REAL-TIME AND PRECISE DETECTION OF FIELD SOYBEAN RUST AND BACTERIAL SPOT BASED ON IMPROVED YOLOV11N

/

基于改进 YOLOv11n 的田间大豆豆锈病与细菌性斑点病实时精准检测

Tianhao WU, Yongcai MA*, Hanyang WANG
College of Engineering, Heilongjiang Bayi Agricultural University, Daqing 163319/China
Tel: +86-18745952020; E-mail: myc1631@163.com
Corresponding author: Yongcai MA
DOI: https://doi.org/10.35633/inmateh-79-18

Keywords: *Soybean Disease Detection, YOLOv11, Lightweight Model, Real-time Detection*

ABSTRACT

To overcome YOLOv11's limitations in complex field environments, this paper proposed SDD-YOLOv11n, a lightweight real-time detector for soybean diseases. The model reconstructed the backbone using GhostConv to minimize redundancy and integrates a C3k2_Star module to enhance small lesion detection against background noise. Additionally, a Detect Efficient (DE) head further compressed the architecture. Experimental results verified the model's efficiency, achieving a parameter count of 1.88 M and a weight size of 3.9 MB—reductions of 27.3% and 25% compared to YOLOv11n, respectively. Furthermore, the model maintained high detection performance with a Mean Average Precision (mAP50) of 75.6% and an F1-score of 69.3%, demonstrating its effectiveness in balancing architectural efficiency and accuracy in complex field environments.

摘要

针对 YOLOv11 在复杂田间环境下存在的局限性，本文提出了一种轻量化的大豆病害实时检测模型——SDD-YOLOv11n。该模型利用 GhostConv 重构主干网络以最大限度地降低特征冗余，并引入 C3k2_Star 模块，旨在抑制背景噪声干扰的同时增强对微小病斑的检测能力。此外，设计了 Detect Efficient (DE)检测头以进一步压缩模型架构。试验结果验证了该模型的高效性，其参数量仅为 1.88 M，权重文件大小为 3.9 MB，较 YOLOv11n 分别降低了 27.3%和 25%。同时，该模型在 mAP50 和 F1-Score 上分别达到了 75.6%和 69.3%，证明了其在检测精度和架构效率之间的良好平衡，能够有效应对复杂田间环境下的检测任务。

INTRODUCTION

As a vital oil and protein source, soybean production in China is integral to the national economy. However, prevalent diseases like soybean rust and bacterial spot can cause yield losses of 10%–80% or even total failure (*Liu et al., 2014*). Since traditional manual identification is inefficient and subjective, developing rapid, accurate detection technologies is essential for precision agriculture and food security (*Xiao et al., 2024*).

In recent years, to address issues regarding the accuracy and efficiency of crop disease identification, numerous scholars have focused on improving deep learning algorithms (*Xu et al., 2025*). *Zhu* proposed the CBF-YOLO algorithm, incorporating CSE-ELAN, Bi-PAN, and FFE modules, to detect common soybean pests in complex environments, significantly enhancing the model's detection capabilities (*Zhu et al., 2024*). *Korade* utilized YOLO for leaf extraction and applied various deep learning models to predict disease status. Their findings indicated that ResNet-50 was superior in capturing images from farms for soybean condition prediction, achieving the highest accuracy across all metrics; furthermore, they generated a confusion matrix to evaluate classification accuracy, verifying the model's robustness *(Korade et al., 2025)*. Using the public RiceDisease dataset, *Gan* proposed the G-YOLO model to improve rice leaf disease detection. By combining a lightweight efficient detection head with a multi-scale pyramid pooling layer, the model increased detection speed while maintaining considerable accuracy, making it highly suitable for deployment on resource-constrained edge devices (*Gan et al., 2025*). Similarly, to facilitate edge deployment for peanut disease identification, Sun developed YOLO-PLNet, a lightweight real-time detector specifically optimized for resource-constrained hardware *(Sun et al., 2025*).

To improve localization capabilities for tiny lesion features on cucumbers, Liu proposed the DCNSE-YOLOv7 algorithm, which introduced deformable convolutions and the SENet attention mechanism to strengthen effective target extraction, employed the K-means++ algorithm for anchor box reclustering, and replaced the loss function with Focal-EIOU (Liu *et al., 2023*). To enhance the identification accuracy of ginger leaf pests and diseases in natural environments, *Lan* introduced the GhostNet network and CA attention mechanism to propose Ginger-YOLOv5s, an improved model based on YOLOv5s that met real-time detection requirements while reducing missed and false detections (*Lan et al.,2024*). *Yang* proposed a lightweight tomato leaf pest and disease model based on YOLOv8n (*Yang et al., 2025*).

By reconstructing the original C2f with the StarBlock module and introducing the MLCA attention mechanism to reduce features along the channel dimension, further lightweighting of the original model was achieved. However, despite these advancements, existing studies generally encounter difficulties in balancing detection efficiency with accuracy. Common issues include excessive model parameters, high memory footprint, and a lack of universality in experimental data.

In summary, although deep learning is widely applied in agriculture, existing high-performance models (e.g., ResNet-50, Faster R-CNN) possess large parameter counts, making them difficult to deploy on agricultural drones and edge devices with limited computing power. While attempts have been made to lightweight the YOLO series, there remains a critical deficiency in maintaining a balance between extreme lightweighting (<4 MB) and performance preservation. Furthermore, optimizations targeting the YOLOv11 architecture for this context are currently insufficient. Consequently, this paper proposes SDD-YOLOv11n. By introducing GhostConv to reconstruct the backbone, integrating the C3k2_Star module, and designing a Detect Efficient (DE) head, this study aims to resolve the aforementioned challenges.

MATERIALS AND METHODS

IMAGE ACQUISITION AND DATASET CONSTRUCTION

The soybean disease dataset utilized in this study was collected at Jianshan Farm in Nenjiang, Heihe City, Heilongjiang Province. The dataset comprises three categories of leaf conditions: Healthy, Bean Rust, and Angular Leaf Spot (as shown in Fig. 1). To enhance image diversity and improve the algorithm's generalization capability, original images were acquired using a Redmi K60 smartphone under varying lighting conditions, weather scenarios, and growth backgrounds. Subsequently, the captured images underwent a manual screening process to select high-quality samples. All selected images were standardized to JPEG format, resulting in a total of 6,523 images.

To ensure the diversity of training samples and mitigate model overfitting, data augmentation techniques were applied to the existing dataset. Common augmentation methods employed include rotation, cropping, brightness adjustment, translation, background masking, and Gaussian noise injection. Specifically, noise injection was utilized to simulate environmental uncertainties encountered in real-world scenarios, thereby enhancing the model's robustness and generalization in complex field conditions (*Sulaiman et al., 2022*). Following data augmentation, the expanded dataset was manually annotated using LabelImg. The class labels were defined as "Healthy," "Bean_Rust," and "Angular_Leaf_spot." Finally, the dataset was randomly partitioned into training, validation, and test sets in a ratio of 8:1:1.

Despite the application of various data augmentation techniques to simulate complex field environments, the current dataset presents certain limitations. The images were exclusively acquired from a single geographic location (Jianshan Farm) using a specific capture device (Redmi K60), and encompass only three primary leaf conditions (Healthy, Bean Rust, and Angular Leaf Spot). Consequently, this restricted range of geographic, climatic, and disease category variations may inherently constrain the model's generalization capability when deployed across more diverse real-world agricultural scenarios. Acknowledging this, future data collection efforts will prioritize expanding the dataset across varied regions, diverse lighting scenarios, and encompassing a broader spectrum of soybean diseases to further enhance model robustness.

As a significant evolution in the series, YOLOv11 integrates novel feature extraction modules to maintain high speed and accuracy (*Khanam et al.,2024*). Its backbone, optimized from YOLOv9's GELAN architecture, features the C3k2 module. Compared to the C2f module in YOLOv8, C3k2 offers superior feature extraction with reduced parameter count and complexity, facilitating deployment across diverse environments (*Wu et al., 2025*). Among the five variants, this study selects YOLOv11n as the baseline. Its minimal computational footprint makes it particularly suitable for resource-constrained agricultural IoT devices (*Wang et al., 2025*).

a. Healthy b. Bean_Rust c. Angular_Leaf_spot

Fig. 1 - H (healthy) vs D (diseased) leaves of soybean

To further optimize the trade-off between detection accuracy and efficiency in YOLOv11n, this study proposes three structural enhancements. First, GhostConv substitutes standard convolutions to minimize computational redundancy and parameter volume. Second, a novel C3k2_Star module is introduced to refine the original C3k2 structure, offering a lighter and more flexible solution for resource-constrained field environments. Finally, a Detect Efficient (DE) head is implemented to replace the standard detection head, significantly reducing computational load while maintaining detection performance. The architecture of the improved SDD-YOLOv11n is illustrated in Figure 2.

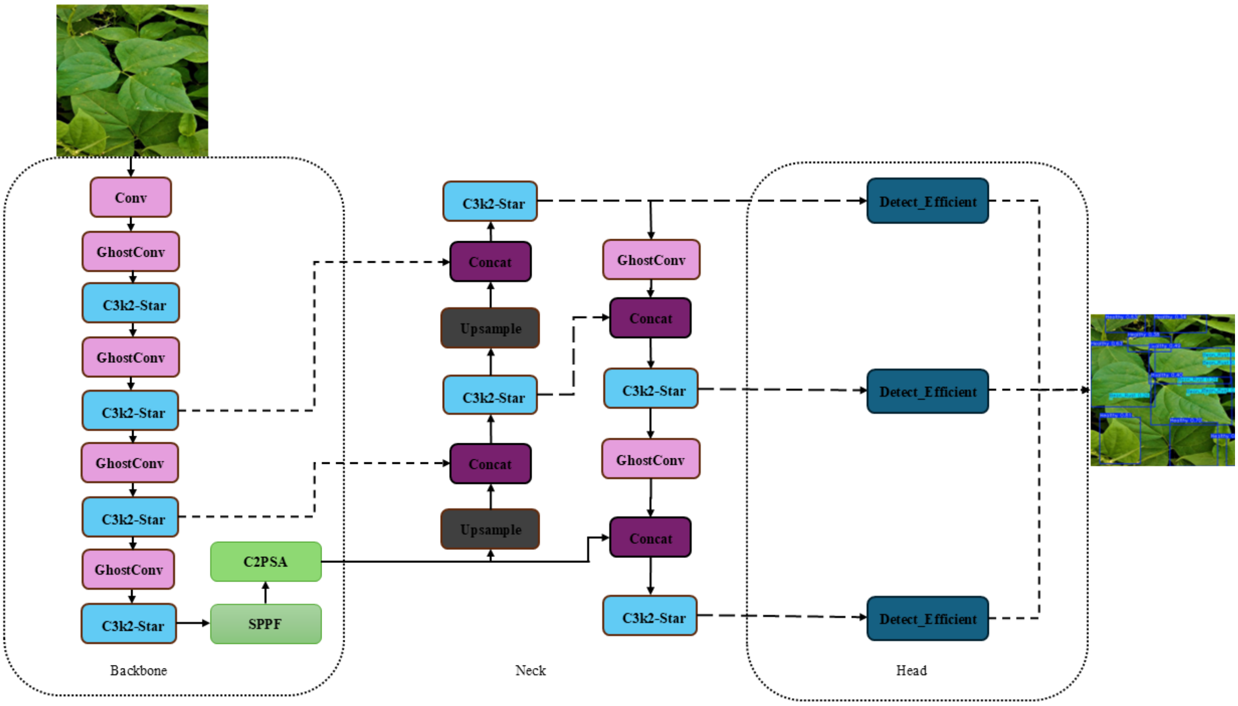

Fig. 2 - The architecture of the SDD-YOLOv11n network

Ghostconv Module

To accommodate resource-constrained edge devices in field applications, the substantial computational overhead caused by redundant feature maps in YOLOv11's standard convolutions presents a challenge (*Han et al., 2020*). Consequently, this study integrates the GhostNet architecture into YOLOv11n by substituting standard convolutions (excluding the initial layer) with GhostConv modules. By leveraging "cheap linear operations" to generate feature maps, this approach significantly compresses parameter volume while preserving critical spatial information. GhostConv achieves efficient feature learning by decoupling primary feature extraction from "ghost" feature generation.

The specific process is as follows:

1) Primary Feature Generation:

A small subset of convolution kernels is selected to extract core primary features. Assuming the input feature map is $X \in \mathbb{R}^{H \times W \times C}$, the number of primary feature channels is m (where $m \ll N$), and the number of convolution kernels is m, the calculation for the primary feature map Y' is defined as shown in Equation (1):

$$Y' = X * f' \quad (1)$$

where f' represents the primary convolution kernels, and the computational cost is approximately $\frac{1}{s}$ of that of standard convolution.

2) Ghost Feature Generation:

For each primary feature, $s-1$ ghost features are generated via cost-effective linear operations, where s represents the generation ratio controlling the total number of features $N = m \times s$. Letting the i-th primary feature be y_i', its corresponding ghost feature generator is $\Phi_{i,j}$. The calculation for the ghost features is defined as shown in Equation (2):

$$y_{i,j} = \Phi_{i,j}(y_i'),\ \forall i = 1,\dots,m,\ \ j = 1,\dots,s-1 \quad (2)$$

Upon completion of these two steps, the primary features and ghost features are concatenated to obtain the final output feature map Y.

In contrast to the full-channel operations in YOLOv11, GhostConv executes standard convolution on a subset of channels and generates the remainder via linear transformations. This mechanism significantly reduces computational complexity and parameter volume while preserving robust feature expression, thereby achieving model lightweighting without compromising detection accuracy.

C3k2_star Module

Derived from the Cross Stage Partial (CSP) architecture, the C3k2 module optimizes gradient flow and reduces redundancy through feature partitioning and fusion. However, to address the strict computational constraints of field environments, further lightweight optimization is required. Consequently, this study introduces the C3k2_Star module to replace the standard C3k2 in both the backbone and neck networks, thereby enhancing deployment efficiency.

The core of the C3k2_Star module lies in the stacking of n Star_Block instances. Derived from StarNet (*Ma et al., 2024*), the Star_Block employs element-wise multiplication for feature modulation. Inputs are processed via two parallel branches to generate a gating signal and a feature stream, respectively, which are then fused through element-wise multiplication to enhance feature representation (*Li et al., 2020)*.

Letting the input feature be X, this core interaction process is defined by Equation (3):

$$Y = (W_1 X) * f(W_2 X) \quad (3)$$

Where W_1 and W_2 represent the weight matrices of the two branches, respectively; f denotes the non-linear activation function; and $*$ represents element-wise multiplication. The structure of the StarBlock module is illustrated in Figure 3.

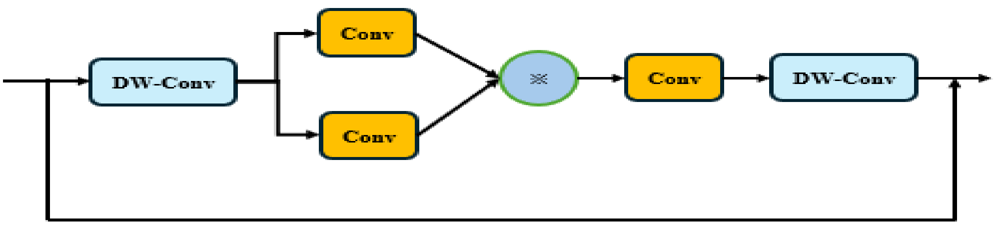

Fig. 3 - Structure of the StarBlock Module

Through element-wise multiplication, the module achieves dynamic channel re-weighting, effectively isolating minute disease symptoms from complex backgrounds by suppressing noise and amplifying relevant features. This "Star operation" significantly enhances representational capability via high-dimensional mapping Figure 3: Structure of the StarBlock Module maintaining a low parameter count, avoiding the substantial computational overhead associated with global attention mechanisms.

The final output undergoes a linear mapping to restore the channel dimensions, incorporating a residual connection as defined in Equation (4):

$$Y_{out} = W_3(Y_{star}) + X \quad (4)$$

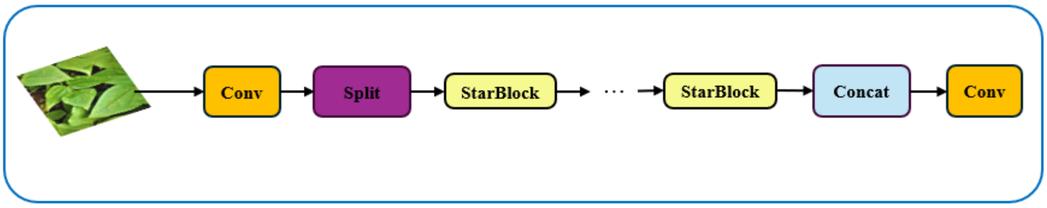

Fig. 4 - Structure of the C3k2_Star Module

Where W_3 represents the output projection matrix, and the addition operation denotes the residual connection. The structure of the C3k2_Star module is illustrated in Figure 4.

Detect_Efficient Module

To address parameter redundancy in YOLOv11n's decoupled head and enhance inference efficiency, this study introduces the DE module. While retaining Distribution Focal Loss (DFL) and Anchor-free strategies to ensure localization precision and multi-scale adaptability, the design implements lightweight optimizations to mitigate the high computational overhead associated with standard dense convolutional layers. The core principle of DFL is defined in Equation (5):

$$\hat{y} = \sum_{i=0}^{n} P\,(y_i) y_i \tag{5}$$

Where $\hat{y}$ represents the predicted bounding box coordinate, and $P(y_i)$ denotes the predicted probability density function. During the discretization process, the continuous regression range is discretized into the interval $[y_0, y_n]$, where $\sum P\,(y_i) = 1$.

To mitigate the high computational cost of dense convolutions, the DE module substitutes standard layers with stacked 3×3 Grouped Convolutions at the input of each branch. This design significantly lowers parameter count and complexity via independent channel processing, achieving an optimal speed-accuracy trade-off suitable for agricultural deployment.

RESULTS

Training Environment and Experimental Setup

Experiments were conducted on a Linux platform featuring an NVIDIA RTX 4090D (24GB) GPU and an 18-core 2.6GHz CPU, utilizing PyTorch 2.2.0 (CUDA 12.1). Training employed the SGD optimizer (momentum=0.937, initial $lr = 0.01$) with an input resolution of 640×640, a batch size of 64, and a duration of 300 epochs. Performance was evaluated via Precision(P), Recall(R), mAP, Parameters, GFLOPs, FPS, and Model Size(M-S).

Model Performance Evaluation

Upon evaluation on the test set, the computational complexity and inference efficiency of the optimized SDD-YOLOv11n model are presented in Table 1.

Table 1

The computational complexity and inference efficiency of the optimized SDD-YOLOv11n

GFLOPs	Parameters	FPS（Inference）	FPS(Comprehensive)	M-S/MB
4.4	1877017	503.59	285.05	3.9

Experimental results demonstrate that SDD-YOLOv11n exhibits significant lightweight efficiency, characterized by 1.88 M parameters, 4.4 GFLOPs, and a 3.9 MB file size, rendering it ideal for edge deployment. Furthermore, the model achieves exceptional speed—503.59 FPS for pure inference and 285.05 FPS for comprehensive processing—fully satisfying the stringent real-time requirements of agricultural disease detection.

The quantitative evaluation results for the SDD-YOLOv11n model are detailed in Table 2.

Table 2

Quantitative Evaluation Results of the SDD-YOLOv11n Model

P / %	R / %	F1-Score / %	mAP50 / %	mAP50-95 /%
68.9	70.3	69.3	75.6	54.2

Results indicate that the model demonstrates robustness with an mAP of 75.6% and a high-threshold mAP of 54.2%. Furthermore, an overall F1-score of 69.3% signifies a well-balanced trade-off between precision and recall.

Ablation Studies of Proposed Improvements

To validate the efficacy of the proposed improvements, a comprehensive ablation study was conducted on the YOLOv11n baseline. By incrementally integrating GhostConv, C3k2_Star, and the DE Head, the specific quantitative results are presented in Table 3.

Table 3

Ablation experiment

Model	Parameters	GFOLPs	M-S/MB	FPS	P%	R/%	mAP50)/%
Base	2582737	6.3	5.2	290.56	68.9	72.0	76.9
Base+a	2257041	5.5	4.6	298.38	69.3	69.9	76.3
Base+b	2472393	6.4	5.1	267.92	69.7	71.1	76.4
Base+c	2313041	5.1	4.7	290.47	68.9	71.5	76.9
Base+a+b	2146713	5.6	4.5	302.05	70.1	69.6	76.1
Base+a+b+c	1877017	4.4	3.9	285.05	68.9	70.3	75.6

Note: "Base" refers to the YOLOv11n model; "a", "b", and "c" denote the GhostConv module, the C3k2_Star module, and the DE module, respectively.

Combining all three modules yielded the best performance, validating the effectiveness of each component and their synergy.

Comparative Analysis of Different Detection Models

To benchmark SDD-YOLOv11n, quantitative comparisons were conducted against mainstream detectors—including Faster R-CNN(F-R) (*Ren et al., 2016*), SSD (*Liu et al., 2024*), and recent YOLO iterations (*Tian et al., 2025, Lei et al., 2025*)—under identical experimental conditions. Performance was evaluated via mAP, Parameters, GFLOPs, FPS, Recall, and Memory Size, with specific results detailed in Table 4.

Table 4

Contrast test

Model	Parameters	GFOLPs	M-S/MB	FPS	P/%	R/%	mAP50)/%
YOLOv11n	2582737	6.3	5.2	290.56	68.9	72.0	76.9
YOLOv12n	2508929	5.8	5.2	266.42	70.2	70.3	76.7
YOLOv13n	2448480	6.1	5.2	250.64	70.1	69.6	76.1
SSD	24103280	30.5	12.5	55.20	61.5	64.0	68.5
F-R	41350110	180.2	25.4	18.50	71.0	68.5	74.5
Ours	1877017	4.4	3.9	285.05	68.9	70.3	75.6

Comparative analysis reveals that while Faster R-CNN achieved the highest Precision (71%), its extensive computational overhead (41.3 M parameters, 18.5 FPS) and SSD's sub-optimal performance limit their practical utility. This aligns with the findings of *Korade*, whose international study demonstrated that while heavy architectures like ResNet-50 can achieve superior accuracy in soybean disease prediction, they often pose significant challenges for real-time farm deployment due to their complexity. In contrast, SDD-YOLOv11n surpasses these traditional models in comprehensive efficiency, demonstrating superior feature fusion in complex farmland environments. Against state-of-the-art YOLO variants (v12n, v13n), the proposed model reduces parameters and GFLOPs by approximately 25% while enhancing inference speed by 18–34 FPS. Achieving this lightweight design with negligible accuracy loss, SDD-YOLOv11n strikes a critical balance for energy-constrained drone applications.

Ultimately, SDD-YOLOv11n addresses the hardware constraints of smart agriculture. With only 1.87 M parameters (approx. 1/22 of Faster R-CNN) and the lowest computational load (4.4 GFLOPs), its minimal memory footprint (3.9 MB) facilitates cost-effective deployment on embedded systems like Raspberry Pi and Jetson Nano. The model's exceptional inference speed (285.05 FPS) not only extends operational endurance but also reserves ample computational margin for real-time decision-making systems.

Visualization of Detection Results

In this section, the detection results obtained from the trained YOLOv11n, YOLOv12n, YOLOv13n, and SDD-YOLOv11n models are visualized. The resulting comparison is illustrated in Figure 5, where bounding boxes are color-coded as follows: dark blue denotes “Healthy” leaves, sky blue indicates “Bean Rust,” and white represents “Angular Leaf Spot.”

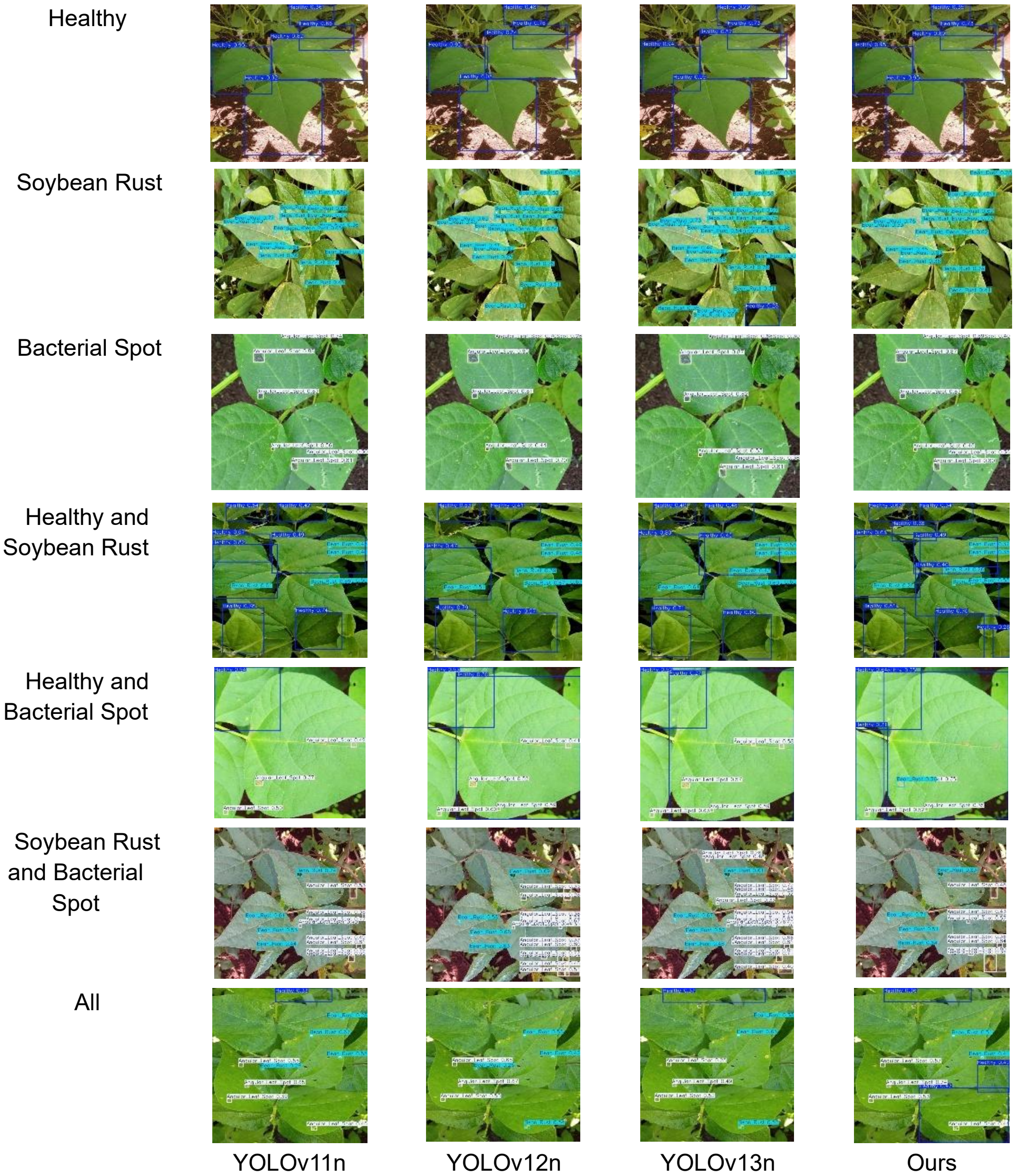

Fig. 5 – Visual Comparison of Detection Results Across Different Models

As illustrated, SDD-YOLOv11n achieves detection accuracy comparable to state-of-the-art benchmarks (YOLOv11n, v12n, v13n). Notably, it demonstrates superior robustness in classifying “bean rust,” avoiding the misclassifications observed in YOLOv12n and v13n. The results confirm the model’s efficacy in distinguishing between healthy, rust-infected, and angular leaf spot samples. Furthermore, by maintaining high detection speeds with minimal resource consumption, SDD-YOLOv11n is highly conducive to real-time deployment on edge devices.

Error Analysis and Robustness Evaluation

Figure 6 illustrates the confusion matrix of SDD-YOLOv11n on the test set, demonstrating exceptional classification consistency within complex field environments. Specifically, the “Healthy” category achieves near-perfect recognition. Misclassifications between the visually similar “Bean Rust” and “Angular Leaf Spot” (104 and 187 instances, respectively) are negligible relative to their vast true positive baselines. This confirms the robust discriminative feature extraction capability of the proposed GhostConv backbone and C3k2_Star module.

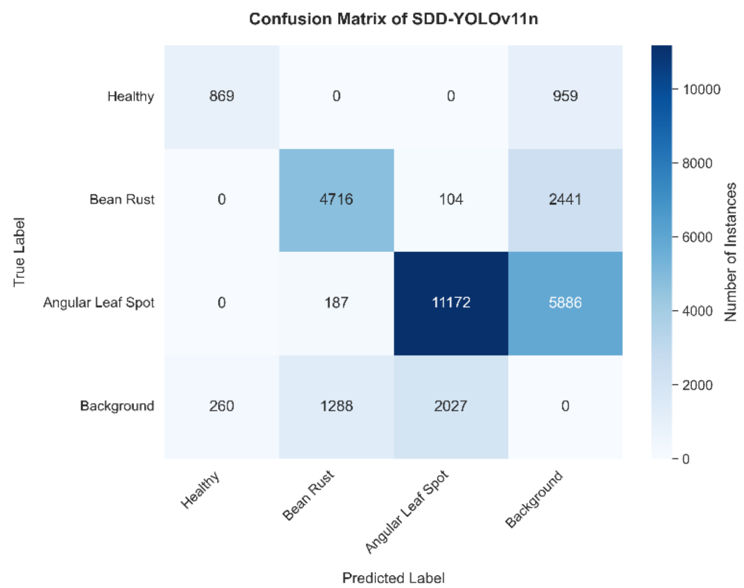

Fig. 6 – Confusion matrix of the SDD-YOLOv11n model on the test set

However, the matrix's background row and column reveal the primary error source: false negatives and false positives induced by background interference, with "Angular Leaf Spot" exhibiting a notably higher missed detection rate. To further interpret these error mechanisms and validate the model's decision-making process, Grad-CAM visualizations (Fig. 7) were evaluated on representative scenarios. This interpretability approach follows the recent advancements by Alhammad, who utilized deep learning combined with explainable AI (XAI) to ensure the transparency and reliability of disease classification in plant pathology *(Alhammad et al., 2025)*. As shown in Fig. 7(a), the model demonstrates exceptional feature focusing capabilities; even under severe leaf overlapping, the high-response regions (dark red) precisely localize on minute, early-stage bean rust lesions, visually confirming the efficacy of the proposed modules. Conversely, Fig. 7(b) and 7(c) expose the model's limitations under severe background noise. Fig. 7(b) illustrates a typical missed detection where a diseased leaf is ignored because the model's attention is erroneously drawn to surrounding weeds, resulting in a low feature response (cold tones) on the actual lesion. Furthermore, Fig. 7(c) presents a false-positive case where a non-target weed is misclassified as a "Healthy" leaf. Crucially, this uncovers an attention defocusing mechanism: the high-response areas are severely scattered across the background soil. This hallucination-like misclassification indicates that homologous green backgrounds and extreme lighting remain the main bottlenecks for the lightweight model's spatial attention allocation. Future work will integrate multi-scale feature fusion strategies to enlarge the global receptive field and optimize detection performance in such challenging scenarios.

(a) Accurate feature focusing on early-stage bean rust

(b) A false-negative case caused by severe background distraction

(c) A false-positive case showing attention defocus caused by complex background noise

Fig. 7 - Grad-CAM visualizations of representative detection results

CONCLUSIONS

To address the requirements for field-based soybean disease detection, this study successfully constructed a high-quality soybean disease image dataset. Based on this dataset and the YOLOv11n architecture, an efficient detection model for soybean rust and bacterial spot—SDD-YOLOv11n—was designed and implemented. The proposed model significantly reduces parameter count, GFLOPs, and model size while maintaining commendable accuracy, thereby providing a theoretical foundation for subsequent applications in actual field environments. The specific conclusions are as follows:

1. Structural Improvements: This study introduced structural optimizations to the YOLOv11n baseline. Specifically, standard convolutional layers were replaced with Ghost modules from the GhostNet architecture. Furthermore, the C3k2_Star module, inspired by StarNet principles, was proposed to supersede the original C3k2 module. Finally, a highly efficient detection head, Detect_Efficient, was developed. Experimental results indicate that, at the cost of a marginal decrease of approximately 1% in Mean Average Precision (mAP), the model achieved reductions of roughly 27.3% in parameter count, 30.1% in floating-point operations, and 25% in model size, while simultaneously increasing the frame rate by 5.5 FPS. This demonstrates a substantial reduction in computational demand while maintaining robust accuracy, thereby satisfying the conditions required for practical detection.
2. Comparative Advantage: Compared to current mainstream detection models, SDD-YOLOv11n exhibits significant advantages in terms of lightweight characteristics. While ensuring high accuracy, the model volume is merely 3.9 MB, and the floating-point operations are limited to 4.4 GFLOPs. In conclusion, SDD-YOLOv11n demonstrates superior performance regarding parameter count, model size, and computational complexity.

Although the SDD-YOLOv11n model proposed in this study achieved promising detection results on the self-constructed soybean leaf disease dataset, there remains room for improvement. Limitations exist regarding the diversity of disease types and the environmental complexity within the dataset due to the challenging nature of data acquisition. Future research will focus on expanding the scale and diversity of the data. Additionally, model compression techniques such as pruning and knowledge distillation will be applied to further lightweight the model, thereby continuing to reduce computational resource consumption and hardware dependency. Future work will also prioritize deployment experiments on edge devices to evaluate the model's real-time performance and stability in actual field environments.

ACKNOWLEDGEMENT

This study was funded by the Key Research and development Plan of Heilongjiang Province(2024ZXDXB45), (LJGXCG2022-107).

REFERENCES

[1] Sarah M. Alhammad, Doaa Sami Khafaga, Walaa M. El hady, Farid M. Samy & Khalid M. Hosny. (2025). Deep learning and explainable AI for classification of potato leaf diseases. *Frontiers in Artificial Intelligence,* Vol. 7*,* pp. 1449329-1449329, Riyadh/Saudi Arabia

[2] Gan B., Pu G., Xing W. (2025). Enhanced YOLOv8 with SDD weight and efficient detection head for detecting rice leaf diseases. *Scientific Reports*, Vol. 15, pp. 22179, London/UK.

[3] Han K., Wang Y., Tian Q. (2020). GhostNet: More features from cheap operations. *Conference on Computer Vision and Pattern Recognition*, pp. 1577-1586, Seattle/USA.

[4] Khanam R., Hussain M. (2024). YOLOv11: An overview of the key architectural enhancements. *arXiv preprint* arXiv:2410.17725, Ithaca/USA.

[5] Korade N.B., Salunke M.B., Bhosle A. (2025). Proactive Soybean Disease Detection through YOLO Leaf Extraction and ResNet-50 Classification to Reduce Crop Loss and Boost Productivity. *International Journal of Engineering Trends and Technology*, Vol. 73, pp. 133, Maharashtra/India.

[6] Lan Y., Sun B., Zhang L. (2024). Identification of ginger leaf diseases and pests in natural scenes based on improved YOLOv5s (基于改进 YOLOv5s 的自然场景下生姜叶片病虫害识别). *Transactions of the Chinese Society of Agricultural Engineering*, Vol. 40, pp. 210-216, Beijing/China.

[7] Lei M., Li S., Wu Y. (2025). YOLOv13: Real-Time Object Detection with Hypergraph-Enhanced Adaptive Visual Perception. *arXiv preprint* arXiv:2506.17733, Ithaca/USA.

[8] Li X., Wang W., Wu L. (2020). Generalized focal loss: Learning qualified and distributed bounding boxes for dense object detection. *Advances in Neural Information Processing Systems*, Vol. 33, pp. 21002-21012, Vancouver/Canada.

[9] Liu M., Liepold C., Swaminathan S. (2014). Progress in research on Asian soybean rust (亚洲大豆锈病研究进展). *Chinese Journal of Oil Crop Sciences*, Vol. 36, pp. 676-684, Wuhan/China.

[10] Liu S., Hu B., Zhao C. (2023). Detection and identification of cucumber leaf diseases and insect pests based on improved YOLOv7 (基于改进 YOLOv7 的黄瓜叶片病虫害检测与识别). *Transactions of the Chinese Society of Agricultural Engineering*, Vol. 39, pp. 163-171, Beijing/China.

[11] Liu W., Anguelov D., Erhan D. (2016). SSD: Single shot multibox detector. *European Conference on Computer Vision*, pp. 21-37, Cham/Switzerland.

[12] Ma X., Dai X., Bai Y. (2024). Rewrite the stars. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 5694-5703, Seattle/USA.

[13] Ren S., He K., Girshick R. (2016). Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 39, pp. 1137-1149, New York/USA.

[14] Sulaiman, N., Che'Ya, N. N., Mohd Roslim, M. H., Juraimi, A. S., Mohd Noor, N., & Fazlil Ilahi, W. F. (2022). The application of hyperspectral remote sensing imagery (HRSI) for weed detection analysis in rice fields: A review. *Applied Sciences*, Vol. 12(5), pp. 2570-2570, Serdang/Malaysia

[15] Sun, J., Feng, Z., Han, J., Xu, F., Zhang, H., & Guo, Y. (2025). YOLO-PLNet: a lightweight real-time detection model for peanut leaf diseases based on edge deployment. *Frontiers in Plant Science*, Vol. 16, pp. 1707501, Zhengzhou/China

[16] Tian Y., Ye Q., Doermann D. (2025). YOLOv12: Attention-centric real-time object detectors. *arXiv preprint* arXiv:2502.12524, Ithaca/USA.

[17] Wang J., Li S., Wang C. (2025). Lightweight tomato pest detection model based on improved YOLOv8n (基于改进 YOLOv8n 的轻量级番茄虫害检测模型). *Transactions of the Chinese Society of Agricultural Engineering*, Vol. 41, pp. 185-195, Beijing/China.

[18] Wu S., Mao Y., Hu H. (2025). Grape fruit and leaf disease detection method based on improved YOLO11n (基于改进 YOLO11n 的葡萄果叶病害检测方法). *Transactions of the Chinese Society of Agricultural Engineering*, Vol. 41, pp. 140-147, Beijing/China.

[19] Xiao C. (2024). Symptom identification and prevention methods of several common soybean diseases (大豆几种常见病害的症状识别及防治方法). *Modernizing Agriculture*, pp. 15-17, Heilongjiang/China.

[20] Xu Y., Wu Q., Zhang B. (2025). Application progress of lightweight deep learning network in crop object detection (轻量级深度学习网络在农作物目标检测的应用进展). *Journal of Chinese Agricultural Mechanization*, Vol. 46, pp. 261-270, Nanjing/China.

[21] Yang S., Zhang P., Wang L. (2025). Lightweight tomato leaf disease and pest identification method integrating improved YOLOv8n and channel pruning (集成改进 YOLOv8n 与通道剪枝的轻量化番茄叶片病虫害识别方法). *Transactions of the Chinese Society of Agricultural Engineering*, Vol. 41, pp. 206-214, Beijing/China.

[22] Zhu Linqi, Li Xiaoming, Sun Hongmin & Han Yingpeng. (2024). Research on CBF-YOLO detection model for common soybean pests in complex environment. *Computers and Electronics in Agriculture,* Vol. 216, pp. 108515, Harbin/China.