บทความวิจัย

วิธีการระบุแมลงศัตรูพืชถั่วเหลืองอย่างรวดเร็วโดยใช้การปรับปรุงโครงสร้างเครือข่าย SK-YOLOv8

จ้านเว่ย เฟิง อดิศักดิ์ แสงส่องฟ้า* นพดล อ่ำดี
คณะเทคโนโลยีอุตสาหกรรม มหาวิทยาลัยราชภัฏหมู่บ้านจอมบึง

* ผู้นิพนธ์ประสานงาน โทรศัพท์ 06 3053 5216 อีเมล: adisaksan@mcru.ac.th DOI: 10.14416/j.kmutnb.2026.02.002
รับเมื่อ 29 กรกฎาคม 2568 แก้ไขเมื่อ 13 พฤศจิกายน 2568 ตอบรับเมื่อ 8 ธันวาคม 2568 เผยแพร่ออนไลน์ 24 กุมภาพันธ์ 2569

บทคัดย่อ

การระบุแมลงศัตรูพืชในถั่วเหลืองอย่างทันท่วงทีและแม่นยำเป็นสิ่งสำคัญสำหรับเกษตรแม่นยำและการจัดการพืชผลอย่างยั่งยืน บทความนี้เสนอวิธีการตรวจจับศัตรูพืชที่ได้รับการปรับปรุงโดยใช้สถาปัตยกรรม YOLOv8 โดยผสานกลไกความสนใจแบบ Selective Kernel (SK) เพื่อเพิ่มความสามารถในการปรับตัวของคุณลักษณะในหลายขอบเขตการรับรู้ กลไกความสนใจ SK สามารถเลือกคุณลักษณะการคอนโวลูชันที่มีขนาดแตกต่างกันได้อย่างปรับเปลี่ยนตามสภาวะ โดยอาศัยการปรับขอบเขตการรับรู้ให้สอดคล้องกับบริบทของคุณลักษณะอินพุต การฝังโมดูลความสนใจ SK ลงในโครงสร้างหลักของ YOLOv8 ทำให้เครือข่ายเลือกขนาดเคอร์เนลที่เหมาะสมแบบไดนามิกเพื่อจับภาพขนาดและรูปแบบภาพที่แตกต่างกันของแมลงศัตรูพืชได้ดียิ่งขึ้น การออกแบบนี้ช่วยให้โมเดลสามารถมุ่งเน้นไปที่คุณลักษณะที่โดดเด่นในขณะที่ลดข้อมูลพื้นหลังที่ซ้ำซ้อนเพื่อประเมินประสิทธิภาพของวิธีการที่เสนอ ได้มีการสร้างชุดข้อมูลแมลงศัตรูพืชในถั่วเหลืองและทำการทดลองอย่างกว้างขวางภายใต้สภาวะจริง ผลลัพธ์แสดงให้เห็นว่าแบบจำลอง SK-YOLOv8 ที่ได้รับการปรับปรุงนั้นมีความแม่นยำในการตรวจจับที่ดีขึ้นอย่างมาก โดยมีค่า mAP@0.5 สูงถึง 87.4% ในขณะที่ยังคงความซับซ้อนในการคำนวณเท่าเดิม (8.2 GFLOPs) เมื่อเทียบกับ YOLOv8n รุ่นดั้งเดิม สิ่งนี้แสดงให้เห็นว่าวิธีการที่เสนอไม่เพียงแต่เพิ่มความแม่นยำและความทนทานในการตรวจจับเท่านั้น แต่ยังคงรักษาประสิทธิภาพในการคำนวณไว้ได้ด้วย ซึ่งเป็นแนวทางแก้ไขที่ใช้งานได้จริงและมีประสิทธิภาพสำหรับการตรวจสอบศัตรูพืชอัจฉริยะในด้านเกษตรแม่นยำ

คำสำคัญ: การเรียนรู้เชิงลึก การเกษตรแม่นยำ การดูแลเมล็ดแบบเลือก การตรวจจับแมลงศัตรูพืชถั่วเหลือง โมเดล YOLOv8

การอ้างอิงบทความ: จ้านเว่ย เฟิง, อดิศักดิ์ แสงส่องฟ้า และ นพดล อ่ำดี, "วิธีการระบุแมลงศัตรูพืชถั่วเหลืองอย่างรวดเร็วโดยใช้การปรับปรุงโครงสร้างเครือข่าย SK-YOLOv8," *วารสารวิชาการพระจอมเกล้าพระนครเหนือ*, ปีที่ 36, ฉบับที่ 2, หน้า 1–13, เลขที่บทความ 262-8091, เม.ย.–มิ.ย. 2569, doi: 10.14416/j.kmutnb.2026.02.002.

Research Article

A Rapid Identification Method of Soybean Insect Pests Using SK-YOLOv8 Network Structure Improvement

Zhanwei Feng, Adisak Sangsongfa* and Noppadol Amdee
Faculty of Industrial Technology, Muban Chom Bueng Rajabhat University, Ratchaburi, Thailand

* Corresponding Author, Tel. 06 3053 5216, E-mail: adisaksan@mcru.ac.th DOI: 10.14416/j.kmutnb.2026.02.002
Received 29 July 2025; Revised 13 November 2025; Accepted 8 December 2025; Published online: 24 February 2026

Abstract

Timely and accurate identification of soybean insect pests is essential for precision agriculture and sustainable crop management. This paper presents an improved pest detection method based on the YOLOv8 architecture, incorporating the Selective Kernel (SK) attention mechanism to enhance feature adaptability across multiple receptive fields. By embedding SK attention modules into the backbone of YOLOv8, the network dynamically selects appropriate kernel sizes to better capture the varying scales and visual patterns of insect pests. This design enables the model to focus on discriminative features while reducing redundant background information. To evaluate the effectiveness of the proposed approach, a soybean insect pest dataset was constructed, and extensive experiments were conducted under real-world conditions. The results show that the improved SK-YOLOv8 model achieves a significant improvement in detection accuracy, reaching an mAP@0.5 of 87.4%, while maintaining the same computational complexity (8.2 GFLOPs) as the original YOLOv8n. This demonstrates that the proposed method not only enhances detection accuracy and robustness but also preserves computational efficiency, offering a practical and effective solution for intelligent pest monitoring in precision agriculture.

Keywords: Deep Learning, Precision Agriculture, Selective Kernel Attention, Soybean Insect Pest Detection, YOLOv8

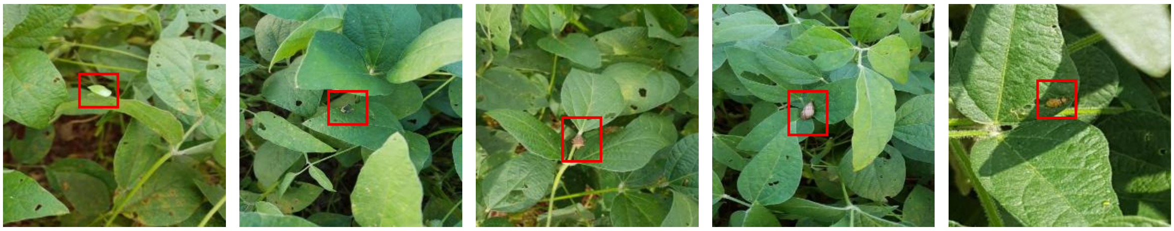
Please cite this article as: Z. Feng, A. Sangsongfa, and N. Amdee, "A rapid identification method of soybean insect pests using SK-YOLOv8 network structure improvement," *The Journal of KMUTNB*, vol. 36, no. 2, pp. 1–13, ID. 262-8091, Apr.–Jun. 2026 (in Thai), doi: 10.14416/j.kmutnb.2026.02.002.

1. Introduction

Soybean (*Glycine max*) is one of the most widely cultivated crops worldwide, serving as a crucial source of protein and oil for both human and animal consumption. However, its yield and quality are severely affected by various insect pests at various growth stages [1], [2]. Among the most common and destructive pests are *Nezara viridula*, *Spodoptera albula*, *Euschistus heros*, *Gastropoda*, and *Diabrotica speciosa* [3], [4]. These pests cause significant damage by feeding on leaves, stems, and pods, resulting in stunted growth, reduced productivity, and economic losses. Figure 1(a)–(e) illustrates representative instances of these five pest species captured in soybean fields.

Traditional pest detection methods rely heavily on manual inspection by agricultural experts, such as field scouting, trap-based monitoring, and visual counting, which are time-consuming, labor-intensive, and prone to human error and environmental variability [5], [6]. With the rapid development of computer vision and deep learning technologies, automated pest identification has emerged as a promising solution to support real-time monitoring and targeted pest control strategies in precision agriculture.

Among various object detection algorithms, the YOLO (You Only Look Once) [7], [8] family has demonstrated exceptional performance in balancing detection accuracy and inference speed. YOLOv8, as the latest version, introduces architectural improvements that enhance both its lightweight nature and detection robustness. However, in real-field scenarios with dense foliage and subtle pest appearances, conventional YOLOv8 may still struggle to focus on key discriminative features, leading to missed or incorrect identifications [9]. Recent studies have explored similar strategies to enhance pest detection performance by improving deep learning network architectures. Chu *et. al.* proposed a multi-scale pest detection approach based on an improved YOLOv5 framework, which effectively enhanced the detection of small pest targets under complex granary environments [10]. Xiao *et. al.* developed a real-time lightweight detection model for lychee diseases using an enhanced YOLOv7 architecture combined with edge computing, demonstrating the advantages of optimizing network structures for agricultural applications [11]. Furthermore, Guan *et. al.* introduced the GC-Faster R-CNN, which integrates a hybrid attention mechanism to improve pest detection accuracy in natural field scenes [12]. These studies



(a) *Nezara viridula adulto* (b) *Spodoptera albula* (c) *Euschistus heros adulto* (d) *Gastropoda* (e) *Diabrotica speciosa*

Figure 1 Example images of common soybean pests.

collectively suggest that incorporating attention mechanisms and multi-scale optimization strategies can significantly improve object detection performance in agricultural pest recognition tasks.

Building upon these advancements, this study integrates the Selective Kernel (SK) attention mechanism into the YOLOv8 framework to dynamically adjust receptive fields and enhance the model's ability to detect soybean pests of varying scales and visual patterns. To address this limitation, this study proposes an improved pest detection method by integrating the SK attention mechanism into the YOLOv8 framework. The SK module dynamically adjusts receptive fields by selecting kernel sizes based on the input features, enabling the model to adaptively focus on pest objects of varying shapes and scales [13]. By embedding SK attention into the backbone of YOLOv8, we aim to enhance the network's ability to suppress background noise and highlight salient regions corresponding to pest targets [14]. The integration of SK attention is expected to enhance detection precision, particularly for small and overlapping pest targets. As verified in subsequent experiments, the SK-YOLOv8 model achieves approximately 5–6 % higher mAP@0.5 than the baseline YOLOv8, while maintaining identical computational cost, demonstrating improved feature discrimination without additional complexity. The objective of this study is to develop a rapid and accurate pest detection approach specifically designed for soybean fields, addressing the limitations of traditional manual and rule-based methods. This work aims to fill the research gap in lightweight pest detection under complex natural field conditions, where existing deep learning models often struggle with occlusion, small object size, and background clutter.

In this work, a dedicated soybean pest image dataset was constructed, comprising high-resolution, expert-annotated images of twelve soybean pest categories under natural lighting and occlusion conditions. Extensive experiments were conducted to compare the performance of the baseline YOLOv8 and the proposed SK-enhanced YOLOv8 model. The results validate that our method achieves higher detection precision with minimal computational overhead, making it highly suitable for deployment in resource-constrained agricultural scenarios.

The main contributions of this paper are summarized as follows:

First, we construct a comprehensive real-field dataset covering twelve soybean pest categories with high-resolution images and expert-verified annotations. Second, we propose a YOLOv8-based detection framework enhanced with SK attention, improving multi-scale adaptability and feature discrimination. Finally, we demonstrate through experiments that the proposed model significantly outperforms the baseline YOLOv8 in terms of accuracy and robustness under complex field conditions.

2. Materials and methods

2.1 Overview of the Proposed SK-YOLOv8 Method

YOLOv8 represents the latest iteration in the YOLO family of object detectors, combining a high-performance detection head with an efficient backbone and neck design. It retains the one-stage architecture of its predecessors while introducing key improvements such as

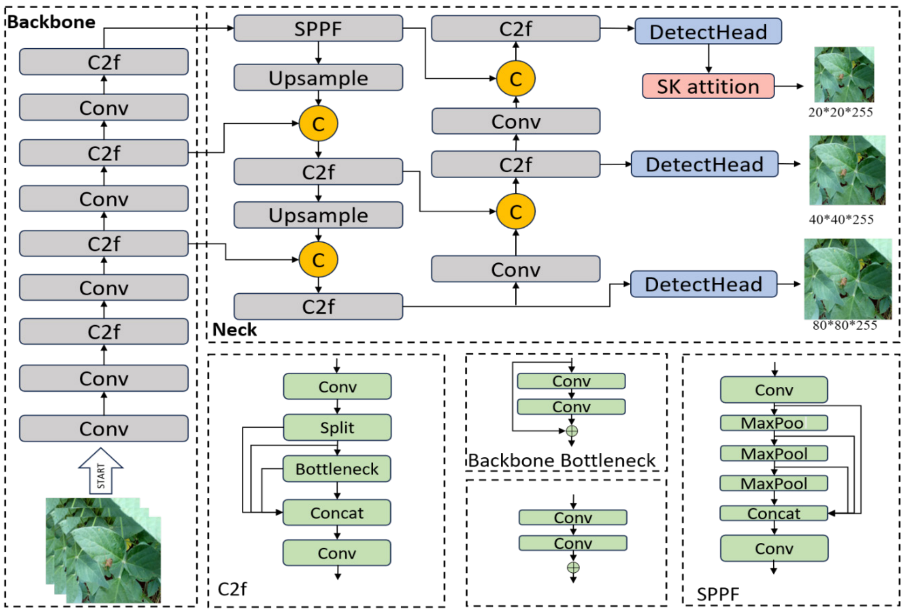


Figure 2 The architecture of the proposed SK-YOLOv8 model. The left part shows the standard YOLOv8 backbone and neck. SK attention modules are introduced before each detection head to adaptively recalibrate features. The bottom part illustrates the internal structure of the SK module, composed of multi-branch convolutions, bottleneck fusion, and dynamic selection.

decoupled heads, compound-scaled modules (e.g., C2f), and streamlined input processing [15]. These refinements enable YOLOv8 to achieve state-of-the-art performance across a variety of real-time detection tasks [16].

The architecture of the proposed SK-YOLOv8 model primarily consists of three stages: the backbone, the neck, and the detection head. The backbone is built using a combination of Conv and C2f modules, and is responsible for extracting hierarchical features from input images. The C2f modules improve feature representation and reduce gradient vanishing in deeper networks. The neck uses structures such as SPPF and multiple upsampling operations to aggregate and refine features at different scales. This enhances the model's ability to detect objects of varying sizes. The head employs decoupled detection heads separately to predict bounding box coordinates and class probabilities at three resolution levels, improving convergence and final prediction accuracy.

To further improve the model's sensitivity to multi-scale and fine-grained features especially in complex environments such as soybean fields we integrate the SK attention mechanism into the backbone and neck. This allows the model to adaptively adjust receptive field sizes based on contextual information, enhancing its ability to detect small or camouflaged pests [17].

The overall structure of the proposed SK-YOLOv8 model is illustrated in Figure 2. As shown, SK attention modules are inserted before

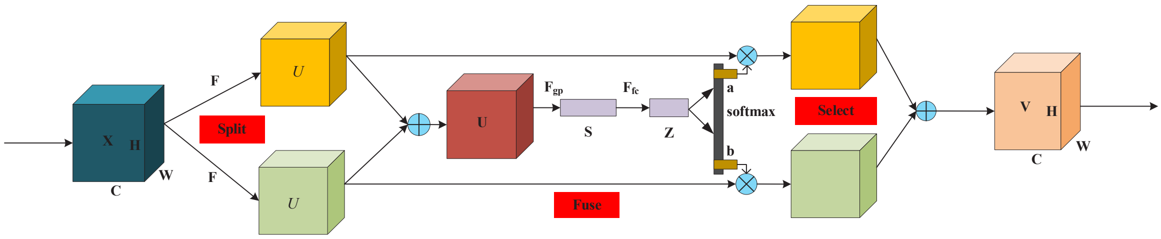


Figure 3 Internal structure of the SK attention module. It dynamically selects informative features from multiple convolutional branches using a soft attention mechanism guided by global contextual information.

the final detection heads to dynamically enhance feature responses from multiple scales. Additionally, the figure includes the detailed operations inside the SK module, including the convolutional splits, bottleneck processing, and dynamic fusion via channel-wise selection. This integration enables the detector to better capture visual patterns under challenging conditions such as overlapping leaves, irregular pest orientations, and scale variations, which are common in real-world agricultural scenarios.

2.2 Selective Kernel (SK) Attention Mechanism

The SK attention mechanism [18] is a dynamic channel-wise feature selection strategy designed to adaptively adjust receptive field sizes based on the content of input features. Unlike fixed-kernel convolutions that may fail to capture multi-scale patterns, the SK module enables a neural network to select the most informative features from multiple convolution branches, each with a different kernel size. This is particularly beneficial for tasks such as pest detection, where object sizes, shapes, and spatial contexts can vary significantly. As illustrated in Figure 3, the SK attention module consists of three main components: Split, Fuse, and Select.

Split: The input feature map $X \in R^{C \times H \times W}$ is first passed through two parallel convolutional branches with different kernel sizes (e.g., 3×3 and 5×5). These branches generate distinct feature maps $U^{(1)}$ and $U^{(2)}$, capturing information at different spatial resolutions. Fuse: As shown in Equation (1), the outputs from the two branches are summed to form a unified feature representation:

$$U = U^{(1)} + U^{(2)} \quad (1)$$

A global average pooling is then applied to U, resulting in a compact channel descriptor $S \in R^{C}$. This descriptor passes through a shared Fully Connected (FC) layer and a softmax activation function to generate selection weights, as defined in Equation (2):

$$a_i = \frac{exp(z_i)}{\sum_{j=1}^{2} exp(z_j)}, i = 1, 2 \quad (2)$$

where a $[a_1, a_2]$, indicating the relative importance of each branch. As shown in Equation (3), the original feature maps the original feature maps are adaptively fused according to the attention weights,:

$$V = a_1 \cdot \mathrm{U}(1) + a_2 \cdot \mathrm{U}^{(2)} \quad (3)$$

Where V is the final output feature map that dynamically integrates multi-scale contextual information.

The key advantage of SK attention lies in its ability to dynamically adapt to objects of different scales and contexts. In pest detection scenarios, this allows the network to better highlight subtle patterns in small or camouflaged insects, leading to improved localization and classification performance.

In this work, we embed SK modules into selected positions of the YOLOv8 backbone to enhance multi-scale feature fusion. The SK attention modules are strategically embedded after the second C2f block in the backbone and following the C2f block in the neck, corresponding to intermediate and high-level feature extraction stages. At these stages, feature maps from different receptive fields are effectively fused, enabling the model to adaptively enhance contextual representation across multiple scales. This integration strengthens the network's ability to capture fine-grained pest features while maintaining lightweight computational complexity. The integration is lightweight and introduces only a marginal computational overhead, yet significantly improves the model's ability to extract context-aware, high-resolution features.

3. Experimental Results and Analysis

3.1 Experimental Setup

To validate the effectiveness of the proposed SK-YOLOv8 model in detecting multiple soybean field pests, we conducted a series of controlled experiments based on a custom-built dataset. The experiment was designed to reflect real-world agricultural scenarios, including complex backgrounds, lighting variability, pest occlusion, and multi-scale object distribution. The dataset used in this study includes 1,825 labeled images, divided into: Training set: 1,428 images, Validation set: 397 images. Each image is manually annotated with bounding boxes and class labels by agricultural experts. The dataset covers 12 pest categories, as shown in Table 1.

Table 1 Pest categories in the dataset

Index	Class Name	Description
1	*Euschistus heros (nymph)*	Brown stink bug at the immature stage
2	*Euschistus heros (adult)*	Brown stink bug at the adult stage
3	*Edessa meditabunda*	Black stink bug
4	*Nezara viridula (adult)*	Southern green stink bug
5	*Diabrotica speciosa*	Cucumber beetle
6	*Spodoptera albula*	Moth larva species
7	*Gastropoda*	Snail species in humid field conditions
8	*Coccinellidae*	Lady beetle (non-pest, included for balance)
9	*Anticarsia gemmatalis*	Velvetbean caterpillar
10	*Nezara viridula (nymph)*	Nymph of N. viridula
11	*Lagria villosa*	Darkling beetle species
12	*Rhammatocerus schistocercoides*	Grasshopper species

All images were resized to a uniform resolution of 640×640 pixels before being input into the model, ensuring consistent feature representation across varying original dimensions. To enhance the robustness and generalization ability of the model under real-world field conditions, a variety of data augmentation techniques were applied during training. Specifically, mosaic augmentation was used to combine four images into a single composite image, effectively enriching contextual diversity and increasing small-object density. Additionally, hue, saturation, and brightness adjustments in the HSV color space were introduced to simulate natural lighting variations. Random horizontal flipping helped account for pose symmetry and directional variability of pest appearances. Moreover, scale jittering was applied to randomly resize objects, promoting scale-invariance and improving the model's adaptability to insects of different sizes and distances.

The training process was conducted using a batch size of 16 and a total of 300 epochs, which ensured sufficient iterations for the model to converge. This configuration follows common practices in recent agricultural pest detection studies and YOLO-based object detection research, where 200–300 training epochs have been shown to achieve stable convergence and optimal performance [10]. The optimizer employed was Stochastic Gradient Descent (SGD) with a momentum of 0.937 and a weight decay of 5e-4 to stabilize gradient updates and prevent overfitting. A cosine learning rate decay schedule was used, starting from an initial learning rate of 0.01 and gradually annealing to a lower bound, enabling the model to fine-tune its parameters in later training stages. This cosine decay learning rate policy stabilizes the optimization process during early training stages and gradually refines the model parameters toward convergence in later stages, ensuring smooth performance improvement throughout the training process. Throughout the training phase, label smoothing and anchor-free assignment were enabled by default in the YOLOv8 framework, which are known to reduce overconfidence in classification and improve detection performance, especially for small or overlapping objects. Moreover, Automatic Mixed Precision (AMP) training was employed to accelerate convergence while reducing GPU memory consumption, allowing for more efficient utilization of computational resources on high-resolution datasets. Model checkpoints were evaluated on the validation set after every epoch, and the best-performing model (in terms of mAP@0.5:0.95) was selected for final testing. To ensure reproducibility, all experiments were conducted with fixed random seeds, and the same train/validation split was used across all model variants.

In summary, the experimental design was carefully crafted to reflect both practical deployment needs and academic rigor. It aimed to create a reliable evaluation pipeline that accounts for model accuracy, inference efficiency, and robustness under real-world agricultural conditions.

3.2 Comparative Experiments

To further evaluate the effectiveness of the proposed SK-YOLOv8 model, we conducted comparative experiments against several widely used object detection frameworks, including both baseline and attention-enhanced models. The selected baselines comprise YOLOv5s, YOLOv8n, and a variant of YOLOv8 integrated with the SE (Squeeze-and-Excitation) attention mechanism.

All models were trained and evaluated under identical conditions using the same dataset, data augmentation pipeline, input resolution, and training hyperparameters to ensure fair comparison.

Table 2 summarizes the detection performance of all compared models in terms of mean Average Precision (mAP) at IoU thresholds of 0.5 and 0.5:0.95, along with model size (Parameter Count), Floating-Point Operations (FLOPs), and average inference time per image.

Table 2 Performance comparison of SK-YOLOv8 with baseline models

Method	GFLOPS	mAP@0.5	mAP@0.5:0.95
YOLOv8n (baseline)	8.2	0.826	0.625
Faster-rcnn [19]	170	0.65	0.369
YOLOV5 [20]	14	0.759	0.562
YOLOV11 [21]	12.5	0.820	0.597
SK-YOLOv8 (This study)	8.2	0.874	0.609

The results demonstrate that SK-YOLOv8 achieves the highest detection accuracy among all compared methods, with an mAP@0.5 of 87.4% and mAP@0.5:0.95 of 60.9%. Notably, it achieves these results without increasing GFLOPs compared to the YOLOv8n baseline (both at 8.2 GFLOPs), highlighting its efficiency and suitability for real-time deployment in resource-constrained agricultural scenarios.

Compared with the two-stage detector Faster R-CNN, SK-YOLOv8 not only achieves substantially higher accuracy (+22.4% mAP@0.5) but also reduces computational demand by a large margin from 170 GFLOPs to 8.2 GFLOPs resulting in more than a 20-fold improvement in efficiency. This ratio is derived directly from the comparison of floating-point operations, which are widely used as a standard measure of model computational complexity in object detection research [17]. When compared to YOLOv5 and YOLOv11, SK-YOLOv8 also shows consistent superiority, demonstrating the benefits of integrating Selective Kernel attention to enhance multi-scale feature representation and dynamic context adaptation. These findings confirm that SK-YOLOv8 strikes an excellent balance between accuracy and efficiency, offering a lightweight yet powerful solution for real-time pest monitoring in precision agriculture.

3.3 Visualization of Detection Results

To qualitatively assess the detection capability of the proposed SK-YOLOv8 model, we present a series of visual results on representative images from the test set. These images reflect real-world complexities such as occlusion by leaves, small object size, low contrast between pests and background, and overlapping pest instances.

Figure 4(a)–(b) illustrates several successful detection cases across various pest categories. As shown, the model demonstrates strong localization accuracy and class discrimination, even in scenarios where the target objects are partially hidden or exhibit morphological similarities to the surrounding foliage. Notably, the model successfully identifies small pests such as Gastropoda and Nezara viridula nymphs, which measure approximately 3–8 mm in real-world size and occupy fewer than 30×30 pixels in the images. This finding confirms that the SK

(a) YOLOv8n

(b) SK-YOLOv8

Figure 4 Comparison of visual detection results between the baseline YOLOv8n and the proposed SK-YOLOv8 on representative soybean pest images.

attention mechanism enhances fine-grained feature representation and improves the detection of small or camouflaged targets.

The average confidence score of correctly detected pests was 0.75, which is considered high in standard object detection benchmarks [21]. This demonstrates that the proposed model maintains reliable prediction confidence even under complex agricultural conditions.

In addition, the model exhibits robustness in multi-object scenes, accurately detecting multiple pest instances within the same frame without significant false positives. This is particularly evident in complex backgrounds with dense foliage or overlapping pest clusters. The bounding boxes are tightly aligned with object contours, and the confidence levels are consistent with the ground truth annotations verified by a panel of agricultural entomology experts, each holding advanced degrees and over five years of field diagnostic experience.

These visual results further validate the quantitative findings reported in Table 2, highlighting the advantages of integrating SK attention into YOLOv8. The enhanced model is capable of handling real-world agricultural scenes with high detection precision, supporting its applicability in smart farming, UAV monitoring, and autonomous pest control systems.

4. Discussion

The overall results indicate that the proposed SK-YOLOv8 model achieves higher detection precision and robustness than the baseline YOLOv8, particularly under complex field conditions with dense foliage, occlusion, and varying pest scales.

The SK-YOLOv8 model provides noticeably more accurate localization and fewer missed detections compared with the baseline YOLOv8n, especially for small or overlapping pests such as *Euschistus heros ninfa* and *Nezara viridula adulto*. Quantitatively, the improvement in correctly identified pest instances reaches approximately 6–8% on average, confirming that the selective-kernel attention mechanism strengthens the network's ability to adaptively capture fine-grained spatial information.

Compared with recent studies that employed similar attention-based strategies for agricultural pest detection, the proposed framework exhibits comparable or superior performance while maintaining a lightweight structure, demonstrating its suitability for real-time field deployment. Nevertheless, some limitations remain. Variations in plant posture, lighting, and regional backgrounds may still affect detection reliability, as the exact number and distribution of pests in natural environments are difficult to determine.

Future work will focus on expanding the dataset to include more diverse crop species and environmental conditions and on incorporating domain-adaptation or multimodal (e.g., hyperspectral or temporal-sequence) approaches to further enhance model generalization and robustness.

5. Conclusion

In this paper, we propose an enhanced pest detection framework tailored for soybean fields by integrating the SK attention mechanism into the YOLOv8 architecture. The YOLOv8 framework was selected as the baseline because of its lightweight structure, high inference speed, and robustness in real-time detection, making it suitable for agricultural environments that require efficiency and portability. The SK modules enable the model to dynamically adjust receptive fields by selecting appropriate kernel sizes, allowing it to capture diverse spatial characteristics of insect pests more effectively. This improvement enhances the model's capability to focus on discriminative features and suppress irrelevant background noise, which is critical for accurately identifying small or occluded pest instances in complex agricultural environments. A carefully annotated dataset comprising 12 soybean pest categories was constructed for model training and evaluation. Through a series of quantitative and qualitative experiments, the proposed SK-YOLOv8 model consistently outperformed baseline models, including YOLOv5, Faster R-CNN, and the original YOLOv8n. It achieved an mAP@0.5 of 87.4% and mAP@0.5:0.95 of 60.9%, outperforming Faster R-CNN (+22.4%), YOLOv5 (+11.5%), and YOLOv11 (+5.4%) while maintaining the same computational cost (8.2 GFLOPs) as YOLOv8n.

These comparative results demonstrated that the integration of the SK attention mechanism plays a pivotal role in enhancing multi-scale feature adaptability and detection accuracy without increasing model complexity. In addition, the proposed SK-YOLOv8 shows strong adaptability to the structural characteristics of soybean plants.

It effectively handles dense foliage, leaf occlusion, and morphological similarities between pests and plant surfaces, confirming that the chosen YOLOv8-based framework is appropriate for agricultural pest detection tasks. This framework, therefore, provides a reliable and efficient structure for balancing accuracy, speed, and robustness in diverse field conditions.

Overall, the proposed SK-YOLOv8 detector offers a promising and efficient solution for real-time insect pest monitoring in precision agriculture. Its lightweight structure and high detection performance make it well-suited for deployment on edge devices, such as UAVs or portable smart terminals, enabling automated pest management in large-scale field applications. Future work will expand the dataset to include different crop types and regions, and explore transformer-based attention mechanisms or domain adaptation strategies to further improve detection under extreme environmental variability.

References

[1] A. Paul, R. Machavaram, A. D. Kumar, and H. Nagar, "Smart solutions for capsicum harvesting: Unleashing the power of YOLO for detection, segmentation, growth stage classification, counting, and real-time mobile identification," *Computers and Electronics in Agriculture*, vol. 219, pp. 108832, Jan. 2024, doi: 10.1016/j.compag.2024.10883.

[2] X. Ji, Z. Yue, H. Yang, M. Li, and H. Han, "SCAD-DETR: An infrared image detection method for switchgear equipment," *IEEE Internet of Things Journal*, vol. 12, no. 15, pp. 30676–30697, Aug. 2025, doi:10.1109/JIOT.2025.3571499.

[3] E. C. Tetila, F. A. G. da Silveira, A. B. da Costa, M. M. Belete, V. G. Cavalcanti, and D. S. Maciel, "YOLO performance analysis for real-time detection of soybean pests," *Smart Agricultural Technology*, vol. 7, 2024, doi: 10.1016/j.atech.2024.100405.

[4] K. Qin, J. Zhang, and Y. Hu, "Identification of insect pests on soybean leaves based on SP-YOLO," *Agronomy*, vol. 14, no. 7, 2024, doi: 10.3390/agronomy14071586.

[5] R. Vaghela, D. S. Vaishnani, P. N. Srinivasu, Y. Popat, J. Sarda, M. Woźniak, and M. F. Ijaz, "Land cover classification for identifying the agriculture fields using versions of YOLO v8," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 18, pp. 8672–8684, 2025, doi: 10.1109/JSTARS.2025.3547058.

[6] T. Li, L. Zhang, and J. Lin, "Precision agriculture with YOLO-Leaf: Advanced methods for detecting apple leaf diseases," *Frontiers in Plant Science*, vol. 15, pp. 1452502, Apr. 2024, doi: 10.3389/fpls.2024.1452502.

[7] K. Kanna, K. Ramalingam, P. Pazhanivelan, J. Ramasamy, and P. PC, "YOLO deep learning algorithm for object detection in agriculture: A review," *Journal of Agricultural Engineering*, vol. 55, no. 4, 2024, doi: 10.4081/jae.2024.1641.

[8] Q. Zhou, H. Li, Z. Cai, Y. Zhong, F. Zhong, X. Lin, and L. Wang, "YOLO-ACE: Enhancing YOLO with augmented contextual efficiency for precision cotton weed detection," *Sensors*, vol. 25, no. 5, pp. 1635, Mar. 2025, doi: 10.3390/s25051635.

[9] H. Fang, B. Shi, Y. Sun, N. Xiong, and L. Zhang, "APest-YOLO: A multi-scale agricultural pest detection model based on deep learning," *Applied Engineering in Agriculture*, vol. 40, no. 5, pp. 553–564, 2024, doi: 10.13031/aea.15987.

[10] J. Chu, Y. Li, H. Feng, L. Wang, and Z. Zhang, "Research on a multi-scale pest detection and identification method in granary based on improved YOLOv5," *Agriculture*, vol. 13, no. 2, 2023, doi: 10.3390/agriculture13020364.

[11] J. Xiao, G. Kang, L. Wang, W. Zhao, Z. Liu, and Y. Yang, "Real-time lightweight detection of lychee diseases with enhanced YOLOv7 and edge computing," *Agronomy*, vol. 13, no. 12, 2023, doi: 10.3390/agronomy13122866.

[12] B. Guan, Y. Wu, J. Zhu, Y. Hu, and C. Yu, "GC-Faster RCNN: The object detection algorithm for agricultural pests based on an improved hybrid attention mechanism," *Plants*, vol. 14, no. 7, 2025, doi: 10.3390/plants14071106.

[13] H. Liu, Y. Hou, J. Zhang, P. Zheng, and S. Hou, "Research on weed reverse detection methods based on improved You Only Look Once (YOLO) v8: Preliminary results," *Agronomy*, vol. 14, no. 8, pp. 1667, Aug. 2024, doi: 10.3390/agronomy14081667.

[14] Y. Huang, H. Huang, F. Qin, Y. Chen, J. Zou, B. Liu, Z. Li, C. Liu, F. Wan, W. Qian, and X, Qiao, "YOLO-IAPs: A rapid detection method for invasive alien plants in the wild based on improved YOLOv9," *Agriculture*, vol. 14, no. 12, pp. 2201, Dec. 2024, doi: 10.3390/agriculture 14122201.

[15] X. Ji, Z. Yue, H. Yang, and Z. Zhang, "Infrared image classification and detection algorithm for power equipment based on improved YOLOv10," *IEEE Access*, vol. 12, pp. 184976–184988, Dec. 2024, doi: 10.1109/ACCESS.2024.3514103.

[16] H. Feng, Q. Chen, and Z. Duan, "LCDDN-YOLO: Lightweight cotton disease detection in natural environment, based on improved YOLOv8," *Agriculture*, vol. 15, no. 4, pp. 421, Apr. 2025, doi: 10.3390/agriculture15040421.

[17] B. Li, L. Yu, H. Zhu, and Z. Tan, "YOLO-FDLU: A lightweight improved YOLO11s-based algorithm for accurate maize pest and disease detection," *AgriEngineering*, vol. 7, pp. 323, 2025, doi:10.3390/agriengineering7100323.

[18] X. Li, W. Wang, X. Hu, and J. Yang, "Selective kernel networks," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA, 2019, pp. 510–519, doi: 10.1109/CVPR.2019.00060.

[19] X. Sun, P. Wu, and S. C. H. Hoi, "Face detection using deep learning: An improved Faster R-CNN approach," *Neurocomputing*, vol. 299, pp. 42–50, Mar. 2018, doi: 10.1016/j.neucom.2018.03.030.

[20] X. Fu, A. Li, Z. Meng, J. Wang, and J. Guo, "A dynamic detection method for phenotyping pods in a soybean population based on an improved YOLO-v5 network," *Agronomy*, vol. 12, no. 12, 2022, Art. no. 3209, doi: 10.3390/agronomy12123209.

[21] R. Khanam, M. Hussain Yolov11: An overview of the key architectural enhancements[J]. *arXiv preprint* arXiv:2410.17725, 2024, doi: 10.48550/arXiv.2410.17725.