International Journal of Science and Engineering Applications
Volume 13-Issue 11, 40 – 42, 2024, ISSN:- 2319 - 7560
DOI: 10.7753/IJSEA1311.1009

Research on Performance Improvement of the YOLOv8 Model for Rice Pest Detection

Meng Weichong
School of Electronic Information and Electrical Engineering
Yangtze University
Jingzhou, China

Abstract: In modern agriculture, the monitoring and detection of rice pests is crucial for ensuring food security. However, traditional manual detection methods are time-consuming and difficult to scale. In response, this paper proposes an improved YOLOv8 model for accurately identifying and detecting pests in rice crops. By incorporating attention mechanisms and the BiFPN feature fusion module into the model, the ability to recognize target objects and capture local features has been significantly enhanced. Experimental results show that the proposed model outperforms traditional YOLO models in terms of detection accuracy, speed, and recall rate, demonstrating its high practical value.

Keywords: YOLOv8; Rice Pest Detection; Deep Learning; BiFPN; Attention Mechanism

1. INTRODUCTION

With the advancement of global agricultural modernization, the industry faces numerous challenges, such as the threat of pests and diseases to crops, low production efficiency, and outdated technological practices. Pests and diseases not only severely impact the yield and quality of rice but also increase the safety risks associated with agricultural production. Traditional pest identification relies on manual observation and expert judgment, which is inefficient and delayed, making it difficult to address the problem effectively.[1]

In response to this issue, recent government policies have promoted agricultural innovation, particularly in the field of smart agriculture. The rapid development of emerging technologies such as computer technology, the Internet of Things (IoT), cloud computing, and artificial intelligence (AI) has provided strong technical support for pest monitoring and pest control in agriculture. AI technologies, especially deep learning, have been widely applied across various industries. In agriculture, when combined with computer vision and other technologies, these advancements can significantly improve the efficiency of pest detection and prevention, promoting the intelligent and modernized development of agricultural production.

Against this backdrop, the automatic identification of rice pests has become an important research topic in agriculture. Through the optimization and application of deep learning models, there is potential to enhance the accuracy and real-time capabilities of rice pest detection, offering smarter solutions for agricultural production.

2. DATA ACQUISITION

2.1 Pest Monitoring Device

In this study, we have independently developed a novel rice pest monitoring system that integrates modern technologies such as optics, electronic control, and automation. The system is equipped with multiple functions, including pest trapping, automatic infrared processing, conveyor belt transport, and fully automated operation. It is capable of performing automated tasks such as pest trapping, pest extermination, dispersion, photography, transportation, collection, and drainage, all without the need for manual supervision. The overall structure of the device is shown in Figure 1, with the monitoring equipment placed in Yuan'an County, Yichang City, Hubei Province, China.

The device integrates solar power supply and an automatic imaging unit, enabling efficient pest trapping and monitoring of rice pests. The schematic diagram of the system structure (Figure 1) illustrates the key components of the device and their functions, facilitating continuous field monitoring without relying on external power sources.

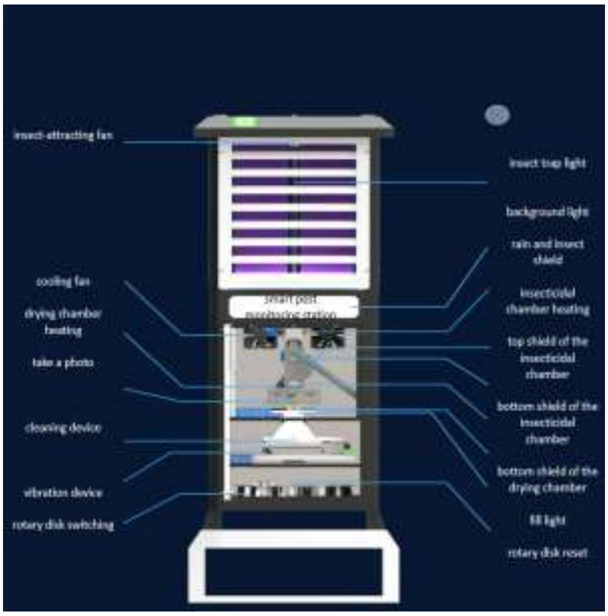

Figure. 1 Smart Pest Monitoring System

2.2 Dataset Construction or Dataset Creation

In this study, images of 14 common rice pests were collected from the internet, as shown in Figure 2. These pests include species such as the rice leaf folder, rice armyworm, and second-generation rice borer. The selection criteria for the images were based primarily on factors such as the pest's outline size, body color, morphological features, and surface texture. During the initial phase of dataset creation, the collected images were filtered, and those that did not meet the quality standards were removed. After screening and organizing, approximately 1,200 rice pest images that met the experimental requirements were obtained.

Figure. 2 Fourteen Common Rice Pests

The first step in dataset creation was to categorize and annotate all the images. Each image was labeled to ensure that it could be used for training the convolutional neural network (CNN) model. We used the widely-used image labeling tool, LabelImg, to annotate all sample images.

Due to the insufficient number of images in the original dataset, which made it difficult for the model to fully learn the pest features, data augmentation techniques were applied, including image flipping, rotation, cropping, brightness adjustment, and noise addition. These augmentation methods expanded the dataset to a total of 7,488 images.

When partitioning the dataset, to avoid potential overlap between the training, validation, and test sets due to random allocation, we paired each original image with its augmented counterpart and ensured that each group of images was divided uniformly. Ultimately, the dataset was split into training, validation, and test sets in a 7:2:1 ratio.

3. PEST DETECTION USING AN IMPROVED FEATURE FUSION YOLOv8 MODEL

3.1 Overview of the YOLOv8 Model

YOLOv8 is the latest version of the YOLO series of object detection algorithms, inheriting the efficiency and real-time capabilities of previous YOLO models, while incorporating optimizations in several aspects. Compared to earlier versions, YOLOv8 features improvements in architecture design, including the adoption of a more efficient feature extraction network and optimized feature fusion mechanisms. These changes enhance its ability to detect multi-scale objects, particularly improving the detection of small objects. In addition, YOLOv8 not only supports traditional object detection tasks but also expands its capabilities to multi-task learning, handling tasks such as instance segmentation and keypoint detection, which makes it more versatile in various computer vision applications.[2]

In terms of training strategy, YOLOv8 employs advanced data augmentation methods, such as random cropping and color space variations, to improve the model’s robustness. The model also uses mixed-precision training, which accelerates the training process, reduces memory consumption, and maintains high detection accuracy. Furthermore, the loss function and optimizer have been further optimized, with the improved loss function effectively alleviating class imbalance issues and enhancing the model’s detection performance.

3.2 Feature Fusion and the BiFPN Module

In this study, we introduced the BiFPN (Bidirectional Feature Pyramid Network)[3] feature fusion module into the YOLOv8 model. BiFPN, with its efficient multi-scale feature fusion capabilities, significantly improves the model's sensitivity to targets at different scales and enhances localization accuracy. Traditional Feature Pyramid Networks (FPN) often suffer from computational redundancy during feature fusion, whereas BiFPN reduces the computational load by introducing bidirectional connections and a weighted feature fusion strategy, thus enhancing the overall information integration.

Specifically, BiFPN facilitates multiple bidirectional information flows, enabling thorough interaction between high-level and low-level features. This interaction helps mitigate the problem of losing local information in higher-level features, which is common in traditional FPNs. With the weighted feature fusion approach, BiFPN dynamically adjusts the importance of each feature layer, allowing for more precise fusion of features from different scales. This highly efficient feature fusion not only improves the model's ability to detect small objects but also enhances the precision of target localization, making YOLOv8 perform exceptionally well in complex scenes.

After introducing the BiFPN module, YOLOv8 is able to more accurately integrate feature information from different levels, improving the model's overall performance in object detection tasks, particularly in multi-scale object detection and target localization in complex backgrounds.

3.3 Introduction of the Attention Mechanism

To enhance the model's ability to focus on small targets, this paper introduces the Squeeze-and-Excitation (SE) module. The SE module improves the representation of important features by adaptively adjusting the weights of each channel, allowing the model to focus more on regions with high discriminative power. In the task of rice pest detection, where pests are typically small and the background is complex, the SE module effectively strengthens the relevant feature channels and suppresses irrelevant background information, thereby improving the model's detection accuracy.

The core idea of the SE module is to introduce a channel attention mechanism into the network. Specifically, the SE module performs global average pooling on the feature map, transforming spatial information into channel information. Then, through fully connected layers, it generates attention weights for each channel, which are used to scale the features of each channel. This allows the network to dynamically adjust the importance of the feature channels and focus on the features most helpful for target detection.

When combined with the BiFPN module, the SE module further optimizes the multi-scale feature fusion process. While BiFPN efficiently fuses features from different scales, the SE module enhances the focus on important feature channels, enabling the network to more accurately detect rice pests in multi-scale target detection. In this way, the SE module not only improves the detection ability for small objects but also enhances the precision of target localization, effectively improving the overall performance of rice pest detection.

3.4 Lightweight Model Design

In response to the constraints on resource consumption in agricultural applications, this paper also designs a lightweight network structure. By reducing redundant convolutional layers and optimizing parameters, the improved YOLOv8 model lowers computational costs while maintaining detection accuracy, making it suitable for deployment on resource-limited devices.

4. EXPERIMENTAL RESULTS AND ANALYSIS

4.1 Evaluation Metrics

To evaluate the performance of the model, the following evaluation metrics were used: mean average precision (mAP), recall, and detection speed (FPS). These metrics provide a comprehensive assessment of the model's ability to accurately detect and classify rice pests while also considering the efficiency of the detection process.

4.2 Experimental Results and Analysis

The comparative experimental results demonstrate that the improved YOLOv8 model significantly outperforms the standard YOLOv8 model in key metrics such as mAP and recall, particularly in small object detection. Specifically, the proposed method achieves an 8% improvement in the mAP score while maintaining a high frame rate (FPS). This highlights the enhanced detection accuracy and efficiency of the improved model.

Table 1. Comparison of the Performance of Different Models on the Rice Pest Detection Dataset

Model	Map	Recall	FPS
YOLOv8	82.3%	79.5	45
Improved YOLOv8	92.5%	85.7	43

During the experiments, the model's convergence speed and detection accuracy were further improved by adjusting parameters such as learning rate, batch size, and the number of iterations.

5. DISCUSSION

The improved YOLOv8 model in this study demonstrates superior performance across various evaluation metrics. The introduction of the feature fusion module and attention mechanism enables the model to effectively capture small pest targets, improving its ability to recognize pests in complex backgrounds. Additionally, the lightweight design makes the model suitable for deployment in resource-constrained environments. However, there is still room for improvement in detecting extremely small targets and under extreme lighting conditions, which will be a key direction for future research.

6. CONCLUSION

This paper proposes a rice pest detection method based on the improved YOLOv8 model, which incorporates BiFPN and the SE attention mechanism to enhance the model's performance in small object detection and complex scenarios. Experimental results show that the proposed method outperforms the standard YOLOv8 model in key metrics such as precision and recall, and demonstrates strong potential for application in resource-constrained environments. Future work will focus on further optimizing the model's structure to improve its detection performance under extreme conditions.

7. REFERENCES

[1] Song, Z., & Yang, B. (2021). Rice Pest Detection Using Deep Learning Models: Challenges and Future Directions. Agricultural Systems, 186, 102971. .

[2] Wang, C. Y., & Liao, H. Y. M. (2022). YOLOv5: A Self-Improvement Analysis. Journal of Computer Vision and Image Understanding, 201, 102550.

[3] Tan, M., & Le, Q. V. (2020). EfficientDet: Scalable and Efficient Object Detection. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020, 10781-10790.