E.; Li, Y.; Zhao, Y.; Mao, B. YOLO-DCPG: A Lightweight Architecture with Dual-Channel Pooling Gated Attention for Intensive Small-Target Agricultural Pest Detection. *Front. Plant Sci.* **2025**, in press. <https://doi.org/10.3389/fpls.2025.1716703>
- [21] Morbekar, A.; Parihar, A.; Jadhav, R. Crop Disease Detection Using YOLO. In *Proc. Int. Conf. Emerging Technol. (INCET)*; **2020**; pp 1–5. <https://doi.org/10.1109/INCET49848.2020.9153986>
- [22] Marvin-DiPasquale, M.; Windham-Myers, L.; Agee, J. L.; Kakouros, E.; Kieu, L. H.; Fleck, J. A.; Alpers, C. N.; Stricker, C. A. Methylmercury Production in Sediment from Agricultural and Non-Agricultural Wetlands in the Yolo Bypass, California, USA. *Sci. Total Environ.* **2014**, *484*, 288–299. <https://doi.org/10.1016/j.scitotenv.2013.09.098>
- [23] Guo, A.; Jia, Z.; Wang, J.; Zhou, G.; Ge, B.; Chen, W. A Lightweight Weed Detection Model with Global Contextual Joint Features. *Eng. Appl. Artif. Intell.* **2024**, *136*, 108903. <https://doi.org/10.1016/j.engappai.2024.108903>

-
- [24] Alhawsawi, A. N.; Khan, S. D.; Rehman, F. U. Enhanced YOLOv8-Based Model with Context Enrichment Module for Crowd Counting in Complex Drone Imagery. *Remote Sens.* **2024**, *16*, 4175. <https://doi.org/10.3390/rs16224175>
 - [25] Liu, Y.; Shao, Z.; Hoffmann, N. Global Attention Mechanism: Retain Information to Enhance Channel–Spatial Interactions. *arXiv Preprint* **2021**, arXiv:2112.05561. <https://doi.org/10.48550/arXiv.2112.05561>
 - [26] Ni, Z.-Y.; Wang, J.-C. A Global Attention Mechanism-Based EfficientNet Model for Road Pavement-Type Identification. *Int. J. Comput. Intell. Syst.* **2025**, *18*, 1–18. <https://doi.org/10.1007/s44196-025-00842-3>
 - [27] Sun, X.; Wu, P.; Hoi, S. C. H. Face Detection Using Deep Learning: An Improved Faster R-CNN Approach. *Neurocomputing* **2018**, *299*, 42–50. <https://doi.org/10.1016/j.neucom.2018.03.030>
 - [28] Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.-Y.; Berg, A. C. SSD: Single Shot Multibox Detector. In *Proc. Eur. Conf. Comput. Vis. (ECCV)*; **2016**; pp 21–37. https://doi.org/10.1007/978-3-319-46448-0_2
 - [29] Khanam, R.; Hussain, M. YOLOv11: An Overview of the Key Architectural Enhancements. *arXiv Preprint* **2024**, arXiv:2410.17725. <https://doi.org/10.48550/arXiv.2410.17725>
 - [30] Lai, Y. CBAM-Enhanced YOLOv8: An Attention-Based Approach for Tomato Disease Detection. In *Proc. 2nd Int. Symp. Agric. Eng. Biol. (ISAEB)*; **2025**; pp 337–346. https://doi.org/10.2991/978-94-6463-910-0_36