Article

Research on Lightweight Kiwifruit Detection Algorithm Based on YOLOv8-BFW

Wei Shan , Dan Zhao , Yurong Yue [*]

Article

Research on Lightweight Kiwifruit Detection Algorithm Based on YOLOv8-BFW

Wei Shan [1,2], **Dan Zhao** [1,2] and **Yurong Yue** [1,2,*]

1   School of Physics and Electronic Information, Huaibei Normal University, Huaibei 235000, China

2   Anhui Key Laboratory of Intelligent Computing and Applications, School of Physics and Electronic Information, Huaibei Normal University, Huaibei 235000,China

*   Correspondence: 12212060813@chnu.edu.cn

Abstract: Aiming at the problems of low recognition efficiency and accuracy and high computing resources of YOLOv8 model in kiwifruit detection, this paper proposes a lightweight kiwifruit detection algorithm YOLOv8-BFW based on YOLOv8. Firstly, the weighted bidirectional feature pyramid network BiFPN module is introduced into the Neck part of YOLOv8 to realize multi-scale feature fusion, so as to improve the efficiency and accuracy of its detection; Secondly, all C2f modules are integrated into the FasterNet network to form a C2f-Faster module to reduce computing resources; Finally, the WIOU boundary loss function is used to replace the original CIOU to improve the regression performance of the network bounding box and accelerate the convergence speed of the model, which can more accurately measure the overlapping area of the box and further improve the efficiency and accuracy of its detection. The experimental results show that the average detection accuracy mAP of the YOLOv8-BFW algorithm on the self-made kiwifruit dataset collected and labeled in this paper reaches 96.9%. Compared with the original YOLOv8 algorithm, the model size is reduced by 32%, the calculation amount is reduced by 22.2%, the parameter quantity is reduced by 29.6%, and the detection speed is increased by 26.7%.

Keywords: kiwifruit detection; lightweight network; YOLOv8n; YOLOv8-BFW; BiFPN; C2f-Faster; WIOU

1. Introduction

Relevant data show that in recent years, the yield of kiwifruit in China and the world has been increasing year by year. China has become one of the countries that mainly grow kiwifruit, and its planting area has exceeded the total planting area of other countries. The detection of kiwifruit has gradually become one of the important issues in the field of agriculture. The degree of automation in the current kiwifruit industry structure in China is low. The basic steps in the kiwifruit industry such as kiwifruit picking are still mainly manual, which is not only time-consuming and laborious, but also has low detection accuracy, which seriously affects the development of the national economy. In recent years, with the continuous development of computer computing power, deep learning technology is widely used in target detection, and the use of machine vision detection in artificial intelligence [1] has become the mainstream development trend. At present, the widely used method based on machine vision target detection has strong ability to express target features and good recognition effect is YOLO (YOLOv1, YOLOv2,...... YOLOv7, YOLOv8, etc.) single-stage algorithm series[2]. Since the publication of YOLOv1 in 2015, YOLO has been iteratively updated, which has witnessed visual object detection in the era of deep learning. Based on the YOLO single-stage target detection algorithm model, the algorithm does not need to generate candidate frames, and directly transforms the visual target positioning problem into a regression problem. Therefore, the fast detection ability of YOLO model can be better applied to visual object detection. However, at present, the target detection algorithm for YOLO series improves the accuracy of the model by improving the

network structure, which will lead to higher complexity and greatly increase the amount of calculation and parameters of the model. Therefore, the research on the lightweight of YOLO series target detection has gradually attracted attention.

The algorithms of the YOLO series, such as YOLOv3[3], YOLOv5, YOLOv6[4], YOLOv8[5], etc., are a general target detection algorithm model. Among them, for the problem of low recognition efficiency and accuracy and high computing resources in YOLOv8, a large number of references are consulted. It can be seen that the efficient and lightweight grape and picking point synchronization detection model based on key point detection proposed in literature [6]. The model uses the BiFPN module in the Neck part of YOLOv8 to achieve multi-scale feature fusion. The experimental results show that the detection accuracy of the model is improved after adding the module. In Reference [7], the walnut recognition method of UAV remote sensing image was proposed, in which the YOLOv8 model was improved, and a lightweight feature extraction backbone FasterNet fusion C2f module was proposed to form a C2f-Faster module. After adding this module, according to the results of ablation experiments, it can be seen that under the average accuracy of model detection, the parameter amount of the model is reduced by 1.38 M, and the size is reduced by 2.7 MB. In Reference [8], the improved lightweight YOLOv8 model was used to detect multi-stage apple fruits in complex orchard environment in real time. The WIOU boundary loss function was used to replace the original CIOU. The experimental results showed that only the loss function was replaced, and the parameters and calculation amount of the model did not change, but the average accuracy of the model was improved.

The above research is based on the different processing of the YOLOv8 target detection model. Some of the improved modules reduce the calculation amount and parameter number of the model, and some improve the detection accuracy of the model. However, in many small and medium-sized kiwifruit processing enterprises, the computing power and resources of the equipment are limited, which requires reducing the number of parameters and calculation of the model while ensuring its detection accuracy.

Based on the above problems, the BiFPN module in Reference [6] is applied to the Neck part of YOLOv8 to improve the average detection accuracy of kiwifruit. The C2f-Faster module in Reference [7]is applied to YOLOv8 to reduce computing resources. The WIOU loss function in [8] is applied to this paper to further improve the efficiency and accuracy of its detection. In summary, a kiwifruit target detection model YOLOv8-BFW based on YOLOv8 is proposed, which aims to solve the problems of low efficiency and accuracy of kiwifruit detection and recognition and high computing resources.

2. The Introduction of YOLOv8 Algorithm

YOLOv8 model is a target detection algorithm based on deep learning developed by Ultralytics, which has good performance in detection speed, accuracy and feature extraction. YOLOv8 is an optimization and improvement based on the YOLOv5 model. It replaces the C3 structure of the backbone and neck network with a gradient-rich C2f module, deletes the convolution operation in the upsampling stage of YOLOv5, and replaces the Head part with the mainstream decoupling head structure to separate classification and detection. The neck network continues the PANet structure of YOLOv5 for multi-scale feature fusion. Compared with the YOLOv5 model, it has achieved significant improvement in performance, flexibility and efficiency. According to the width and depth of the network model, it is divided into five different structures : n, s, m, l and x. According to the requirements of kiwifruit detection in this paper, considering the real-time performance of target detection, this paper improves the YOLOv8n network with the smallest and fastest speed as the benchmark model. Its benchmark network structure is shown in Figure 1 below.
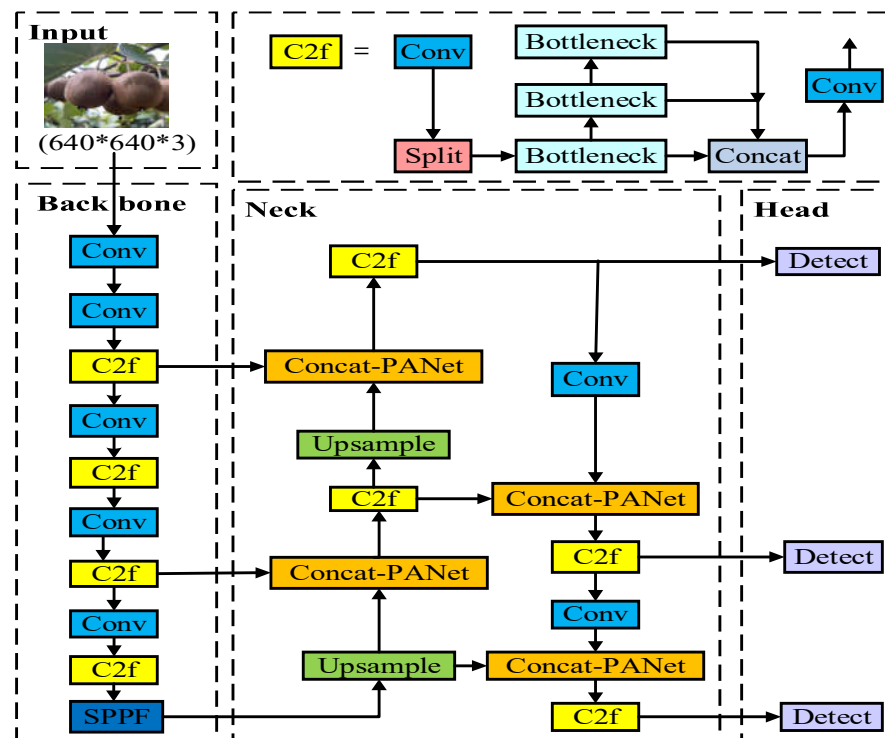
Figure 1. YOLOv8 network structure.

3. Improved Algorithm YOLOv8-BFW Based on YOLOv8

3.1. Concat-BiFPN Module

In order to ensure a certain detection accuracy during model training and realize real-time detection of kiwifruit,the Neck part of YOLOv8 was modified in this study. The weighted feature pyramid network BiFPN structure [9] is used to replace the original path aggregation network PANet structure. In the PANet network structure, the feature pyramid FPN block transfers the high-level features to the low-level from top to bottom, and supplements the semantic features of the low-level. The path aggregation structure PAN block transfers the low-level information to the high-level in a bottom-up manner, so that the low-level information is also fully utilized, and finally realizes the effective fusion of multi-scale features. However, bidirectional feature fusion enhances the representation ability of the backbone network. However, this complex network structure will inevitably lead to an increase in the number of parameters and calculations.

Therefore, this paper introduces a more lightweight and efficient weighted bidirectional feature pyramid network BiFPN (Bidirectional Feature Pyramid Network), which is a new network structure for computer vision tasks such as target detection and semantic segmentation. The network structure diagram of FPN, PANet and BiFPN is shown in Figure 2 below. The BiFPN network is proposed on the basis of FPN and PANet networks. FPN is a top-down unidirectional feature fusion structure. PANet is a fixed symmetric bidirectional feature fusion structure. BiFPN introduces bidirectional feature fusion and dynamic network structure design, which makes feature fusion more comprehensive and adaptable. The implementation is to add an additional weight to each input during feature fusion, so that the network can learn the importance of each input feature, so as to make better use of semantic information and fine-grained information and improve the performance of target detection.
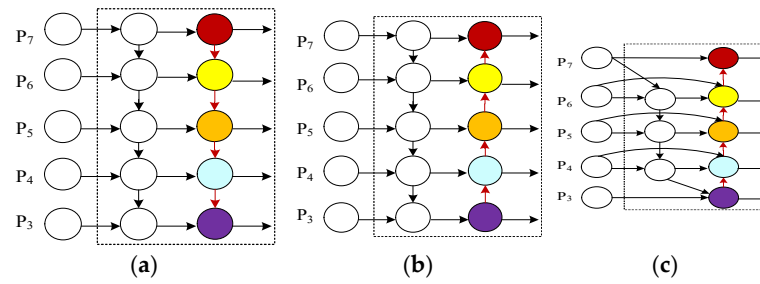
Figure 2. (a) FPN **(b)** PANet **(c)** BiFPN.

In YOLOv8, the PANet structure is replaced by the weighted bidirectional feature pyramid network BiFPN, and the original Concat-PANet module is replaced by the Concat-BiFPN module to improve the average detection accuracy of the model. The improved Neck structure is shown in Figure 3 below.
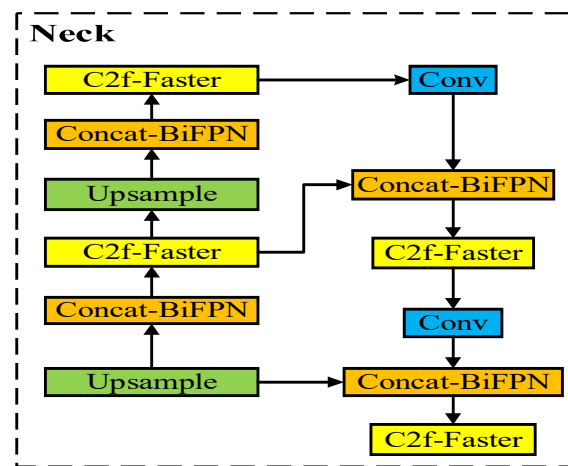

Figure 3. The improved Neck network structure diagram.

3.2. Concat-BiFPN Module

In order to ensure the low consumption of computing resources during model training, the FasterNet lightweight network [10] is cited. It is an efficient lightweight target detection neural network model. The structure diagram is shown in Figure 4 below, which mainly includes four levels. Each level contains a FasterBlock module for down sampling or increasing the number of channels.
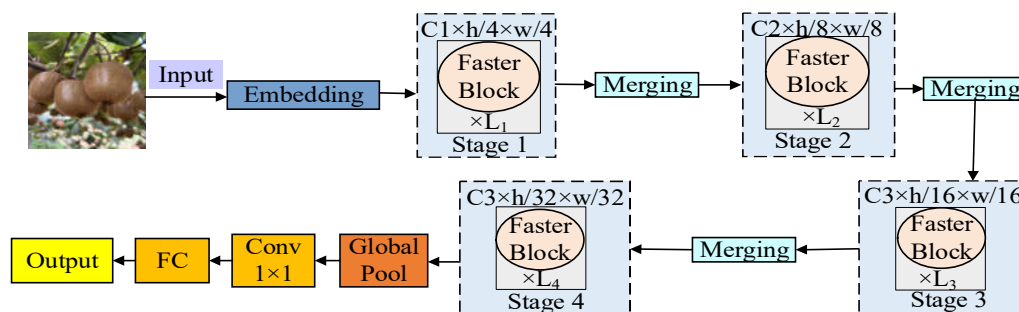

Figure 4. The feature fusion network structure diagram.

The FasterBlock module of each level is composed of a local convolution PConv layer followed by two point-by-point convolution PWConv (or Conv 1×1) layers, and the structure diagram is shown in Figure 5 below. The core of FasterNet is PConv, whose structure is shown in Figure 6 below. Unlike conventional convolutions, partial convolutions do not mechanically apply the same

convolution kernel to all parts of the input data. On the contrary, it dynamically determines the scope of the convolution kernel based on the validity of the data, that is, whether the data points are missing or damaged. This operation not only improves the robustness of missing data, but also can more effectively extract and utilize the remaining information, so as to have better device computing power.
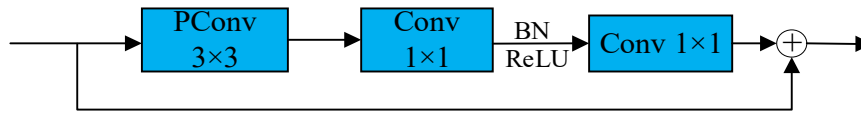

Figure 5. FasterNet Block module structure diagram.

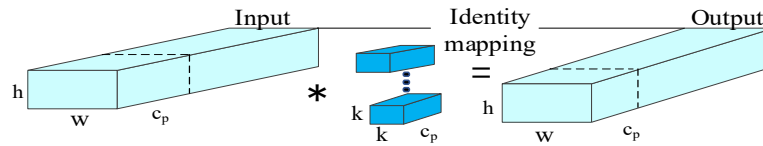**Figure 6.** PConv network structure diagram.

For the above Figure 6 it is assumed that the length and width of the input feature map are h and w, respectively, the number of input channels is cn and c, the channel participating in the convolution is cp, the size of the convolution kernel is k, and the convolution rate is r. The expression is as shown in formula (1), the FLOPS expression of the calculation amount is as shown in formula (2), and the MAC calculation formula of the memory access is as shown in formula (3).

$$r = \frac{c_p}{c} = \frac{1}{4} \tag{1}$$

$$FLOP\ S = h \times w \times k^2 \times c_p^2 \tag{2}$$

$$M_{AC} = h \times w \times k^2 \times 2c_p \approx h \times w \times 2c_p \tag{3}$$

As shown in the above formula, usually 1/4 is selected as the convolution part during operation. Therefore, compared with conventional convolution, PConv reduces the calculation amount of 1/16 and the memory access amount of 1/4, and reduces the operation of floating point number, thus realizing the lightweight of the model to a certain extent.

In YOLOv8, the Bottleneck in the C2f module is composed of two continuous Conv modules. The Conv module includes three steps:convolution Conv2d, activation BN, and pooling SiLU. The parameters are large and the computational complexity is high. The structure is shown in Figure 7.
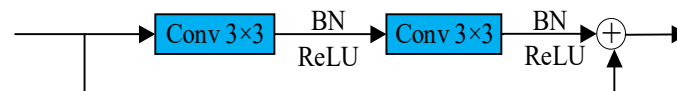

Figure 7. Bottleneck structure diagram.

The Bottleneck module in the C2f module of YOLOv8 is replaced by the FasterBlock module, which uses residual connections for input and output features. The number of channels in the middle layer is expanded, and the middle layer uses a normalized layer. By reducing the number of convolution operations, the number of networks is greatly reduced and the inference speed is accelerated. The C2f-Faster module contains different convolutions, connections, and built-in function Split modules. The model can fully and effectively utilize the information of all channels while reducing the number of parameters, and maintain the diversity of features extracted by the model. The C2f-Faster module structure diagram is shown in Figure 8 below.
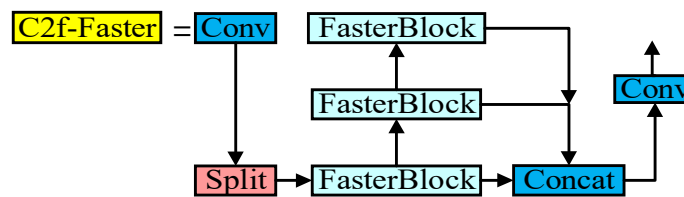
Figure 8. C2f-Faster module structure diagram.

3.3. WIOU Loss Functionon

In deep learning, the loss function is one of the core components, which measures the difference between the predicted results of the model and the real values. By minimizing the value of the loss function, the model can gradually improve its performance during training. The loss function provides a clear optimization goal for the neural network and is an important bridge connecting data and model performance.

The expression of   IoU is shown in formula (4).

$$IoU = \frac{area\ of\ overlap}{area\ of\ union} \tag{4}$$

The intersection area and union area are the same as the meaning of the set in mathematics, and the graphic representation is shown in Figure 9.
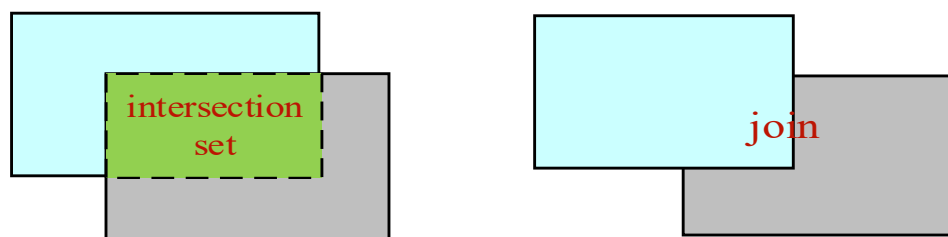
Figure 9. Intersection and union represent graphs.

In YOLOv8, Binary Cross Entropy (BCE) is used as the classification loss, and DFL Loss + CIOU Loss is used as the regression loss. BCE Loss is used to deal with category imbalance, and DFL Loss is used to deal with the problem of category imbalance and background category. In the detection box regression, CIOU is used as the loss function to deal with the overlap between the prediction box and the real box. However, since the CIOU 's aspect ratio only describes the relative relationship between the prediction frame and the real frame width or height, there is a certain ambiguity, and the balance problem of difficult and easy samples is not considered, which will lead to slower convergence speed and lower model detection efficiency.

In this paper, a dynamic non-monotonic focusing mechanism (wise intersection of union, WIoU) is used to replace the original loss function CIOU, and the ' outlier ' is used instead of IoU to evaluate the performance of the anchor frame, that is, a weight factor is introduced to adjust the degree of overlap between different size target frames. The weight factor is calculated according to the size of the target box, and usually the smaller target box will be given a larger weight. The WIoU loss function is to sum the WIoU values of all the prediction boxes as the final loss value, which can effectively reduce the competition of high-quality anchor boxes, reduce the unfavorable gradient caused by low-quality samples, and focus on the anchor boxes of ordinary quality, so as to improve the overall performance of the detection model.

3.4. YOLOv8-BFW Network Structure

In order to improve the average detection accuracy of the model, reduce the number of parameters and calculations of the network, and realize the lightweight of the network, this paper has made the following three improvements for the YOLOv8 model. Firstly, the original path aggregation network PANet structure used in the neck connection part, that is, the Concat-PANet module is replaced by the Concat-BiFPN module. Secondly, all the C2f modules in YOLOv8 are integrated into the FasterNet lightweight network to form a new C2f-Faster module. Finally, the WIOU boundary loss function is used to replace the original CIOU to improve the regression performance of the network bounding box and accelerate the convergence speed of the model, which can more accurately measure the overlapping area of the box. The improved network structure is shown in Figure 10 below.
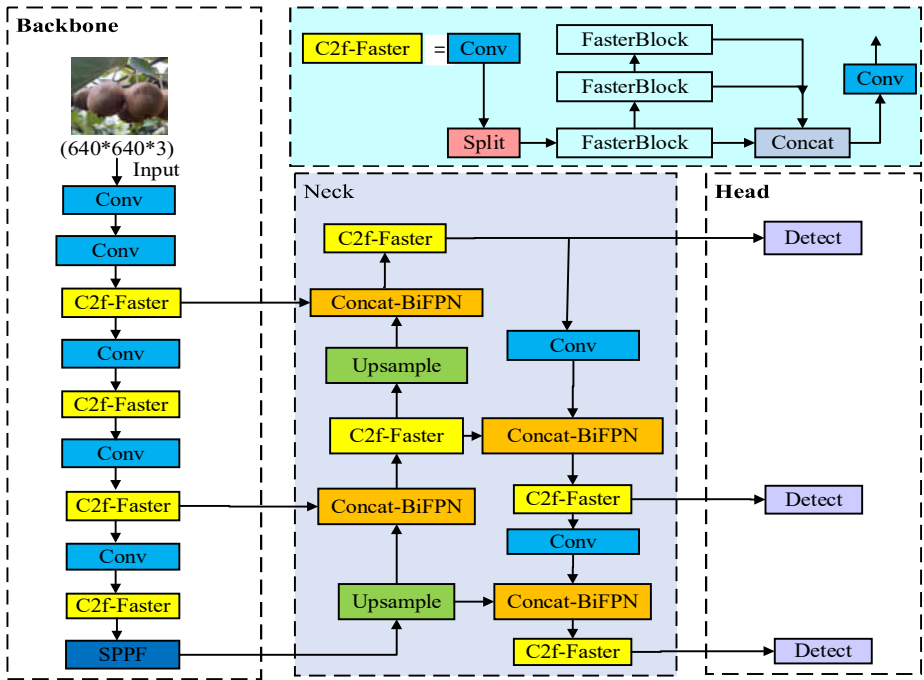

Figure 10. The network structure diagram of improved YOLOv8.

4. Experiment and Analysis

4.1. Experimental Platform

All the experiments in this paper are based on the same experimental environment, which can reduce the error caused by the change of experimental conditions and improve the comparability and repeatability of the results. In terms of hardware and software configuration, this paper uses Pytorch as a deep learning framework, which is flexible and allows users to write custom operations and models using the Python language. At the same time, it also supports GPU acceleration, which can use GPU 's parallel computing power to accelerate the training and reasoning of deep learning models. The specific configuration parameters are shown in Table 1 below.

Table 1. Experimental Hardware and Software Configuration.

parameter	configuration
learning_rate	0.01
epochs	300
batch-size	64
workers	4

Table 2 shows some relevant settings of the model during training. Among them, workers are used to set whether to use multithreading to read data. The learning rate is the step size of each iteration, and epochs are the total number of rounds of training.

Table 2. Experimental training parameter settings.

parameter	configuration
system environment	Windows 11
GPU	RTX 4090 Laptop GPU
deep learning framework	Pytorch 2.1.1
language	Python 3.10
memory	64G
CPU	lntel(R) Core(TM) i9-14900KF
accelerated environment	CUDA 12.1

4.2. Data Set

The experiment uses the self-made YOLO data set as the sample training and testing and validation set. The pictures used are taken by two different types of mobile phones and tro drones in the kiwifruit production base of zhouzhi county, xi'an city, shaanxi province. A total of more than 4000 pictures were taken. After deleting the blurred, repeated and non-research target pictures, the remaining 2856 pictures were sorted and labeled to make the kiwifruit data set. The following Table 3 is the relevant information of the pictures used in the experiment. For the problem that the size of the pictures is not uniform, the YOLOv8 network will automatically normalize its size to 640×640×3 at the input end. The size of the original picture and the normalized picture are shown in Figure 11 and Figure 12. In this experiment, the pictures in the kiwifruit data set were divided according to the division ratio of training set : validation set : test set = 8 : 1 : 1.

Table 3. Image-related information used in the experiment.

Shooting tool model	OPPO A32	iPhone 13 Pro Max	Trotterlo drone
Shooting size	3120×4160×3	3024×4032×3	2595×1936×3
normalized sizes		640×640×3	

Figure 11. Three data set samples

Figure 12. Normalized samples of three datasets

4.3. Experimental Evaluation Index

The evaluation indexes used to measure the average accuracy of the model detection are mainly the accuracy rate P, the recall rate R, the average accuracy of the detection mean mAP, the most important of which is mAP. The accuracy rate P represents the proportion of the targets predicted by the model that are really the targets, and the recall rate R represents the ability of the model to find

all the targets. AP represents the average precision value under different recall levels. The higher the AP is, the less the detection error is. The mAP is obtained by averaging the AP values of all categories, which provides an overall measure of model detection performance. The experiment summarizes four results, true positive (TP) refers to the manually marked kiwifruit is correctly detected, false positive (FP) refers to the object that is wrongly detected as kiwifruit, true negative (TN) refers to the negative sample that the system predicts to be negative, and false negative (FN) reflects the missing kiwifruit [11]. The expressions of AP and mean average precision (mAP) parameters are as follows.

$$AP_i = \int_0^1 p(r)dr \tag{5}$$

$$mAP = \frac{1}{N}\sum_{i=1}^{N} AP_i \tag{6}$$

Since only one category is marked in kiwifruit recognition, mAP is equal to AP. The higher the mAP is, the more accurate the prediction result of the model is. MAp@0.5 refers to the average accuracy of all images in each category if IoU is assigned to 0.5, and then averages all categories.

The number of model parameters refers to the number of parameters that need to be learned in the model. These parameters usually include weights and biases, which directly affect the performance indicators such as fitting ability, generalization ability and convergence speed of the model. The smaller the number of parameters is, the lower the complexity of the model is, and the computing resources and time required in the training and prediction process are relatively small.

The computing power of the model refers to the number of floating-point operations per second. The unit used to evaluate the computing speed is mainly used as an indicator to describe the hardware performance. It is used to describe the unit of the actual running speed of the deep learning model on the GPU, that is, the task of training and reasoning under how much computing power the model provides on the GPU. It is usually used to evaluate the computational complexity of the model and the demand for hardware computing power. The smaller GFLOPs means that the model performs less computation per unit time, and the overall computational efficiency is higher.

FPS represents the number of image frames processed by the model in one second, which is an important index to measure the inference speed of the model. The higher the FPS is, the faster the inference speed of the model is. The calculation formula is as follows.

$$FPS = \frac{1}{Total\ Time} \tag{7}$$

Among them, Total Time represents the total time spent.

5. Comparative Analysis of Experimental Results

5.1. The Ablation Experiment of the Improved Process and Its Result Analysis

In this paper, based on the yolov8n network as the benchmark model, the ablation experiments are carried out on the improved C2f-Faster and Concat-BiFPN modules and the single and combined improvement tasks of WIOU loss function, which proves the effectiveness of the improved method proposed in this paper on the kiwifruit detection task. The specific experimental results are shown in Table 4.

Table 4. Ablation experiment

algorithm	mAP@0.5/%	Size(MB)	Params/10^6	GFLPs	FPS/(frame/s)
YOLOv8n	96.6	6.23	3.05	8.3	232.2
YOLOv8n+Concat-BiFPN	96.7	5.55	2.78	8.1	256.4

algorithm	mAP@0.5%	Size(MB)	Params/10^6	GFLOPs	FPS/(frame/s)
YOLOv8n+C2f-Faster(Backbone)	96.6	5.26	2.64	7.0	232.6
YOLOv8n+C2f-Faster(Neck)	96.6	5.40	2.72	7.6	232.6
YOLOv8n+C2f-Faster(Backbone+Neck)	96.6	4.61	2.30	6.3	263
YOLOv8n+C2f-Faster(Backbone+Neck)+ Concat-BiFPN	96.8	4.26	2.11	6.3	263.2
YOLOv8n+C2f-Faster(Backbone+Neck)+ Concat-BiFPN +WIOU	96.9	4.26	2.11	6.3	294.1

It can be seen from Table 4 that after the first step, the feature fusion network BiFPN is used to replace the neck connection module, the mAP@0.5 and detection speed of the model are improved, and the parameter quantity and calculation amount of the model and the size of the model are slightly reduced, indicating that the module is effective for improving the average detection accuracy of the model. In the second step, after replacing the C2f module in the neck of the YOLOv8n model with C2f-Faster, the mAP@0.5 value and detection speed of the model did not change compared with YOLOv8n, but the parameter quantity and GFLOPs of the model decreased slightly. After replacing the C2f module in the trunk of the YOLOv8n model with C2f-Faster, the parameter quantity and GFLOPs of the model decreased slightly. After replacing the overall C2f module with C2f-Faster, the parameter quantity and GFLOPs of the model decreased the most. And the detection speed of the model is greatly improved. Therefore, this paper uses the C2f-Faster module to replace the overall C2f module to reduce the parameter quantity and calculation amount of the model and realize the lightweight of the model. After replacing the WIOU loss function in the third step, the number of parameters, the amount of calculation and the size of the model did not change, but the mAP@0.5 value of the model increased by 0.3%, indicating the effectiveness of WIOU in improving the average detection accuracy of the model.

Table 4. Ablation experiment

algorithm	mAP@0.5 %	Size(MB)	Params/10^6	GFLOPs	FPS/(frame/s)
YOLOv8n	96.6	6.23	3.05	8.3	232.2
YOLOv8n+Concat-BiFPN	96.7	5.55	2.78	8.1	256.4
YOLOv8n+C2F-Faster(Backbone)	96.6	5.26	2.64	7.0	232.6
YOLOv8n+C2F-Faster(Neck)	96.6	5.40	2.72	7.6	232.6
YOLOv8n+C2F-Faster(Backbone+Neck)	96.6	4.61	2.30	6.3	263
YOLOv8n+C2F-Faster(Backbone+Neck+ Concat-BiFPN	96.8	4.26	2.11	6.3	263.2
YOLOv8n+C2F-Faster(Backbone+Neck)+ Concat-BiFPN +WIOU	96.9	4.26	2.11	6.3	294.1

The above Table 5 is a comparative analysis of YOLOv8n and YOLOv8n-BFW. The experimental results show that the average detection accuracy of the YOLOv8n-BFW model proposed in this paper is 0.3% higher than that of the YOLOv8n model, the model size is reduced by 32%, the calculation amount is reduced by 22.2%, the parameter quantity is reduced by 29.6%, and the detection speed is increased by 26.7%. While ensuring the accuracy of the model detection, the lightweight detection of

kiwifruit was realized. The average detection accuracy of YOLOv8n and YOLOv8n-BFW on the kiwifruit dataset is shown in Figure 13.

**Table 5.** Comparative analysis of YOLOv8n and YOLOv8n-BFW

evaluating indicator	map@0.5/%	Size(MB)	Params/10^6	GFLOPs	FPS/(frame/s)
YOLOv8n	96.6	6.23	3.05	8.3	232.2
YOLOv8n-BFW	96.9	4.26	2.11	6.3	294.1
performance change	0.3 ↑	32% ↓	22.2% ↓	29.6% ↓	26.7% ↑

(a)   (b)
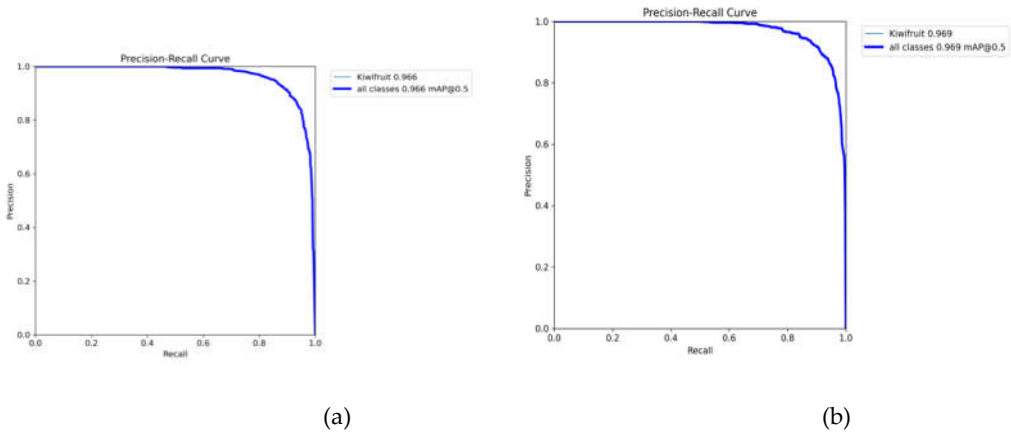
Figure 13. (a) YOLOv8n   mAP@0.5 **(b)** YOLOv8n-BFW mAP@0.5.

5.2. Comparison of Detection Models

Some of the test results of the experiment are shown in Figure 14. Compared with the file marked by the picture, it is found that the six targets marked have been detected, but the confidence of the improved model has increased slightly, as shown in the blue box marking part of the picture.

doi:10.20944/preprints202501.0400.v1

(a)   (b)
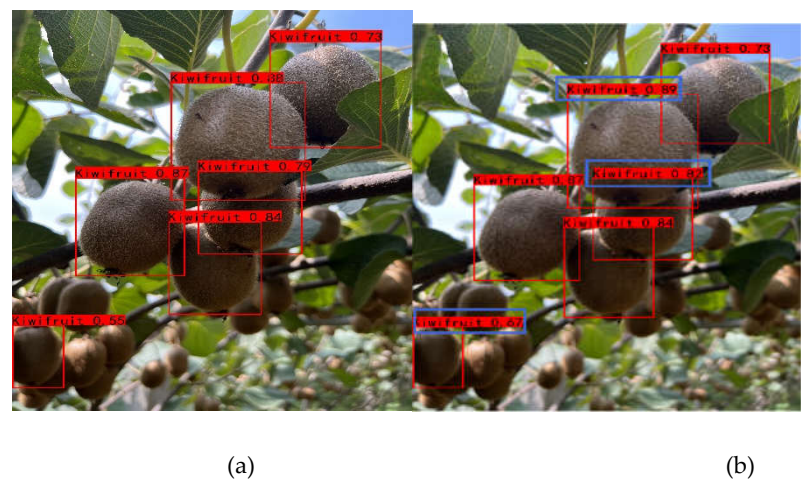
Figure 14. (a) YOLOv8n Prediction of experimental results **(b)** YOLOv8n-BFW Prediction of experimental results.   .

In order to more intuitively reflect the recognition performance of the algorithm in this paper, the algorithm in this paper is compared with different algorithms of YOLO series. In table 6, the performance indexes of five algorithms for kiwifruit recognition are given respectively. From this, it can be seen that among the five algorithms, the average accuracy mAP@0.5 and detection speed FPS of this paper have reached the highest, while its params and GFLOPs values have reached the lowest, indicating that the algorithm in this paper is more excellent as a whole, and it also ensures the recognition accuracy while realizing the lightweight of the network.

Table 6. Comparison of experimental evaluation indexes of different algorithms.

network model	map@0.5/%	Size(MB)	params/10^6	GFLOPs	FPS/(frame/s)
YOLOv3n	96.5	791	10.4	283.3	263
YOLOv5n	96.6	5.02	2.50	7.1	115
YOLOv6n	96.7	8.28	4.23	11.8	256.4
YOLOv8n	96.6	6.23	3.05	8.3	232.2
our method（YOLOv8-FBW）	96.9	4.26	2.11	6.3	294.1

6. Conclusion

This paper proposes a lightweight network for kiwifruit detection. The network is based on YOLOv8n.In order to achieve lightweight, the PANet module in the original network Neck is replaced by the weighted bidirectional feature pyramid network BiFPN module, and the C2f-Faster module is formed by referring to the FasterNet lightweight network. At the same time, the original CIOU loss function is replaced by the WIOU boundary loss function. The experimental results show that compared with the original YOLOv8n, the model size is reduced by 32%, the calculation amount is reduced by 22.2%, the parameter quantity is reduced by 29.6%, the detection speed is increased by 26.7%, and the detection accuracy of kiwifruit is increased by 0.3%. The experimental results show that this scheme is feasible. Compared with other algorithms in the YOLO series, the YOLOv8n-BFW network model proposed in this paper requires less computing resources and faster speed while maintaining high detection accuracy. It is more conducive to deployment to terminal equipment with limited computing resources, and has certain reference significance in kiwifruit production and processing.

Author Contributions: Conceptualization: W.S and D.Z; Methodology: W.S. and D.Z; Resources: D.Z and Y.Y; Validation: D.Z; Writing Original Draft: D.Z; Writing Review and Editing: W.S and Y.Y. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: The data supporting the results of this study were publicly available in [Shan, Wei (2024), ' Kiwifruit Detection ', Mendeley Data, V1, doi : 10.17632 / stm3cb6y7r.1].

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Jiang Y, Li X, Luo H, et al. Quo vadis artificial intelligence?[J]. Discover Artificial Intelligence, 2022, 2(1): 4.
2. Terven J, Córdova-Esparza D M, Romero-González J A. A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS[J]. Machine Learning and Knowledge Extraction, 2023, 5(4): 1680-1716.
3. Redmon J, Farhadi A. Yolov3: An incremental improvement[J]. arXiv preprint arXiv:1804.02767, 2018.
4. Li C, Li L, Jiang H, et al. YOLOv6: A single-stage object detection framework for industrial applications[J]. arXiv preprint arXiv:2209.02976, 2022.
5. Varghese R, Sambath M. YOLOv8: A Novel Object Detection Algorithm with Enhanced Performance and Robustness[C]//2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS). IEEE, 2024: 1-6.
6. Chen J, Ma A, Huang L, et al. Efficient and lightweight grape and picking point synchronous detection model based on key point detection[J]. Computers and Electronics in Agriculture, 2024, 217: 108612.
7. Wu M, Yun L, Xue C, et al. Walnut Recognition Method for UAV Remote Sensing Images[J]. Agriculture, 2024, 14(4): 646.
8. Ma B, Hua Z, Wen Y, et al. Using an improved lightweight YOLOv8 model for real-time detection of multi-stage apple fruit in complex orchard environments[J]. Artificial Intelligence in Agriculture, 2024, 11: 70-82.
9. Tan M, Pang R, Le Q V. Efficientdet: Scalable and efficient object detection[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020: 10781-10790.
10. Chen J, Kao S, He H, et al. Run, don't walk: chasing higher FLOPS for faster neural networks[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2023: 12021-12031.
11. Yang G, Wang J, Nie Z, et al. A lightweight YOLOv8 tomato detection algorithm combining feature enhancement and attention[J]. Agronomy, 2023, 13(7): 1824.