Research on Improved YOLOv5 Rice Leaf Disease Detection Algorithm

Hongyi Xia[1], Lixin Tan[1,*]

[1]College of Information and Intelligence Science and Technology, Hunan Agricultural University, Hunan 410125, China

[*] Corresponding Author

Abstract

In view of the complex characteristics, multi-scale and low efficiency of rice diseases, deep learning was used to construct the DEFFN-YOLOv5 rice leaf disease algorithm and study rice bacterial blight, rice blast and brown spot. In order to improve the accuracy of disease detection, the original YOLOv5 algorithm was improved and the PixelShuffle upsampling module was introduced to restore image details. In addition, feature extraction capabilities are enhanced, and deformable convolution and lightweight ECA channel attention modules are introduced. By using BiFPN to improve the PAN module, information interaction is enhanced and the model's understanding and positioning capabilities are improved. Experiments have shown that the average accuracy (mAP) of the improved DEFFN-YOLOv5 algorithm in target detection reaches 86%, which is 3% higher than the original YOLOv5 algorithm. At the same time, the computational requirements are reduced by 4.6 GFLOPs, which is 27.85% less than the original YOLOv5 algorithm. These improvements make DEFFN-YOLOv5 perform better in rice disease detection.

Keywords

DEFFN-YOLOv5, rice leaf disease detection, PixelShuffle.

1. Introduction

Rice is susceptible to disease, which not only reduces its quality but also affects yields. Traditional disease detection methods rely on manual observation and experience of tea farmers, are time-consuming and inaccurate, and cannot meet actual needs[1]. In the field of plant disease detection, the current mainstream target detection algorithms based on deep learning are divided into two types: one is a two-stage algorithm based on regional suggestion frame prediction, such as Fast R-CNN[2] and Faster R-CNN[3] etc.; another one is to use a single-stage algorithm based on regression problems, such as SSD[4] and YOLO[5]. The single-stage detection algorithm has excellent speed and accuracy in target detection, but there are problems such as missed detection and poor detection of small targets. As the depth of the convolutional network increases, the performance of the model may degrade, which is not ideal for handling small object detection tasks. Based on YOLOv5, this study proposes an improved DEFFN-YOLOv5's rice leaf disease detection method designs the DEFFN module to replace the convolution module, PAN structure, and C3 module in the original network, and combines it with the PixelShuffle[6] upsampling module to optimize the model. The improved model has better receptive fields, information capture capabilities, and better performance, which has very positive significance for rice production activities.

2. Model structure design

2.1. YOLOv5 model

YOLO is a target detection algorithm based on regression ideas. Compared with traditional two-stage models such as R-CNN and Faster R-CNN, YOLO, a single-stage detection method, performs better in terms of detection speed and detection accuracy. YOLOv5 mainly consists of three parts: the backbone network Backbone, the detection layer Neck and the prediction layer Head.

2.2. DEFFN structure for deep fine-grained information fusion

YOLOv5 uses the nearest neighbor interpolation method for upsampling, which considers the gray value of the closest pixel and ignores the influence of other pixels, resulting in jagged artifacts in the image, reducing image quality, and affecting the accuracy of object detection. To improve this problem, the original upsampling module was replaced with the PixelShuffle upsampling module to more effectively handle models that are sensitive to details and edge information. The backbone network of YOLOv5 adopts CBS structure to show good detection performance on the COCO dataset. However, for small targets, feature extraction is limited, resulting in reduced detection accuracy. The original YOLOv5 uses the FPN structure to fuse features, which has one-way information transfer limitations. To solve this problem, the following measures are taken: the PAN structure is introduced and channels are added from bottom to top to fuse the bottom and top-level information, but PAN increases the model complexity. In addition, the Neck part of the YOLOv5 model uses the C3 module, which may lose detail information during the convolution and pooling process, especially when processing small objects or high-resolution images. To improve these problems, the DEFFN structure is proposed, as shown in Figure 1.

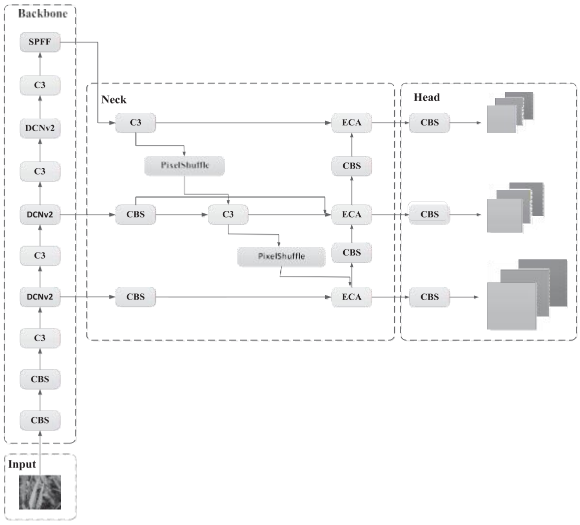

Figure 1: DEFFN structure of deep fine-grained information fusion

DEFFN replaces the three CBS structures in the original backbone network and introduces DCNv2 deformable convolution[7] so that the convolution kernel can be dynamically adjusted to adapt to object feature extraction of different shapes. This is very helpful for extracting information about diseased areas, especially when dealing with complex-shaped objects. In addition, the neck part of the YOLOv5 model is optimized to improve the PAN-based BiFPN (Bi-Directional Feature Pyramid Network, bidirectional feature pyramid network)[8] structure. The BiFPN structure improves feature fusion efficiency and reduces computing resource requirements through node optimization and channel increase. The weight normalization mechanism is introduced to enhance the model's perception of different target situations. It can also effectively integrate different levels of feature map information in the prediction stage to solve interference factors such as noisy images, which helps to improve the performance and robustness of the YOLOv5 model. , especially when dealing with complex scenes and multi-scale target detection. Finally, the Neck part introduces the ECA (EfficientChannel Attention) attention mechanism[9] to replace the C3 module in the original network and focus more computing resources on the image ontology information.

In the DEFFN structure, replace the C3 module with the ECA module. The C3 module has far more parameters than the required batch normalization parameters, and the ECA module has approximately 23.68% of the parameters of the C3 module. This means that the ECA module significantly reduces the number of parameters of the model and improves the efficiency of computing and storage resources while maintaining the same performance. The DEFFN structure incorporates PixelShuffle, which enhances feature extraction and multi-scale information fusion capabilities, while improving model efficiency and object detection robustness, especially under limited computing resources. This structure not only improves performance, but also better adapts to environments with limited computing resources, making the model more scalable and practical. Therefore, the DEFFN structure fused with PixelShuffle is of great significance for object detection tasks.

3. Experiment and result analysis

3.1. Experimental data

The research objects are bacterial blight, rice blast, and brown spot. The rice leaf disease data used comes from flying paddle, with a total of 1448 images.

The machine selected 1,158 images as the training set, 144 images as the verification set, and 146 images as the test set.

3.2. Experimental environment and parameter configuration

Training and testing based on YOLOv5 improved algorithm on aconfiguration with PCIntel(R)Core(TM) i5-13600K CPU@3.50GHz,NVIDIA GeForce RTX3060 12GB and Set up the SGD stochastic gradient descent optimizer on a Windows 11 system with 32GB of running memory. The initial learning rate is 0.01, which is used to update the weight values. The input image size is 640×640 pixels, the batch size is 16, and the rounds are 100. The code running environment is python, the version is 3.8.16, and the Pytorch version is 1.13,1. The evaluation index used is mean average precision (Mean Average Precision, mAP).

3.3. Result analysis

In order to objectively evaluate the effectiveness of the improved algorithm, the same samples and parameter configurations were used to conduct comparative experiments with existing target recognition algorithms with higher comprehensive performance. The experimental results are shown in Table 1. As can be seen from the table, the performance of the single-stage network model is significantly better than that of the two-stage network model. The network designed for rice disease target detection outperforms mainstream target detection models on most evaluation indicators. In the control group, SSD, Faster R-CNN and YOLO other series of detection. The accuracy is lower than YOLOv5s. Compared with YOLOv5s, mAP(@0.5)and mAP(@0.95) were increased by 3% and 2% respectively, and the number of parameters was reduced by 0.77M. While improving accuracy, the number of parameters and calculation amount were significantly reduced. Compared with YOLOv3-tiny, YOLOv5m and YOLOv7 of the same series, mAP(@0.5) has increased by 10.6%, 7% and 17.4% respectively, and its parameter amount and calculation amount have also been reduced a lot, and the detection speed has been improved. Compared with the dual-stage detector FasterR-CNN, it has obvious advantages in accuracy and detection speed.

Table 1: Overall detection performance evaluation

Model	AP(@0.5)			mAP (@0.5)	mAP (@0.95)	Parameter amount (M)	Calculation amount (GFLPs)
	Brown spot	Rice blast	Bacterial blight				
YOLOv3-tiny	0.709	0.914	0.638	0.754	0.38	8.67	12.9
YOLOv5s	0.871	0.916	0.704	0.83	0.401	7.02	15.8

YOLOv5m	0.887	0.931	0.553	0.79	0.395	20.86	47.9
YOLOv7	0.637	0.846	0.576	0.686	0.289	36.49	103.2
FasterR-CNN	0.821	0.883	0.476	0.727	0.363	/	/
SSD	0.997	0.905	0.618	0.840	/	/	/
Ours(DEFFN-YOLOv5)	0.844	0.931	0.806	0.86	0.421	6.25	11.4

4. Conclusion

In order to improve the accuracy of rice disease detection, the YOLOv5 algorithm was improved, and the PixelShuffle upper sampling module and DEFFN structure were introduced to improve disease identification and spot detection capabilities. The improved DEFFN-YOLOv5 algorithm has improved the average accuracy (mAP) of target detection by 3% compared to the original version; the computing requirements have been reduced by 4.6GFLOPs, a reduction of 27.85%. Experiments have proven that DEFFN-YOLOv5 has leading performance in rice disease detection tasks, providing an important reference for future automated rice disease detection.

References

[1] Koutroumbas, S. D., Katsantonis, D., Ntanos, D. A., et al. (2009). Blast disease influence on agronomic and quality traits of rice varieties under Mediterranean conditions. Turkish Journal of Agriculture & Forestry, 33(5), 487-494.

[2] Derya, A., Eser, S., Esin, D., et al. (2023). A new super-resolution Faster R-CNN model based detection and classification of urine sediments. Biocybernetics and Biomedical Engineering, 43(1), 58-68.

[3] Ren, S., He, K., Girshick, R., et al. (2017). Faster R-CNN: Towards real-time object detection with region proposal networks. IEEE Transactions on Pattern Analysis & Machine Intelligence, 39(6), 1137-1149.

[4] Chen, L., Zhang, Z., & Peng, L. (2019). Real-time detection method based on improved SSD. Progress in Lasers and Optoelectronics, 56(1), 109-115.

[5] Redmon, J., Divvala, K. S., Girshick, B. R., et al. (2016). You only look once: Unified, real-time object detection. CoRR, 779-788.

[6] Shi, W., Caballero, J., Huszár, F., et al. (2016). Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 2651-2659). IEEE.

[7] Dai, J., Qi, H., Xiong, Y., et al. (2017). Deformable convolutional networks. IEEE, 764-773.

[8] Tan, M., Pang, R., Le, Q. V. (2020). EfficientDet: Scalable and efficient object detection. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (pp. 10778-10787).

[9] Wang, Q., Wu, B., Zhu, P., et al. (2020). ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks. In 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 8244-8253). IEEE.