

ERL-YOLOv8n: A Potato Disease Detection Network for Complex Field Environments

Feng Gong^{1,2}, Yingsheng Chen^{2,*}

¹*School of Computer Science and Engineering, Chongqing Three Gorges University, Chongqing, China*

²*Big Data and Information Security Engineering Technology Research Center, Chongqing College of Humanities, Science and Technology, Chongqing, China*

**Corresponding Author*

Abstract: To address the challenge of balancing accuracy and efficiency in potato disease detection models under complex field conditions, this study proposes an improved lightweight detection algorithm, ERL-YOLOv8n. The algorithm optimizes the YOLOv8n model in three key aspects: First, the Receptive Field Attention Convolutional Block Attention Module (RFCBAMConv) is introduced into the backbone network to enhance the model's adaptive perception of disease spot features at various scales. Second, an EPSA (ECA and Polarized Self-Attention) module is embedded in the neck network to improve the model's anti-interference capability in complex backgrounds by fusing multi-dimensional feature information. Finally, the LSGE (Large Selective Kernel Network and Spatial Group-wise Enhance) attention mechanism is incorporated to synergistically optimize the backbone and neck, effectively improving detection accuracy for large-scale targets and complex scenes. Ablation and comparative experiments demonstrate that the ERL-YOLOv8n model achieves significant improvements in key performance metrics while maintaining high detection speed. Compared to the original YOLOv8n model, its precision, recall, and mAP@50 increased by 1.4, 5.7, and 2.3 percentage points, respectively. The improved model exhibits enhanced robustness and superior detection performance, particularly in the early identification of early and late blight, providing reliable technical support for practical applications in precision agriculture.

Keywords: Potato Disease; Object Detection; YOLOv8n; Deep Learning; Attention Mechanism.

1. Introduction

In recent years, China's potato industry has entered a new stage of rapid development. With the continuous optimization of cultivation techniques and rising market demand, both the planting scale and annual output of potatoes have maintained steady growth. It is noteworthy that during the process of industrial expansion, some production areas have encountered the problem of field management measures failing to keep pace, such as in nutrient management and pest and disease control during the crop growth stage. This extensive management model not only restricts the potential for yield improvement but also significantly increases the risk of crops being infected by various diseases [1], causing serious economic losses to growers. The conventional approach to managing potato diseases depends primarily on the empirical assessment of agricultural experts. This reliance on visual inspection, however, presents distinct drawbacks, including high labor costs, low detection efficiency, and unstable accuracy. These limitations are especially acute when addressing the monitoring needs of large-scale plantations. To break through this impasse, the field of agricultural technology has started to implement intelligent solutions [2].

Intelligent agricultural technology is now rapidly driving industrial change. Within the field of crop disease recognition, a complete technological ecosystem has gradually formed around deep learning. By building custom models and training them with massive volumes of leaf imagery, many research groups have successfully enabled the intelligent identification of crop diseases. In the specific niche of potato disease detection, academia has seen the emergence of numerous groundbreaking research outcomes:

To address the challenges of late blight detection, the team led by Wang Linbai innovatively optimized the CenterNet-SPP model architecture. By constructing a multi-level feature extraction network, their solution achieves synergistic optimization of disease center point localization and target size regression, significantly enhancing detection sensitivity [3]. Taking a different approach, Niu Yuxia et al. developed a layered feature alignment network. This method achieved a breakthrough in few-shot learning scenarios by fusing textual semantics and visual features through multi-modal mapping technology [4]. Of particular note, the team led by Sun Jianming combined the Efficient Net v2 network with an attention mechanism. Through the synergistic application of a pyramid feature fusion strategy and a visual Transformer encoder module, their model achieved an accuracy of 98.26%, setting a new benchmark in the field [5]. Regarding real-time detection, the contributions of Liu Kaiqi with the YOLOv3 framework are noteworthy. By incorporating a Spatial Pyramid Pooling module, his YOLOv3-SPP algorithm markedly enhances accuracy without compromising detection speed [6]. Meanwhile, the latest YOLOv8n-Potato algorithm from Zeng Liang's team marks a two-fold achievement. By redesigning the feature fusion network with a CAA-HS-FPN architecture and combining it with a lightweight detection head and an optimized PIoU loss function, the model's precision metric improved by 2.4 percentage points, even as the number of parameters was reduced by 42% [7].

Although these innovations have significantly improved detection accuracy in laboratory environments, critical challenges remain in practical application scenarios. Field operational environments are subject to uncontrollable factors such as variations in lighting and complex background interference. Consequently, the performance advantages of existing models on idealized datasets are difficult to fully translate to real-world planting scenarios [8]. To address this industry pain point, this study focuses on overcoming the challenge of disease recognition in complex environments. By optimizing the architecture of the YOLOv8 model, we develop a detection system suitable for open-air cultivation environments. The algorithm is specifically adapted for two highly prevalent diseases, early blight and late blight, to provide a reliable technical solution for practical

agricultural production.

2. Materials and Methods

2.1 Test Environment and Parameter

The experiments were conducted on a server equipped with an Intel Xeon Platinum 8481C CPU operating at 2.70GHz, an NVIDIA GeForce RTX 4090D graphics card (GPU), and 24 GB of RAM. The algorithm was implemented in Python. During training, the batch size was set to 32, the number of epochs to 220, and the initial learning rate to 0.001. The input images were resized to a resolution of 640×640 pixels.

2.2 Dataset Acquisition and Preprocessing

The dataset used in this study was sourced from the publicly available Plant Village dataset. A total of 2000 images were selected, encompassing three categories: early blight leaves, healthy leaves, and late blight leaves. Subsequently, the labeling tool was used to annotate these three classes of data. The dataset was then partitioned into training, testing, and validation sets according to a 7:2:1 ratio. Data augmentation was performed to expand the dataset, a process that enhances the model's training performance and robustness. As shown in Figure 1, the augmentation process resulted in a final dataset of 4000 images.



Figure 1. Original Diseased Leaf Images and Images Generated by Data Augmentation

2.3 Evaluation Metrics

The evaluation metrics used in this experiment are Precision, Recall, mAP50, mAP50-95, and GFLOPs (to measure model computational complexity). The formula for Precision (P) is expressed as:

$$P = \frac{TP}{TP + FP} \quad (1)$$

The formula for Recall (R) is expressed as:

$$R = \frac{TP}{TP + FN} \quad (2)$$

The mean Average Precision (mAP)

metric mAP50 represents the average precision calculated at an Intersection over Union (IoU) threshold of 50%. In contrast, mAP50-95 represents the mAP value averaged over a range of IoU thresholds from 50% to 95%. The formula for its calculation can be expressed as:

$$\text{mAP50} = \frac{1}{C} \sum_{j=1}^C \frac{\text{TP}}{\text{TP} + \text{FN}} \# \quad (3)$$

2.4 Research Methods

2.4.1 The YOLOv8 Model

YOLOv8 is a detection algorithm from the YOLO family, developed by Ultralytics for tasks such as object detection and instance segmentation. It features a three-part decoupled architecture ("Backbone-Neck-Head"). Its backbone significantly improves feature extraction by using the C2f module with cross-stage dense connections and an adaptive Meta-ACON activation function. Meanwhile, the head introduces a dynamic decoupled design, separating classification from regression and utilizing an anchor-free mechanism. Its loss function is bifurcated, comprising a BCE Loss for classification and a combination of Distribution Focal Loss and CIoU Loss for regression. Our work builds upon the YOLOv8n model, introducing improvements primarily to its convolutional modules, as well as its backbone and neck architectures.

2.4.2 The ERL-YOLOv8n Algorithm

YOLOv8n is a lightweight object detection model that, despite its advantages, struggles with high error rates and unstable performance when detecting potato diseases in challenging field conditions like fluctuating light and leaf occlusion. Nevertheless, its end-to-end architecture and decomposable modules offer ample room for enhancement. Based on this potential, we introduce ERL-YOLOv8n, an improved model for potato leaf disease detection. As shown in Figure 2, the ERL-YOLOv8n network features three key architectural upgrades. To better capture complex features, we replace the backbone's standard Conv layers with RFCBAMConv [9] modules. To boost detection accuracy in complex scenarios, we introduce the EPSA attention mechanism into the neck. Finally, to improve accuracy on large targets and in cluttered backgrounds without sacrificing real-time speed, we insert the LSQE attention mechanism between the backbone and neck. As a result, our approach achieves higher accuracy at a lower computational cost,

demonstrating superior performance in the detection of potato leaf diseases.

2.4.3 The EPSA Module

In object detection tasks, environmental interference poses a significant challenge that cannot be overlooked. To enable the model to adapt to detection in various complex environments and thereby enhance its accuracy, this study designed a novel attention mechanism, EPSA (ECA and Polarized Self-Attention). This mechanism was developed by integrating the advantages of two existing attention mechanisms: ECA [10] (Efficient Channel Attention) and PSA [11] (Polarized Self-Attention). A key characteristic of EPSA is its ability to balance multi-dimensional feature capture with computational efficiency.

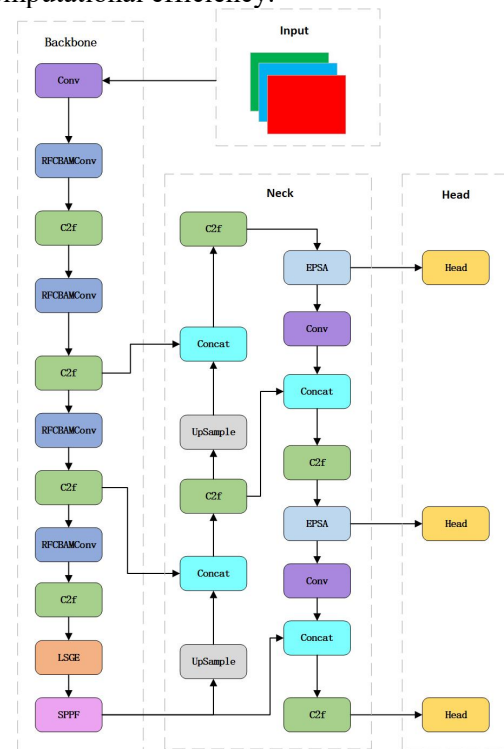


Figure 2. Architecture of the ERL-YOLOv8n Model

The EPSA attention mechanism optimizes image feature extraction by combining the advantages of ECA and PSA. The ECA module uses adaptive average pooling and a one-dimensional convolution to capture inter-channel dependencies. In contrast, the PSA module captures multi-dimensional features through channel and spatial attention, but at a high computational cost. EPSA strikes a balance between these two approaches by using a learnable weight parameter (α) to achieve comprehensive feature extraction. While ECA is known for its efficiency, its granularity is limited;

PSA offers comprehensive feature extraction but incurs significant computational overhead. By fusing these two mechanisms and adjusting the alpha parameter, EPSA optimizes the computational load, achieving a balance between efficiency and performance.

The EPSA attention mechanism achieves comprehensive and rich feature extraction by

integrating the advantages of ECA and PSA, all while preserving computational efficiency. This mechanism is incorporated into the ninth layer of the backbone network to enhance the efficiency and accuracy of feature extraction. The architecture of the module is illustrated in Figure 3.

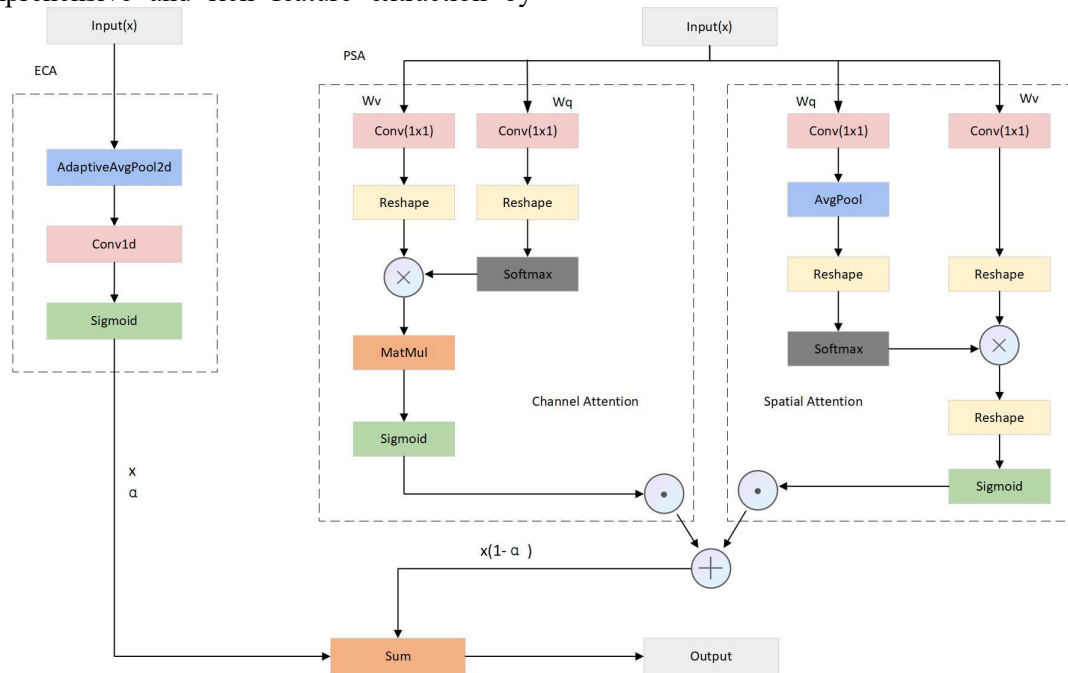


Figure 3. Architecture of the EPSA Module

First, within the ECA module, the global context is captured by applying global average pooling to the input feature map. A one-dimensional convolution is then used to fuse channel information, with the kernel size being determined by an adaptive function. Subsequently, a Sigmoid activation function is applied to generate the weights for each channel. In parallel, the PSA module generates spatial attention weights and a new feature map using its own 1D convolution, global pooling operations, and a Sigmoid function. Finally, the output from PSA is dynamically fused with the output from ECA, governed by a learnable weight. The corresponding mathematical formula is shown below.

$$F_{\text{EPSA}}(X) = \alpha \cdot F_{\text{ECA}}(X) + (1-\alpha) \cdot F_{\text{PSA}}(X) \quad (4)$$

2.4.4 The RFCBAM Module

This research addresses two primary challenges. First, disease samples collected in the field often exhibit indistinct feature boundaries and complex morphological variations. Second, conventional convolution operations have distinct limitations during the feature extraction stage [12]. To overcome these technical

bottlenecks, this study incorporates an adaptive perceptual convolution mechanism into the backbone network architecture. This technical approach, which integrates multi-scale receptive field optimization with dual-dimension attention regulation, significantly enhances the model's analytical capability for complex pathological signs while controlling computational resource consumption.

The performance of conventional Convolutional Neural Networks (CNNs) in disease recognition is primarily constrained by two factors. First, the fixed parameters of convolutional kernels struggle to adapt to the diverse feature variations across different diseased regions, resulting in insufficient sensitivity for capturing subtle pathological signs. Second, the weight allocation method, often based on global averaging, dilutes the representational strength of critical disease regions [13]. This dual limitation frequently causes traditional models to suffer from the loss of fine-grained features and the misclassification of key areas during leaf spot detection.

Therefore, Receptive Field Attention Convolution is introduced. By leveraging the

synergistic interaction of spatial perception and channel attention mechanisms, it constructs dynamic convolutional kernels with region-adaptive capabilities. This, in turn, enhances the

model's ability to extract multi-scale features from complex diseases. Its architecture is illustrated in Figure 4.

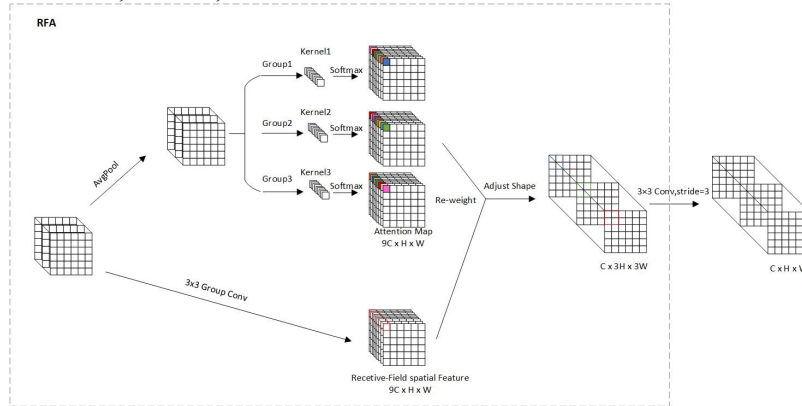


Figure 4. Architecture of the RFACnv Module

First, the input feature map is down sampled using an average pooling layer to reduce computational complexity while preserving principal features. Subsequently, a grouped convolution layer performs feature extraction on the pooled feature map; by dividing channels into multiple groups, grouped convolution enhances the model's representational capacity. The resulting feature maps are then divided into groups, where each group interacts with its corresponding kernel and generates an attention weight map via a softmax function. These weight maps indicate the importance of different spatial locations. The attention map is then multiplied element-wise with the original receptive field features, which re-weights the original features to emphasize important regions. Following this, the shape of the re-weighted feature map is adjusted to accommodate subsequent convolutional operations. Finally, the receptive field's spatial features are processed based on the weights from the attention map, then resized to the appropriate dimensions to produce the final output of the Receptive Field Attention Convolution. The computational process of the Receptive Field Attention Convolution is detailed in Equation (5).

$$F \&= \text{Softmax} \left(g^{1 \times 1} (\text{AvgPool}(X)) \right) \times \text{ReLU} \left(\text{Norm} (g^{k \times k} (X)) \right) \quad (5)$$

$$\&= A_{rf} \times F_{rf}$$

In Equation (5), $g^{1 \times 1}$ represents a grouped convolution of size $i \times i$, k is the size of the convolution kernel, and Norm denotes the normalization function. X is the input feature map, and the output F is obtained by multiplying the attention map A_{rf} with the transformed receptive field spatial features F_{rf} .

Spatial attention mechanisms focus on receptiv

e field spatial features and, when combined with aggregation mechanisms, overcome the problem of shared parameters in standard convolutions. Existing spatial attention mechanisms, such as the Convolutional Block Attention Module [14] (CBAM), already account for long-range information by generating a global context through global average pooling or global max pooling. Therefore, this study introduces RFCBAM Conv, a Receptive Field Attention Convolution module that combines RFA with CBAM. Its architecture is illustrated in Figure 5.

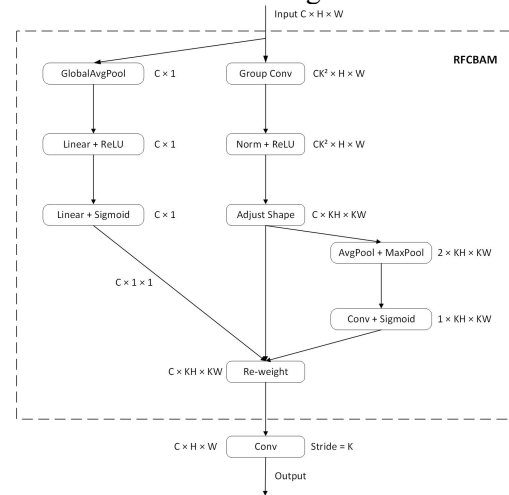


Figure 5. Architecture of the RFCBAMConv Module

2.4.5 The LSGE Module

To address key challenges in detecting leaf spots against complex backgrounds, such as significant scale variations and high levels of background interference, this study proposes a leaf disease detection method based on locally sensitive group enhancement. By fusing multi-scale feature optimization with a dynamic

attention mechanism, this method overcomes the performance bottlenecks of traditional models that rely on single-path attention mechanisms. It provides a high-precision,

lightweight solution for disease detection in complex agricultural scenarios [15]. The architecture of the LSGE module is illustrated in Figure 6.

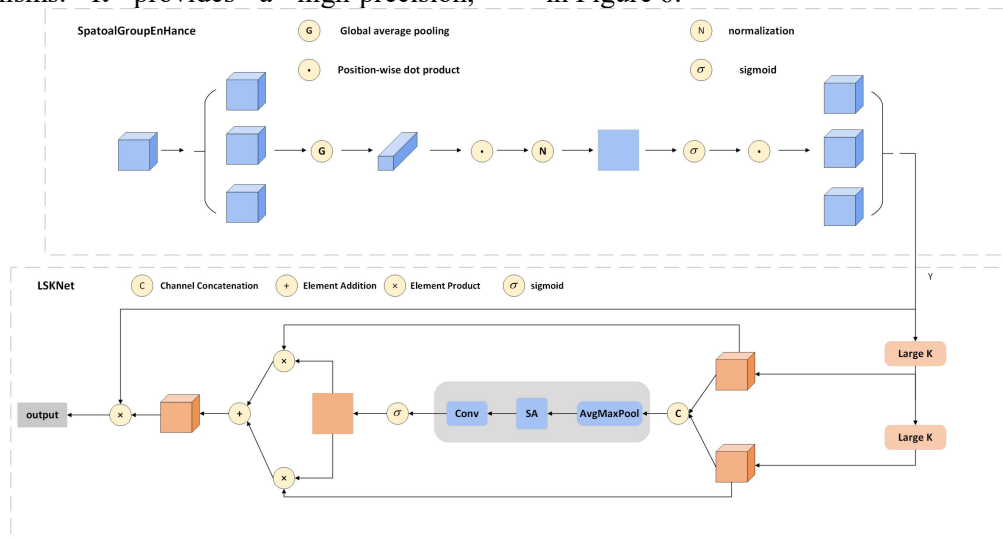


Figure 6. Architecture of the LSGE Module

This study combines the Spatial Group-wise Enhance [16] (SGE) and the Large Selective Kernel Network [17] (LSKNet) mechanisms to enhance the model's ability to capture and understand image features.

First, a grouped convolution operation is performed on the input feature map. By partitioning the input feature map into multiple groups and applying convolution operations independently within each group, this process generates several sets of intra-group feature maps, achieving an initial separation and extraction of features.

Subsequently, Global Average Pooling (GAP) is applied to each intra-group feature map. This operation compresses each feature map along its spatial dimensions by calculating the average value for each channel. The result is a corresponding global feature vector, which achieves both dimensionality reduction and the aggregation of global information.

Next, a dot product operation is performed between the global feature vector and the original feature map to generate attention weights. This operation highlights key feature regions by establishing a correlation between global information and local features. To ensure the weights conform to a probability distribution, the generated attention weights are sequentially processed with normalization and a Sigmoid activation function. Subsequently, the processed attention weights are applied to the original feature map via a multiplication operation,

which enhances the feature map's capacity to represent key information.

During the feature fusion stage, the feature maps from the multiple groups are first concatenated along the channel dimension. On this basis, a large-kernel convolution is applied. By expanding the receptive field, this operation captures a broader range of contextual information, thereby enhancing the network's ability to understand global semantics.

To further enhance the representational capacity of the feature map, a self-attention mechanism is introduced. This mechanism adaptively adjusts the weights of different feature components by calculating the degree of correlation between elements within the feature map. Concurrently, both average pooling and max pooling operations are utilized to capture information from the feature map from different perspectives. Average pooling preserves the overall, holistic information of the features, while max pooling highlights the most salient local information.

Finally, the different types of features acquired through the aforementioned operations are fused and then processed by a Sigmoid activation function to generate the final attention weights. These weights are applied to the original feature map to produce the final output feature map. Through this design, which combines multi-scale feature capture with an efficient attention mechanism, the model's performance in image recognition tasks is effectively enhanced [18].

Table 1. Ablation Study Results

Model	EPSA	LSGE	RFCBAM	Precision (%)	Recall (%)	mAP@50(%)	mAP@50~95(%)
YOLOv8n	×	×	×	89.7	70.0	84.3	69.6
1	√	×	×	89.9	73.5	85.3	69.6
2	×	√	×	90.3	72.4	84.9	70.1
3	×	×	√	87.0	78.3	84.3	69.2
4	√	√	×	90.0	72.1	83.6	69.5
5	√	×	√	88.7	73.8	86.1	69.8
6	×	√	√	89.1	75.2	85.0	70.1
ERL- YOLOv8n	√	√	√	91.1	75.7	86.6	70.6

3. Results and Discussion

3.1. Ablation Study

To validate the effectiveness of the individual improved modules and their final combination in the ERL-YOLOv8n model, an ablation study was designed. In this study, comparative experiments were conducted using different combinations of the three proposed modules. To ensure the fairness of the evaluation, all experiments were performed on the same dataset and with the same hyperparameters.

The results of the ablation study are presented in Table 1. Compared to the original YOLOv8n, the ERL-YOLOv8n model—which simultaneously incorporates the RFCBAMConv, EPSA, and LSGE modules—achieved improvements of 1.4, 5.7, 2.3, and 1.0 percentage points on the four key metrics of

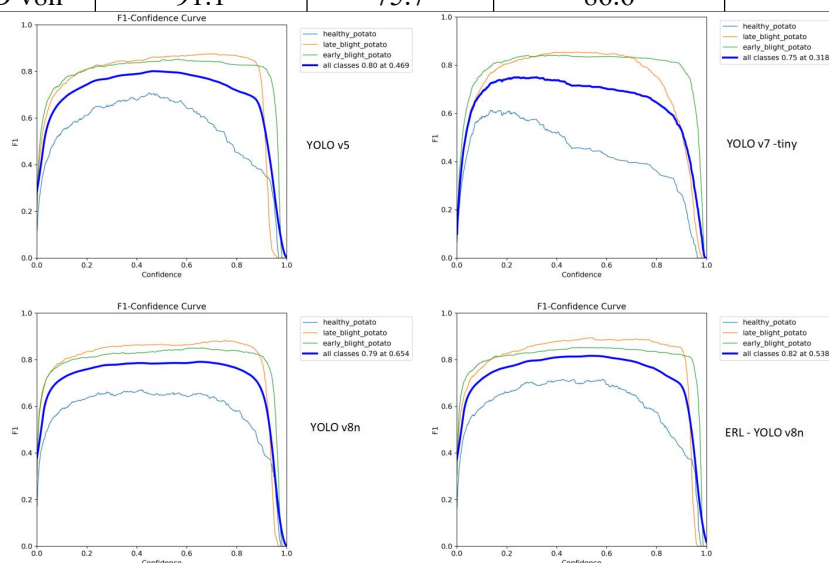
Precision, Recall, mAP@50, and mAP@50~95, respectively, after 220 epochs of training. This demonstrates that the improved algorithm proposed in this study, ERL-YOLOv8n, achieves significant results in practical applications.

3.2 Comparative Experiments

To evaluate the performance of the ERL-YOLOv8n algorithm for the task of potato leaf disease detection, we conducted a comparative analysis against several other models: YOLOv5, YOLOv7-tiny, and the original YOLOv8n. All experiments were conducted under identical conditions, including the same environmental setup, dataset, and hyperparameters. The performance differences are summarized in Table 2, and their respective F1-Confidence curves are compared in Figure 7.

Table 2. Comparative Experiment Results

Model	Precision (%)	Recall (%)	mAP@50 (%)	mAP@50~95 (%)
YOLO v5	89.1	73.1	84.6	68.7
YOLO v7-tiny	87.1	67.1	79.4	61.2
YOLO v8n	89.7	70.0	84.3	69.6
ERL - YOLO v8n	91.1	75.7	86.6	70.6

**Figure 7. Comparison of F1-Confidence Curves for the Different Models**

3.3 Detection Results

To validate the effectiveness of the algorithm's improvements, this study conducted a comparative experiment using samples of three common potato leaf diseases. As illustrated in Figure 8, when tested on the same dataset, the ERL-YOLOv8n model demonstrates superior identification capabilities compared to the original YOLOv8n model. The experimental data show that the improved model performs better in recognizing complex morphologies and small-sized lesions. For the late blight samples, the original model was prone to misclassifying leaves with tiny, early-stage brown spots as healthy. In contrast, the new model can effectively capture these subtle features, significantly reducing missed detections of early-stage lesions. For the early blight samples, which are often characterized by indistinct lesion borders and irregular shapes (especially in the initial stages of infection), the recognition accuracy of the improved model was 12.6 percentage points higher than the original model.

This indicates that, under complex field conditions, the new model can more reliably identify leaves exhibiting the early signs of early blight.

In-depth analysis reveals that the improved model, through its optimized feature extraction mechanism, captures detailed features such as leaf texture and spot distribution in its shallow layers, while simultaneously enhancing its ability to process mid-level semantic information. This dual optimization enables the model to maintain strong feature discrimination capabilities even against complex backgrounds. This is particularly evident when dealing with the semi-transparent lesions of early-stage late blight, where the model can achieve accurate judgments by cross-validating features across multiple dimensions. The comparative experiments confirm that the improved network architecture effectively boosts the recall rate for small object detection, offering a new approach to resolving misclassification issues in agricultural image detection.

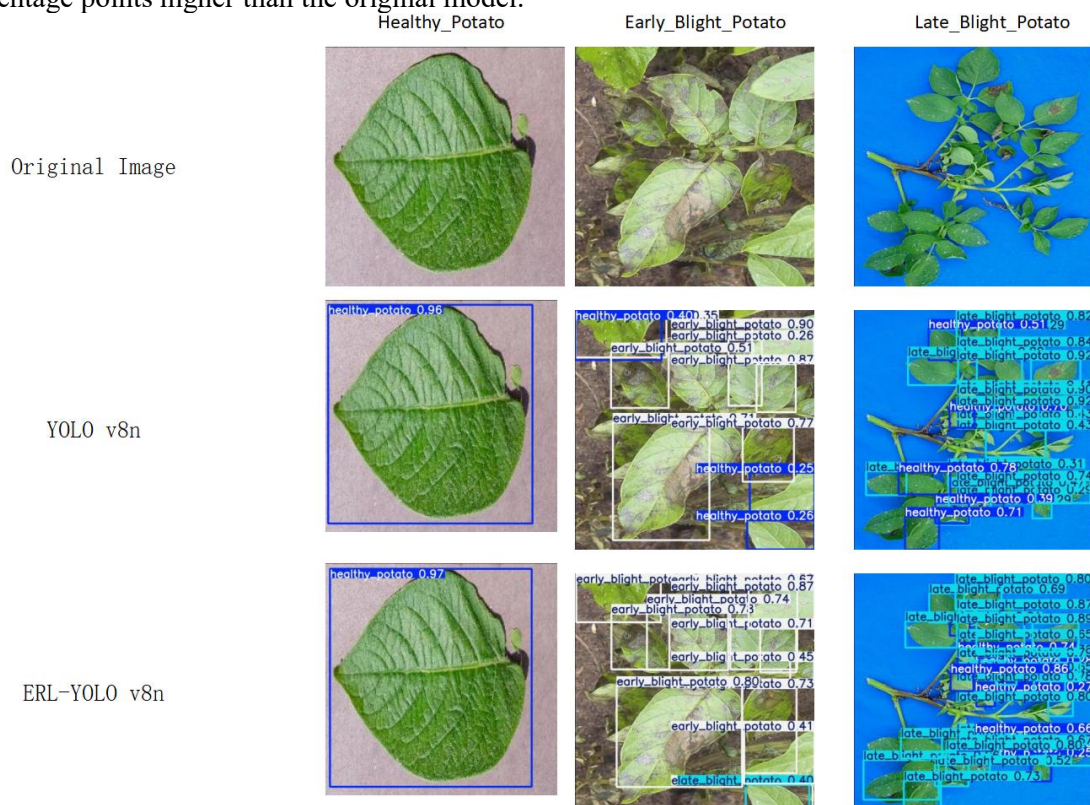


Figure 8. Comparison of Disease Detection Results

4. Conclusion and Discussion

Targeting the critical industry issue of balancing accuracy and efficiency in agricultural disease recognition, this research

proposes an innovative detection framework named ERL-YOLOv8n. Through a systematic optimization of the existing algorithm architecture, the interaction between the attention mechanism and the feature extraction

module was fundamentally redesigned. Experimental results demonstrate that the model, enhanced by synergistic multi-module optimization, exhibits significant advantages in complex field scenarios. It not only effectively captures minute lesion features but also shows particularly notable improvements in detection accuracy under conditions of leaf occlusion. Compared to the baseline model, the improved ERL-YOLOv8n achieves a breakthrough 2.3 percentage point increase in the mean Average Precision metric while maintaining computational efficiency, thus validating the effectiveness of the structural optimization strategy. This research has achieved positive results in improving the YOLOv8n model, demonstrating superior performance on the potato leaf disease detection task. These promising outcomes provide robust support for the future development of object detection and offer a more accurate and efficient solution for practical applications.

Acknowledgments

This work was supported by the University-level Research Project of Chongqing College of Humanities, Science and Technology (Grant No. CQRKZK202005).

References

- [1] Yuan Y J. Potato Planting Technology and Common Pest and Disease Control Strategies—A Case Study of Ningjin County. *Southern Agricultural Machinery*, 2023, 54(06): 66-68, 86.
- [2] Javed R, Imran K, Ghulam A, et al. Multi-Level Deep Learning Model for Potato Leaf Disease Recognition. *Electronics*, 2021, 10(17): 2064-2064.
- [3] Wang L B, Zhang B, Yao J F, et al. Potato Leaf Disease Recognition and Lesion Detection Based on Convolutional Neural Networks. *Journal of Chinese Agricultural Mechanization*, 2021, 42(11): 122-129.
- [4] Niu Y X, Sun Z H, Ren W, et al. Few-Shot Potato Disease Leaf Detection Based on a Layered Feature Alignment Network. *Journal of Chinese Agricultural Mechanization*, 2024, 45(02): 250-258.
- [5] Sun J M, Bi Z Y, Niu L D. Recognition of Potato Leaf Diseases Using an EfficientNet v2 Network with a Feature Fusion Transformer. *Jiangsu Agricultural Sciences*, 2024, 52(08): 166-176.
- [6] Liu K Q. Research on Potato Leaf Disease Recognition Based on YOLOv3-SPP. *Automation Application*, 2022, (06): 21-24+28. DOI: 10.19769/j.zdhy.2022.06.006.
- [7] Zeng L, Peng Y. Potato Leaf Disease Detection Algorithm Based on Improved YOLOv8. *Journal of Luoyang Institute of Science and Technology (Natural Science Edition)*, 2024, 34(03): 62-69.
- [8] Gao X, Zhang Y. Detection of Fruit using YOLOv8-based Single Stage Detectors. *International Journal of Advanced Computer Science and Applications (IJACSA)*, 2023, 14(12).
- [9] Zhang X, Liu C, Yang D, et al. RFACConv: Innovating spatial attention and standard convolutional operation. *arXiv preprint arXiv:2304.03198*, 2023.
- [10] Q. W, B. W, P.Z, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2020, 11531-11539.
- [11] Huajun L, Fuqiang L, Xinyi F, et al. Polarized self-attention: Towards high-quality pixel-wise mapping. *Neurocomputing*, 2022, 506: 158-167.
- [12] Zhong T. Research on Methods for Crop Leaf Disease Type Identification and Detection Based on Deep Learning. *Tianjin University of Technology*, 2024. DOI: 10.27360/d.cnki.gtlgy.2024.001314.
- [13] Zhou Y F, Liu D Y, Zhou Y P. Crop Disease Leaf Detection Based on Multimodal Feature Alignment. *Journal of Chinese Agricultural Mechanization*, 2024, 45(07): 180-187. DOI: 10.13733/j.jcam.issn.2095-5553.2024.07.027.
- [14] Woo S, Park J, Lee J Y, et al. CBAM: convolutional block attention module // *Lecture Notes in Computer Science*. Cham: Springer International Publishing, 2018: 3-19.
- [15] Wang Z T, Zou Y B, Wu C L, et al. Beyond Single Perception: An Agricultural Pest Detection Algorithm MRA-YOLOX. *Computer Engineering and Applications*, 2024, 60(16): 206-216.
- [16] Li X, Hu X, Yang J. Spatial Group-wise Enhance: Improving Semantic Feature Learning in Convolutional Networks.

- CoRR, 2019, abs/1905.09646
- [17]Y. Li, Q. Hou, Z. Zheng, M. -M. Cheng, J. Yang and X. Li, "Large Selective Kernel Network for Remote Sensing Object Detection," 2023 IEEE/CVF International Conference on Computer Vision (ICCV), Paris, France, 2023, pp. 16748-16759, DOI: 10.1109/ICCV51070.2023.01540.
- [18]Nagachandrika B, Prasath R, Joe P I. An automatic classification framework for identifying type of plant leaf diseases using multi-scale feature fusion-based adaptive deep network. Biomedical Signal Processing and Control, 2024, 95(PA): 106316-.