# Agricultural Pest Detection Based on the Fusion of CR_ULK_YOLO Algorithm

Weijiafa
School of Electronic Information and Electrical Engineering
Yangtze University
HuBei, China

**Abstract**: In response to the unsatisfactory performance of YOLOv8n in detecting pests under complex backgrounds and dense small-object scenarios, as well as its relatively slow convergence during training, this paper proposes an improved network model named CR_ULK_YOLO (CreToNext_SPPF_UnirepLK_YOLO). Based on the YOLOv8n framework, the proposed model integrates the concept of combining large-kernel and small-kernel convolutions. A novel SPPF_UnirepLK spatial pyramid network with ultra-large kernels is designed to enhance the abstraction level of spatial features. In the neck network, a multi-branch feature fusion structure, CreToNext, is introduced to optimize both the training efficiency and inference capability of the model. Furthermore, the Shape-IoU (SIoU) loss function is adopted to accelerate model convergence and improve recognition accuracy.

A customized dataset named IP9-AC, containing nine categories of agricultural pests, was constructed based on the IP102 dataset. Experimental results show that the improved CR_ULK_YOLO model achieves a mean Average Precision (mAP<sub>50–95</sub>) of 73.4%, a recall rate of 88.4%, and a precision of 94.6%, which are 4.2%, 3.6%, and 3.4% higher, respectively, than those of YOLOv8n. Comparative experiments with other mainstream detection models demonstrate that CR_ULK_YOLO exhibits superior interpretability and practical detection performance, providing valuable insights for accurate pest identification in complex agricultural environments.

**Keywords**: YOLOv8, CreToNext,Pest Detection

## 1. INTRODUCTION

Agriculture serves as the cornerstone of social development, and its stability and prosperity are closely linked to national economic growth and people's livelihoods. However, pest infestation has long been a major threat to crop yield and quality [1], and in recent years, the issue of pest damage has become increasingly severe. Traditional pest detection methods mainly rely on manual observation or on-site expert consultation, which suffer from low efficiency, high subjectivity, and poor timeliness, making them difficult to apply on a large scale [2–3].

With the continuous advancement of artificial intelligence and the introduction of classic convolutional neural network models such as VGG [4] and ResNet [5], deep learning–based agricultural pest recognition has become a rapidly emerging research field [6].

Rajan et al. [7] developed a pest detection approach based on the Support Vector Machine (SVM), which classifies images according to their color and texture features to achieve automatic pest detection. However, this method performs poorly under complex background conditions. Liu et al. [8] proposed a hybrid approach combining the Faster R-CNN [9] and R-FCN [10] frameworks to detect pests in grain storage environments. The core idea of this method is to incorporate the fully convolutional layer from R-FCN into the traditional Faster R-CNN structure to enhance detection robustness across different scales and background complexities.

As a typical two-stage object detection algorithm, Faster R-CNN first generates candidate regions through the Region Proposal Network (RPN) and then classifies and refines the bounding boxes using a convolutional neural network. However, this approach suffers from high computational complexity and slow detection speed when processing large-scale datasets, limiting its suitability for real-time applications.

To address these challenges, Teng et al. [11] proposed a novel network named MSR-RCNN (Multi-Scale Super-Resolution RCNN) for pest detection. This network integrates a Multi-Scale Super-Resolution (MSR) feature enhancement module that enables finer extraction of small pest details. By performing multi-scale feature extraction on the input image, the MSR module generates high-resolution feature maps, improving the model's capability to recognize small targets. Additionally, MSR-RCNN introduces a feature fusion mechanism to combine multi-scale information, thereby enhancing detection accuracy for diverse and visually similar pests. Experimental results on multiple public datasets demonstrate that MSR-RCNN significantly improves both precision and recall for small-object detection, showing strong practical potential in real-world scenarios such as pest monitoring in grain warehouses. Nevertheless, due to its complex architecture and heavy computational cost, MSR-RCNN requires higher hardware and computational resources for deployment.

These studies and improvements provide valuable insights and technical references for further optimizing detection speed and accuracy in pest detection tasks. On the self-constructed LLPD-26 dataset, MSR-RCNN achieved a detection accuracy of 67.4%.

The YOLO (You Only Look Once) series [12–16] represents a family of single-stage object detection algorithms known for their high detection speed and real-time performance [17]. Yang Kaihang et al. [18] improved the YOLOv5 model for rice pest detection by introducing a Coordinate Attention (CA) module into the backbone network to enhance spatial and directional feature sensitivity, and by using a-CIoU as the bounding-box loss function, effectively improving detection

International Journal of Science and Engineering Applications
Volume 14-Issue 11, 41 – 44, 2025, ISSN:- 2319 - 7560
DOI: 10.7753/IJSEA1411.1009

accuracy. However, this approach still struggles with densely distributed small-object pests.

Kang Feilong et al. further conducted comprehensive studies on crop pest detection technologies based on computer vision and proposed a deep learning–based pest detection method suitable for mobile applications, providing technical support for rapid and accurate crop pest identification [19–20].

Despite the progress achieved by the aforementioned methods, most of them have only been validated on small-scale and idealized datasets with relatively simple environmental conditions. In real agricultural scenarios, the complex field environment often causes occlusions between pests and crops, resulting in incomplete or partially visible pest targets in the captured images. Moreover, pests frequently appear in large groups and high densities, leading to a target density far exceeding that of typical object detection tasks.

To address these challenges, this study proposes an agricultural pest detection algorithm that integrates the CreToNext module with an improved SPPF_UniRepLKA model. By introducing a novel network architecture and optimized training strategy, the proposed method effectively mitigates several limitations of the original YOLOv8n algorithm, including insufficient feature extraction for agricultural pests, poor adaptability to complex backgrounds, and suboptimal detection performance for small objects.

| model | Precision | Recall | mAP50-95 | fs-1 |
|---|---|---|---|---|
| FasterR-CNN | 82.1% | 76.2% | 58.3% | 35.5 |
| YOLOv4-Tiny | 88.7% | 78.5% | 62.8% | 33.5 |
| YOLOv5s | 90.5% | 80.1% | 65.3% | 35.1 |
| YOLOv8n | 91.2% | 84.0% | 69.2% | 39.0 |
| Ours | 94.6% | 88.4% | 73.4% | 34.2 |

## 2. Improved YOLOv8 Algorithm

The CreToNext module was proposed by DAMO-YOLO [20] as an efficient network architecture specifically designed for complex feature extraction tasks. This architecture inherits the concept of residual connections, where each residual block consists of a CBS block and a re-parameterized convolution block (RepConv) [21–22].

Based on the maximum entropy principle, the module achieves high performance under low-latency constraints by incorporating spatial pyramid pooling and focus mechanisms within a ResNet-like [23–24] or CSP-like [25] structure.

In this study, the CreToNext module is introduced into the neck network for multi-scale feature fusion, replacing the conventional C2f module to enhance the feature extraction capability of the model. The structure of the CreToNext module is illustrated in Figure 1.

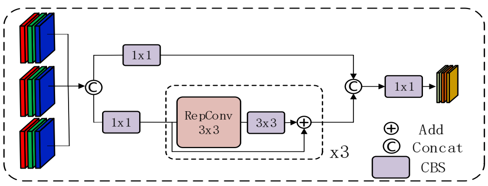


Fig.1 Structure of CreToNex

In the CBS block, the traditional ReLU activation function is replaced with Swish, a smooth and continuous nonlinear function. The smooth nature of Swish allows for more gradual gradient variations, which helps stabilize the training process.The core of the CreToNext module lies in its sparse network structure, where the input data are divided into two separate branches, each performing independent local feature extraction. The extracted local features are then reintegrated to obtain a richer and more comprehensive receptive field representation.The RepConv block adopts different parameterization strategies during the training and inference stages, as illustrated in Figure 2. The structural diagram of RepConv shows its distinct configurations for both phases. During training, RepConv employs two convolution operations — a 3×3 convolution kernel (Conv2d) and a 1×1 convolution kernel (Conv2d) — each followed by Batch Normalization (BN). The outputs of the 3×3 and 1×1 convolutions are summed to form the final output feature map. In the inference phase, however, the RepConv structure is simplified to a single 3×3 convolution operation. This simplification merges the multiple convolution and normalization operations used during training into a single equivalent operation, thereby reducing computational cost and improving inference speed and efficiency. The purpose of this structural design is to optimize computational resource utilization during inference without compromising model performance.

## 3. Dataset Construction and Experimental Platform Setup

The dataset used in this study is built upon the publicly available IP102 dataset [30–31], and is specifically designed for agricultural pest detection in Chinese farmlands. Considering that rice, wheat, and maize are the most commonly cultivated crops in China, nine representative pest species that commonly affect rice and maize were selected to construct the IP9-AC dataset.After screening and data augmentation, the IP9-AC dataset contains a total of 2,688 images. The dataset construction process consists of several key steps:Selection of common pest species,Sample augmentation and correction,Dataset partitioning, and Format conversion.To ensure diversity and representativeness, the dataset was divided into training, validation, and test sets at a ratio of 7:2:1. This partitioning strategy ensures comprehensive evaluation of the model's performance across different training stages.Table 1 presents the detailed categories of the selected pest species, while Figure 7 displays sample images of several agricultural pests. The construction of the IP9-AC dataset not only provides a valuable data resource for pest detection research but also lays a solid foundation for future studies in intelligent agricultural monitoring and pest identification.

Fig.2 Images of common agricultural pest

| train | val | test | labels |
|---|---|---|---|
| 285 | 93 | 57 | grubs |
| 210 | 49 | 19 | Needleworm |
| 199 | 53 | 37 | Wheat ed Spider |
| 196 | 48 | 19 | aphids |
| 221 | 69 | 26 | Rice leaf roller |
| 197 | 72 | 40 | Blowing the scales |
| 180 | 64 | 24 | Oides decempunctata |
| 178 | 52 | 24 | blister beetle |
| 200 | 53 | 23 | Cicadellidae |
| 1866 | 553 | 269 | |

**Tab. 1 Pest data set information**

## 4. Experimental Environment and Parameter Settings

**Tab 2 Experimental results of different models**

The experiments were conducted on a Windows 11 operating system equipped with an Intel Core i5-13490F CPU (base frequency 2.5 GHz, turbo frequency up to 4.8 GHz), 32 GB RAM, and an NVIDIA GeForce RTX 4070 GPU. The computational environment was configured with CUDA 11.3 and cuDNN 8.2.0, using Python 3.8 and the PyTorch 1.13.0 deep learning framework.

The initial learning rate was set to 0.01, momentum to 0.937, and weight decay to 0.0005. The input image size was fixed at 640 × 640 pixels, with a batch size of 32, 250 training epochs, and an IoU threshold of 0.7.

The model was trained and evaluated on nine common agricultural pests, including grubs, wireworms, wheat red spiders, aphids, rice leaf rollers, mealybugs, ladybugs, blister beetles, and leafhoppers.

Model performance was comprehensively assessed using Precision (P), Recall (R), mean Average Precision (mAP<sub>50–95</sub>), and detection speed (fps$^{-1}$).

As illustrated in Figure 8, the comparison of detection results among different algorithms demonstrates that the proposed method achieves superior detection performance and accuracy, particularly in challenging scenarios where the background closely resembles the target, the objects are partially occluded, or the pest targets are small and densely distributed. The proposed algorithm effectively alleviates issues of missed detections and low detection accuracy, thereby achieving more robust and reliable pest identification under complex field conditions.

Original Faster R-CNN YOLOv4-Tiny YOLOv5s YOLOv8n Ours

Fig.8 Comparison of detection results of different algorithms

## 5. Conclusion

This study focuses on the detection and recognition of common agricultural pests affecting major crops in China, including rice, maize, and wheat. To address the challenges of complex background information, small and densely distributed pest targets, and high intra-class similarity, an optimized version of the lightweight YOLOv8n model was developed. The proposed model achieved improved detection performance for target pest species through architectural and feature-level enhancements.Specifically, the CreToNext module was introduced to replace the C2f module in the neck network. This modification enhances the efficiency of information flow, strengthens the representation of multi-scale features, and improves the inference performance of the model while maintaining a lightweight structure.Experimental results demonstrate that, compared with the original YOLOv8n model, the proposed CR_ULK_YOLO achieves improvements of 3.4% in Precision, 3.6% in Recall, and 4.2% in mAP<sub>50–95</sub>, respectively.The results further indicate that CR_ULK_YOLO exhibits superior detection capability for small and densely distributed pest targets such as aphids, as well as improved accuracy across other pest categories. The proposed algorithm demonstrates high stability and robustness in complex farmland environments, providing an effective and practical visual detection approach for agricultural pest monitoring and control.

## 6. References

[1] Bowman, M., Debray, S. K., and Peterson, L. L. 1993. Reasoning about naming systems. .

[2] [1] Tang X., Xiang Q., Huang X., et al. Detection of Agricultural Pests Based on Improved YOLOv5 Algorithm[J]. Journal of South China Normal University (Natural Science Edition), 2023, 55(4): 42–49.

[3] [2] Peng H., Xu H., Gao Z., et al. Insect Pest Detection of Field Crops Based on Improved YOLOF Model[J]. Chinese Society for Agricultural Machinery, 2023, 54(4): 285–294.

[4] [3] Zhai Z., Cao Y., Xu H., et al. Review of Key Techniques for Crop Disease and Pest Detection[J]. Chinese Society for Agricultural Machinery, 2021, 52(7): 1–18.

[5] [4] Simonyan K., Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition[J]. arXiv preprint arXiv:1409.1556, 2014.

[6] [5] He K., Zhang X., Ren S., Sun J. Deep Residual Learning for Image Recognition[C]// Proceedings of the

IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 770–778.

[7] [6] Bian K. C., Yang H. J., Lu Y. H., et al. Overview of the Application of Deep Learning in the Detection and Identification of Pests and Diseases in Agriculture[J]. Software Guide, 2021, 20(3): 26–33.

[8] [7] Rajan P., Radhakrishnan B., Suresh L. P. Detection and Classification of Pests from Crop Images Using Support Vector Machine[C]// 2016 International Conference on Emerging Technological Trends (ICETT). IEEE, 2016: 1–6.

[9] [8] Liu Z., Zhang Z., Yang G., Sun H., et al. Application of Object Detection Algorithm in Identification of Rice Weevils and Maize Weevils[C]// Proceedings of the 2018 2nd International Conference on Deep Learning Technologies (ICDLT'18). 2018: 76–80.

[10] [9] Bangrong W., Jun W., Xiaofeng X., et al. Gas Mask Wearing Detection Based on Faster R-CNN[J]. Journal of Ambient Intelligence and Smart Environments, 2024, 16(1): 57–71.

[11] [10] Li C., Guo H., Ma D., et al. Comparative Analysis of the Accuracy of Deep Learning Algorithms for Oil Tank Detection in Remote Sensing Images[J]. Hydrographic Surveying and Mapping, 2020, 40(2): 52–56.

[12] [11] Teng Y., Zhang Y., Dong J., et al. MSR-RCNN: A Multi-Class Crop Pest Detection Network Based on a Multi-Scale Super-Resolution Feature Enhancement Module[J]. Frontiers in Plant Science, 2022, 13: 810546.

[13] [12] Redmon J., Divvala S., Girshick R., et al. You Only Look Once: Unified, Real-Time Object Detection[J]. arXiv preprint arXiv:1506.02640, 2016.

[14] [13] Redmon J., Farhadi A. YOLO9000: Better, Faster, Stronger[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA, 2017: 6517–6625.

[15] [14] Redmon J., Farhadi A. YOLOv3: An Incremental Improvement[J]. arXiv preprint arXiv:1804.02767v1, 2018.

[16] [15] Liu J., Wang X. Tomato Diseases and Pests Detection Based on Improved YOLOv3 Convolutional Neural Network[J]. Frontiers in Plant Science, 2020, 11: 898–910.

[17] [16] He Y., Chen D., Peng L. Research on Object Detection Algorithm of Economic Forestry Pests Based on Improved YOLOv5[J]. Journal of Chinese Agricultural Mechanization, 2022, 43(4): 106–115.

[18] [17] Duan X., Wang S., Zhao Q., et al. Detection Method of Major Corn Pests Based on Improved YOLOv4 Model[J]. Shandong Agricultural Sciences, 2023, 55(10): 167–173.

[19] [18] Yang K., Zhu Z., Luo X., et al. Research on Rice Pest Detection Algorithm Based on Improved YOLOv5 Model[J]. Electronic Technology, 2023, 52(8): 34–38.

[20] [19] Li P., Liu Y., Shi X., Zhang S. Agricultural Pest Detection Method Based on Multi-Scale Residual Spatial Attention Lightweight U-Net[J]. Jiangsu Agricultural Sciences, 2023, 51(3): 187–196.

[21] [20] Kang F., Li J., Liu T., et al. A Review of Image Recognition Application Technologies for Multiple Types of Crop Diseases and Pests[J]. Jiangsu Agricultural Sciences, 2020, 48(22): 22–27.

[22] [21] Xu X., Jiang Y., Chen W., et al. Damo-YOLO: A Report on Real-Time Object Detection Design[J]. arXiv preprint arXiv:2211.15444, 2022.

[23] [22] Wang H., Wang B., Ge C. Reparameterized YOLOv8 Pavement Disease Detection Algorithm[J]. Journal of Computer Engineering & Applications, 2024, 60(5): 191–199.

[24] [23] Aghaei A., Ebrahimi Moghaddam M., Alzheimer's Disease Neuroimaging Initiative. Brain Age Gap Estimation Using Attention-Based ResNet Method for Alzheimer's Disease Detection[J]. Brain Informatics, 2024, 11(1): 16–34.

[25] [24] Cao Y., Zhu X., Zhao Y., et al. Plant Leaf Disease and Pest Identification Based on Improved ResNet[J]. China Agricultural Mechanization and Chemistry Journal, 2021, 42(12): 175–175.

[26] [25] Huang H. Research on Motor Imagery Classification Combining CSP and Deep Learning[D]. Anhui University, 2022: 42–52.

[27] [26] Xiuhuan D., Shixin L., Jixiang Z. YOLOv5s Object Detection Based on Sim-SPPF Hybrid Pooling[J]. Optoelectronics Letters, 2024, 20(6): 367–371.

[28] [27] Ding X., Zhang Y., Ge Y., et al. UniRepLKNet: A Universal Perception Large-Kernel ConvNet for Audio, Video, Point Cloud, Time-Series, and Image Recognition[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2024: 5513–5524.

[29] [28] Chen Q., Si J. Application Layer Protocol Identification Method Integrating SENet and Transformer[J]. Journal of Computer Science and Exploration, 2024, 18(3): 805–817.

[30] [29] Gao Q., Pan Y., Zhu L., et al. Improved YOLOv5 Remote Sensing Object Detection Method Based on SIoU Function[J]. Yangtze River Information and Communication, 2022, 35(11): 5–8.

[31] [30] Guo J., Wang R., Nan J., et al. YOLOv5 Model Integrated with GRN Attention Mechanism for Insect Pest Recognition[J]. Transactions of the Chinese Society of Agricultural Engineering (CSAE), 2024, 40(8): 159–170.

[32] [31] Gan Y., Guo Q., Wang C., et al. Recognizing Crop Pests Using an Improved EfficientNet Model[J]. Transactions of the Chinese Society of Agricultural Engineering (CSAE), 2022, 38(1): 203–211.