WHEAT GRAINS AUTOMATIC COUNTING BASED ON LIGHTWEIGHT YOLOv8
/
基于轻量化的 YOLOv8 小麦籽粒自动计数研究

Na MA[*)], Zhongtao LI, Qingzhong KONG
College of Information Science and Engineering, Shanxi Agricultural University, Taigu/China
Tel: +86-13834188480; E-mail: manasxau@163.com
DOI: https://doi.org/10.35633/inmateh-73-50

Keywords: *wheat grains, counting, object detection, YOLOv8*

ABSTRACT
In order to accurately and quickly achieve wheat grain detection and counting, and to efficiently evaluate wheat quality and yield, a lightweight YOLOv8 algorithm is proposed to automatically count wheat grains in different scenarios. Firstly, wheat grain images are collected under three scenarios: no adhesion, slight adhesion, and severe adhesion, to create a dataset. Then, the neck network of YOLOv8 is modified to a bidirectional weighted fusion BiFPN to establish the wheat grain detection model. Finally, the results of wheat grain counting are statistically analyzed. Experimental results show that after lightweight improvement of YOLOv8 with BiFPN, the mAP (mean Average Precision) value of wheat grain detection is 94.7%, with a reduction of 12.3% in GFLOPs. The improved YOLOv8 model now requires only 9.34 ms for inference and occupies just 4.0 MB of memory. Compared with other models, the proposed model in this paper performs the best in terms detection accuracy and speed comprehensively, better meeting the real-time counting requirements of wheat grains.

摘要
为了准确、快速实现小麦籽粒目标检测和计数，更为高效的对小麦品质和产量进行评估，提出一种轻量化的 YOLOv8 算法实现不同场景下小麦籽粒自动计数。首先采集无粘连、轻微粘连、重度粘连 3 种场景下的小麦籽粒图像创建数据集。然后将 YOLOv8 的neck 网络改成双向加权融合的 BiFPN，建立小麦籽粒检测模型。最后，对小麦籽粒计数结果进行统计。试验结果表明，YOLOv8 进行 BiFPN 的轻量化改进后，小麦籽粒检测 mAP 值为 94.7%，GFLOPs 减少了 12.3%。改进后的 YOLOv8 模型推理时间仅需 9.34ms，内存占用仅为 4.0MB。将改进后的 YOLOv8 模型与其他模型相比，本文提出的模型在检测精度和速度综合方面性能最优，更能满足小麦籽粒实时计数要求。

INTRODUCTION
Wheat grains are one of the primary cereal crops for humans, possessing abundant nutritional value (*Sun et al., 2021; Liu et al., 2022*). Rich in starch, proteins, fats, vitamins, and minerals, wheat grains serve as essential sources of energy and nutrients for the human body. Additionally, they are crucial for wheat breeding programs. These programs typically involve cultivating and researching wheat in small experimental fields to estimate yield. Once the actual yield meets the expected targets, the cultivation is scaled up for widespread planting. Yield estimation has always been a critical indicator in wheat breeding, and obtaining accurate counts of wheat grains through statistics is a vital factor for yield prediction (*Zhang et al., 2023*). Therefore, rapid detection and counting of wheat grains can efficiently evaluate wheat quality and yield. Moreover, it enables the adoption of more rational and scientific planting methods during wheat promotion to achieve increased production.

The traditional method for counting wheat seeds involves manually counting the wheat grains, requiring counters to be constantly attentive and patient (*Wang et al., 2023*). The counting process is extremely tedious, time-consuming, and labor-intensive, making it prone to subjective errors and leading to inconsistent results among different individuals. Additionally, the reproducibility is poor, and it is difficult to directly identify high-quality seeds with the naked eye. Therefore, there is a need for an objective, reproducible, fast, and economically reliable method of measurement.

In recent years, deep learning has been widely applied in different fields with promising results, and the integration of traditional agriculture with deep learning has become a trend (*Saleem et al., 2021; Attri et al., 2023*). *Su et al.* (*2023*) proposed an improved YOLOv5 model for detecting brown spot regions in adzuki beans. The improved YOLOv5 model achieved detection accuracy, recall rate, and mAP of 94.7%, 88.2%, and 92.5%, respectively.

Wang et al. (2024) proposed a method called YOLO-EfficientNet for identifying the growth status of hydroponic lettuce. It achieved excellent scores of 95.78 for Val-acc, 94.68 for Test-acc, 96.02 for Recall, 96.32 for Precision, and 96.18 for F1-score. *Gai et al. (2023)* proposed an improved YOLOv4 deep learning algorithm for detecting cherry fruits. It is proposed to increase the network based on the YOLO-V4 backbone network CSPDarknet53 network, combined with DenseNet. The density between layers, the a priori box in the YOLO-V4 model, is changed to a circular marker box that fits the shape of the cherry fruit. The results show that the mAP value obtained by using the improved YOLO-V4 model network in this paper is 0.15 higher than that of YOLOv4. To address weed detection and identification amidst complex field backgrounds, *Peng et al. (2023)* integrated the CBAM attention mechanism into YOLOv7. The enhanced algorithm achieved an average precision (mAP) of 91.15%, surpassing the original YOLOv7 algorithm by 2.06%. Researchers have conducted extensive studies on seed counting using deep learning methods. *Zhao et al. (2023)* proposed a P2PNet-Soy method for automatic counting of soybean seeds. The superiority of the proposed P2PNet-Soy in soybean seed counting and localization over the original P2PNet was confirmed by a reduction in the value of the mean absolute error, from 105.55 to 12.94. *Xiang et al. (2023)* proposed a method called YOLO POD for counting soybean pods. Building upon YOLO X, they added a block for predicting pod quantity, modified the loss function, and introduced the Convolutional Block Attention Module (CBAM). The results show that the R2 between the predicted and actual pod counts by YOLO POD reached 0.967, which is an improvement of 0.049 over YOLO X. A lightweight real-time wheat seed detection model, YOLOv8-HD, based on YOLOv8, was previously proposed. YOLOv8-HD achieves an average precision (mAP) of 99.3%, which is 16.8% higher than YOLOv8. The memory footprint of the YOLOv8-HD model is 6.35 MB (*Ma et al., 2024*).

From the above research, it can be seen that there is an urgent need for a method that can perform fast and accurate real-time counting of seed. Therefore, this study proposes a wheat grain real-time detection model named YOLOv8-BIFPN, focusing on counting accuracy and speed. Firstly, images of wheat grains in different scenarios were collected as the dataset. Then, building upon the YOLOv8 base model (*Yang et al., 2023; Hussain et al., 2023*), BiFPN with bidirectional weighted fusion was introduced to accelerate the model's execution speed and achieve wheat grain object detection (*Yu et al., 2024; He et al., 2023*). Furthermore, the improved YOLOv8-BiFPN was compared with current mainstream object detection algorithms. The results showed that the proposed algorithm has significant advantages in both wheat grain detection accuracy and speed, meeting the practical demand for wheat grain counting and laying a foundation for the development of precision agriculture technology.

MATERIALS AND METHODS

Data Acquisition and Pre-processing

The wheat grains used in the experiment are of the variety Longmai 6197. These grains are compact, resistant to lodging, and have stable high yield characteristics, making them a new variety of drought-resistant and high-yielding wheat. Three scenarios were set up for the experiment: wheat with slight adhesion, wheat without adhesion, and wheat with severe adhesion. A total of 300 images were collected, with 100 images for each scenario. Some of the data are shown in Figure 1, where the wheat grains are placed on a black cushion to enhance the contrast of the images. The photographic equipment used in the experiment was the Huawei P40 Pro Plus smartphone. A certain amount of wheat grains were mixed with wheat husks and straw, and then lightly shaken to randomly scatter them on the experimental platform, ensuring the random distribution of wheat grains, husks, and straw.

(a) no adhesion (b) slight adhesion (c) severe adhesion

Fig.1 - Partial Data Collection

Using the LabelImg software for annotation, wheat grains are labeled as "w", husks are labeled as "k" and stalks are labeled as "g". After annotating the data, it is randomly divided into training, validation, and test sets in a ratio of 7:2:1. Finally, there are 210 images in the training set, 60 images in the test set, and 30 images in the validation set.

YOLOv8 Model

YOLOv8 is an advanced SOTA model developed by Ultralytics. Building upon the success of previous YOLO versions, it introduces new features and improvements, further enhancing its performance and flexibility. Based on the principles of speed, accuracy, and ease of use, YOLOv8 is an excellent choice for tasks such as object detection, image segmentation, and image classification (*Terven et al., 2023*). YOLOv8 is capable of performing detection, classification, and segmentation tasks. The main improvement in YOLOv8 is the C2f module, depicted in Figure 2. It is a further refinement of ELAN, primarily replacing the convolutions in ELAN's computational modules with Bottleneck modules. ELAN serves as the main module for the Backbone. ELAN consists of three parts: Cross Stage Partial; Computational module; 1x1 convolutions, as shown in Figure 3.

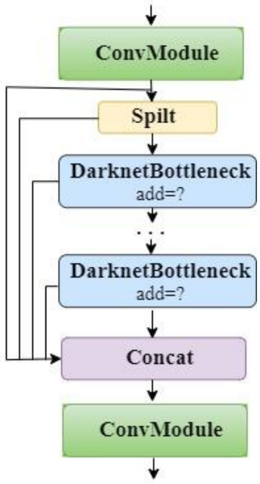


Fig. 2 - C2f model

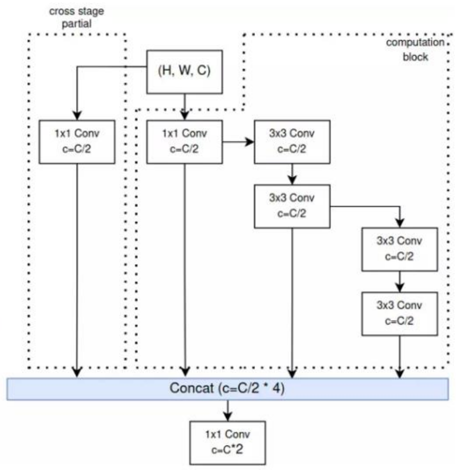


Fig. 3 - ELAN model

BiFPN Model

BiFPN (Bi-directional Feature Pyramid Networks) is an enhanced version of the FPN network structure. FPN relies on a backbone and incorporates bottom-up, top-down, feature fusion, lateral connections, and feature integration, primarily for object detection tasks. This structure is weighted and bidirectional, involving both top-down and bottom-up structures. By constructing bidirectional channels, it achieves cross-scale connections, directly fusing features from the feature extraction network with relatively large features from the bottom-up path. It preserves shallower semantic information without losing too much deep semantic information.

BiFPN features bidirectional connections, adaptive feature adjustment, modular design, efficiency, and performance improvement (*Chen et al., 2021*). It allows information to propagate bidirectionally between different resolution levels, better integrating low-level and high-level features, thereby enhancing object detection accuracy. The specific structure of BiFPN is illustrated in Figure 4.

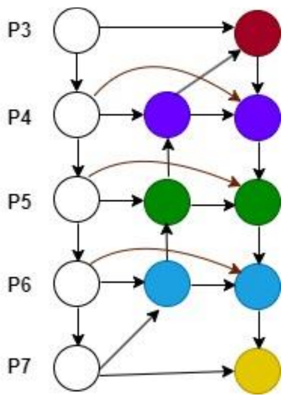


Fig. 4 - BiFPN model

Improved YOLOv8

The counting of wheat grains is easily disrupted by wheat stalks and husks, especially since husks closely resemble wheat grains. Additionally, accurately detecting clumped wheat grains presents a significant challenge. In order to enhance the detection performance of wheat grains, especially under conditions of adhesion and impurities, the YOLOv8 neck network was replaced with BiFPN. The introduction of BiFPN allows information to propagate across different resolutions, aiding the model in better understanding the features of targets in various scenarios. The improved structure diagram is shown in Figure 5.

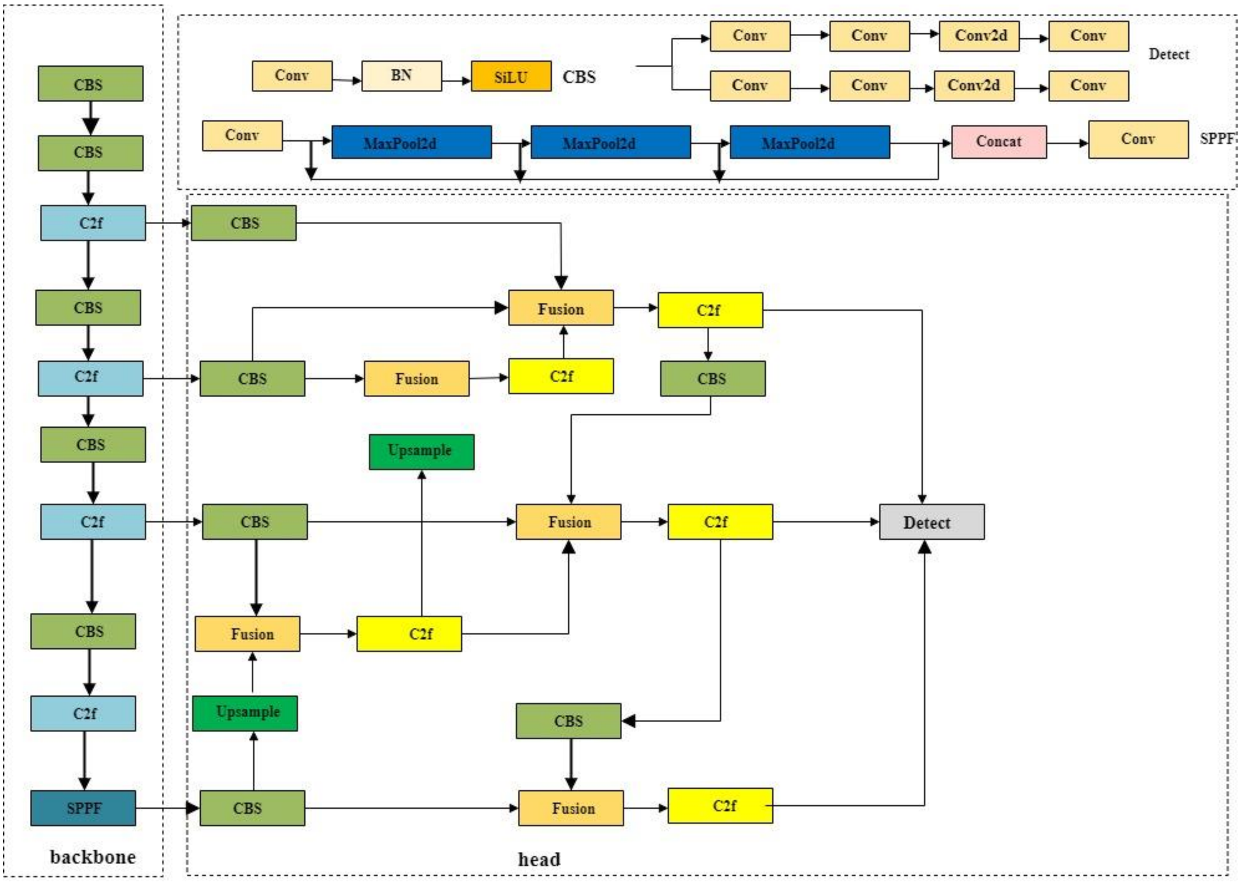


Fig. 5 - Improved YOLOv8 Architecture Diagram

Classic detection models YOLOv3, YOLOv5, YOLOv6, YOLOv8 were used to detect wheat seed data. The model with the highest detection accuracy and fastest detection speed was selected from the experimental results, and further improved using BiFPN to select the optimal wheat grain counting model.

Model Evaluation Metrics

Wheat grain object detection needs to balance accuracy and speed. Therefore, this experiment employs precision (P), recall (R), and mean average precision (mAP). The confusion matrix is a metric used to evaluate model results, which is part of model assessment. It calculates metrics such as precision and recall. TP: True Positive, predicting positive class as positive; FP: False Positive, predicting negative class as positive; FN: False Negative, predicting positive class as negative; TN: True Negative, predicting negative class as negative.

Precision calculation formula is:

$$P=\frac{TP}{FP+TP} \tag{1}$$

Recall calculation formula is:

$$R=\frac{TP}{TP+FN} \tag{2}$$

mAP calculation formula is:

$$AP=\frac{TN+TP}{TP+TN+FP} \tag{3}$$

$$mAP=\frac{\sum_{i=1}^{C} AP_i}{C} \tag{4}$$

FLOPs stands for floating point operations, which can be used to measure the complexity of algorithms/models. GFLOPs represents one billion floating point operations. In the field of deep learning, especially during the training and inference processes of models, GFLOPs are often used to describe the computational requirements of models and the computing capabilities of hardware.

Grad-CAM heat maps are a visualization method used in object detection to understand the decision-making process of deep neural networks in image classification tasks (*Wang et al., 2023*). Through these heat maps, it can be observed which regions of the input image are crucial for the network model's decision-making during classification. The computation of the heat map involves calculating the gradient weights of the target class with respect to the feature maps of the network's final layer, multiplying these weights with the feature maps to obtain the heat map. This process visualizes the network model's attention to different regions. When observing the heat map, the importance of corresponding regions can be judged based on the intensity of colors. Brighter colors indicate that the features in that region significantly contribute to the detection results, while darker colors suggest less significant contributions.

RESULTS

The operating system used for the experiment was Windows 11. The GPU model was NVIDIA GeForce RTX 2060. The CPU model was Intel(R) Core(TM) i7-10875H CPU @ 2.30GHz. The system memory was 16GB and the solid-state drive had 512GB. The GPU acceleration libraries used were CUDA 11.6 and cuDNN 8.7. The Python version used was Python 3.8, and the deep learning framework PyTorch 1.13.1. The image size for deep learning training was 640×640 pixels, with 12 images per batch, and 100 epochs of training. In the experiment, a pre-trained model was used for transfer learning. Applying pre-training as the basis for transfer learning in a new task can significantly improve model performance, reduce training time, and consume fewer computational resources.

Comparison of Detection Performance of Different Models

Using YOLO series algorithms to detect wheat grains, the results are shown in Table 1. From Table 1, it can be seen that the detection accuracy of YOLOv3 (*Lawal et al., 2021*) and YOLOv6 (*Li et al., 2022*) is not high and the speed is slow. Although YOLOv8 has high detection accuracy, its detection GFLOPs are not as good as YOLOv5 (*Jocher et al., 2022*). Therefore, in order to balance detection speed and accuracy, improvements were made to YOLOv8 and YOLOv5.

Table 1

Results of different YOLO algorithms detection

Scene	Category	P/%	R/%	$mAP_{0.5}$/%	GFLOPs
YOLOv3	Mean	89	87.2	93.5	282.2
	Wheat seeds	98.5	97.9	99.4	
	Straw	79.4	80	86.5	
	Husk	89	83.7	94.5	
YOLOv5	Mean	90.6	70.3	86.4	7.1
	Wheat seeds	99	95.6	99.2	
	Straw	78.7	50	68	
	Husk	94.1	65.1	92	
YOLOv6	Mean	87.2	76.6	84.8	11.8
	Wheat seeds	96.9	97.2	99.2	
	Straw	71.1	65	69.5	
	Husk	93.6	67.6	85.8	
YOLOv8	Mean	97.6	96.8	95	8.1
	Wheat seeds	99.9	97.4	99.3	
	Straw	79.5	80.2	86.9	
	Husk	95.9	93	98.8	

Improved Algorithm Performance

After improving YOLOv8 and YOLOv5 with BiFPN, the results are shown in Table 2. YOLOv5-BiFPN achieved an AP of 99.3% for wheat grain detection, an AP of 76.7% for husk detection, an AP of 97% for straw detection, and an mAP of 91.4%. YOLOv8-BiFPN achieved an AP of 99.5% for wheat grain detection, an AP of 87.3% for husk detection, an AP of 97.4 for straw detection, and an mAP of 94.7%. The GFLOPs of YOLOv5-BiFPN were 6.4, and those of YOLOv8-BiFPN were 8.1, the same as YOLOv5. From the above experimental results, it can be seen that the mAP of YOLOv8-BiFPN in multiple scenarios is 94.7%, which is respectively 8.3% higher than YOLOv5, and the AP of wheat grain, straw, and husk is respectively 0.3%, 19.3%, and 5.4% higher than YOLOv5, and 0.2%, 10.6%, and 0.4% higher than YOLOv5-BiFPN. Although the GFLOPs of YOLOv5-BiFPN decreased, its detection accuracy did not surpass that of YOLOv8.

Table 2

Improved algorithm detection results

Scene	Category	P/%	R/%	$mAP_{0.5}$/%	$mAP_{0.5:0.95}$/%	GFLOPs
YOLOv5-BiFPN	Mean	87.7	85.5	91.4	67.9	6.4
	Wheat seeds	98	96.6	99.3	81.7	
	Straw	67.7	73.4	76.7	52.9	
	Husk	97.4	86.4	97	69	
YOLOv8-BiFPN	Mean	95.6	91.2	94.7	75.1	7.1
	Wheat seeds	99	98.7	99.5	83.9	
	Straw	93.7	75	87.3	63.4	
	Husk	94.2	1	97.4	78	

To further compare the performance of the improved algorithms, this study plotted the loss functions of YOLOv3, YOLOv5, YOLOv6, YOLOv5-BiFPN, and YOLOv8-BiFPN, as shown in Figure 6. From Figure 6, it can be observed that the box_loss, cls_loss, and dfl_loss curves of YOLOv3, YOLOv5, YOLOv6, YOLOv5-BiFPN, and YOLOv8-BiFPN all continuously decrease during 100 epochs of training until they stabilize. Compared to other YOLO series algorithms, YOLOv8-BiFPN exhibits faster convergence speed on the training set, effectively extracting wheat grain features and accelerating the model convergence rate.

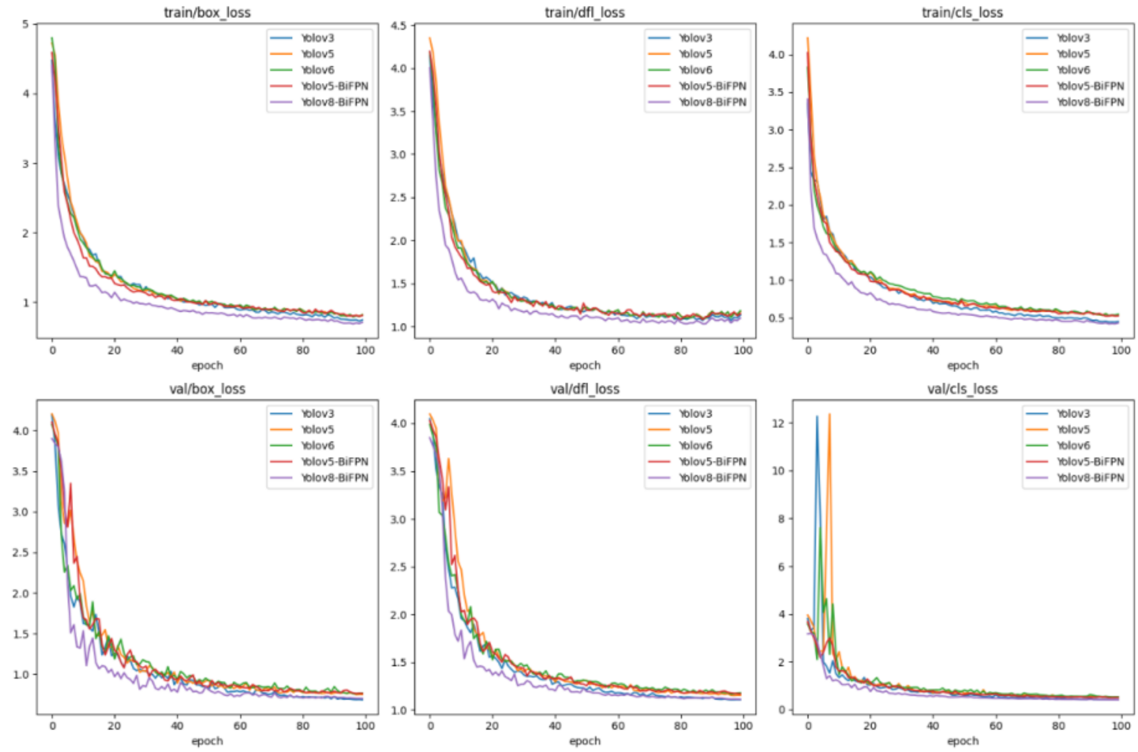


Fig. 6 - Comparison of loss functions for different models

This study compared the precision, recall, $mAP_{0.5}$, and $mAP_{0.5:0.95}$ of YOLOv3, YOLOv5, YOLOv6, YOLOv5-BiFPN, and YOLOv8-BiFPN models, as shown in Figure 7. It can be observed from Figure 7 that YOLOv8-BiFPN outperforms other algorithms in all aspects.

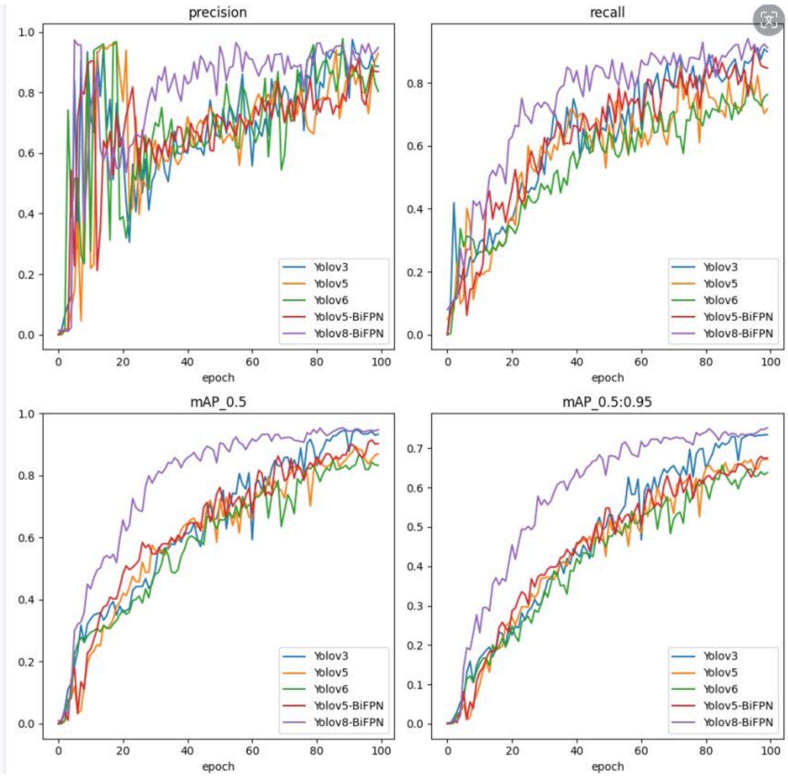


Fig. 7 - Performance curves of different algorithms

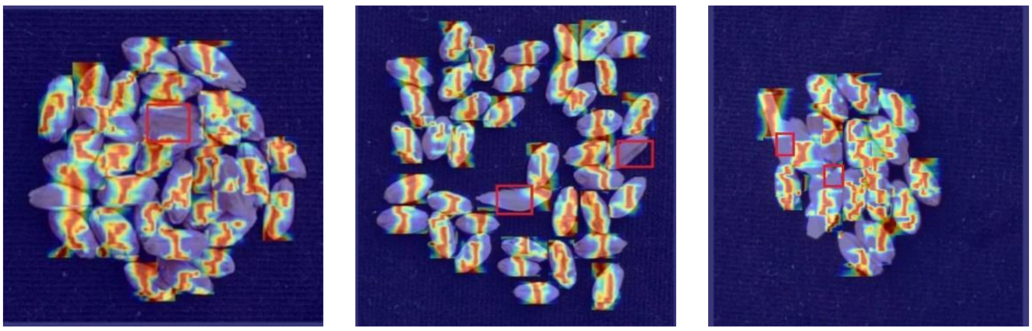
Table 3 compares the speed of the YOLOv8-BiFPN and YOLOv5-BiFPN models using parameters such as FPS, Latency, and size. FPS indicates the number of photos that can be processed per second, serving as a measure of detection speed. Latency represents the inference time, while size denotes the model's memory footprint. From Table 3, it can be observed that although YOLOv8-BiFPN has slightly larger memory requirements, both the inference speed and detection speed of YOLOv8-BiFPN are superior to those of YOLOv5-BiFPN.

Table 3

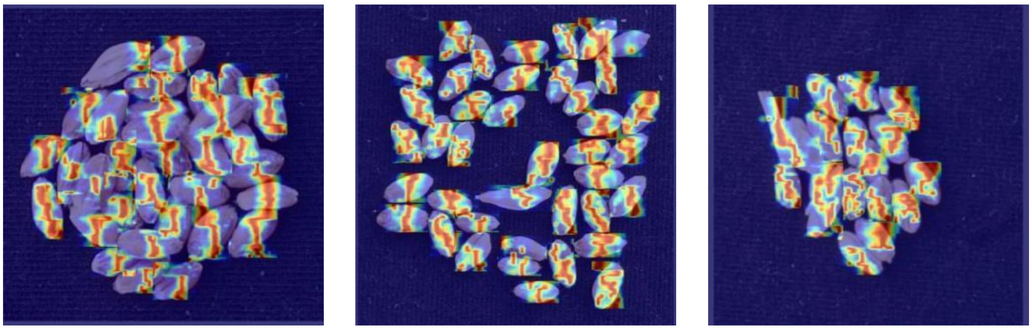
Comparison of Lightweighting Between YOLOv8-BiFPN and YOLOv5-BiFPN Models

Model	Latency	FPS	size
YOLOv8-BiFPN	9.34ms	107.1	4.0M
YOLOv5-BiFPN	12.61ms	79.3	3.5M

To better compare the feature extraction capabilities of the models, heatmaps of wheat grain detection using YOLOv8-BiFPN and YOLOv5-BiFPN were plotted as shown in Figure 8. The portion circled in red boxes in the figure indicates areas where the feature extraction of YOLOv5-BiFPN is less prominent. This demonstrates that the feature extraction capability of YOLOv8-BiFPN is superior to that of YOLOv5-BiFPN.



(a)YOLOv8-BiFPN Heatmaps



(b)YOLOv5-BiFPN Heatmaps

Fig. 8 - Comparison of Heatmaps

Counting Wheat Grains

Using the YOLOv8-BiFPN model for counting in different scenarios yields the results depicted in Figure 9. From the figure, it's evident that the YOLOv8-BiFPN model can effectively detect wheat grains, stalks, and husks in various scenes, accurately determining their quantities.

(a)no adhesion (b)slight adhesion (c)severe adhesion

Fig. 9 - Counting results in different scenarios

The wheat grain target detection and counting results were obtained using YOLOv8-BiFPN and YOLOv5-BiFPN models under scenarios of no adhesion, slight adhesion, and severe adhesion, with a total of 60 images, 20 for each scenario. The counting results are presented in Tables 4 and 5. Calculation errors include errors in wheat grain counting, wheat husk counting, and wheat straw counting. Detection errors encompass misidentification, missed detection, and duplicate detection. Due to the prevalence of detection errors, such as mistaking wheat husks for wheat grains or wheat husks for wheat straws, detailed classification of detection errors was not provided here.

Due to situations where counting errors and detection errors occur simultaneously, the total correct count for various scenarios does not equal 20. From the Tables 4 and 5, it can be observed that YOLOv8-BiFPN exhibits high detection and counting accuracy. In scenarios without adhesion, the counting and detection accuracy of YOLOv8-BiFPN is 70%, with a counting error rate of 25% and a detection error rate of 15%. In scenarios with slight adhesion, the counting and detection accuracy is 75%, with a counting error rate of 15% and a detection error rate of 15%. In scenarios with severe adhesion, the counting and detection accuracy is 55%, with a counting error rate of 25% and a detection error rate of 40%.

YOLOv5-BiFPN, on the other hand, achieves a counting and detection accuracy of 35% in scenarios without adhesion, with a counting error rate of 30% and a detection error rate of 45%. In scenarios with slight adhesion, the counting and detection accuracy is 40%, with a counting error rate of 25% and a detection error rate of 60%. In scenarios with severe adhesion, the counting and detection accuracy is 10%, with a counting error rate of 60% and a detection error rate of 55%.

Table 4

YOLOv8-BiFPN counting results

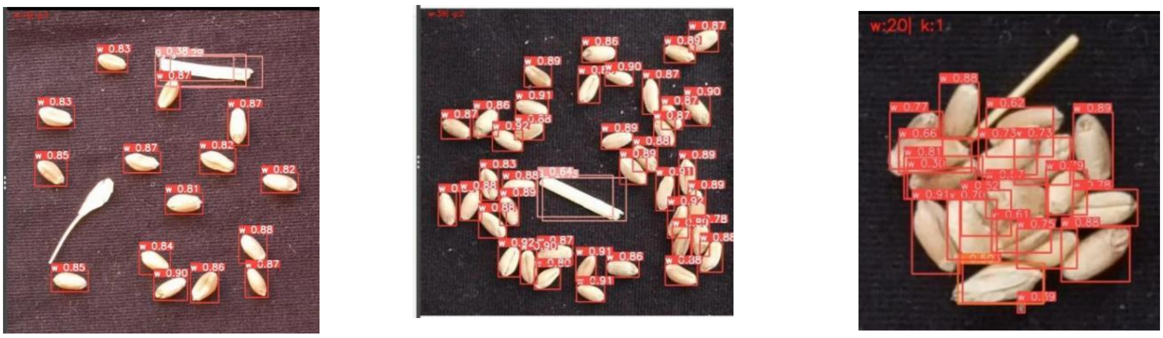
Scene	counting and detection correct	counting errors	Detection errors
no adhesion	14	5	3
slight adhesion	15	3	3
severe adhesion	11	5	8

Table 5

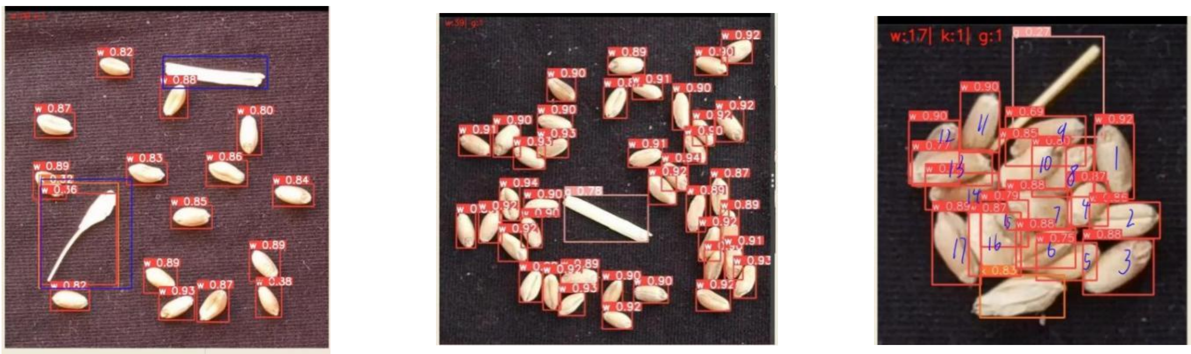
YOLOv5-BiFPN counting results

Scene	counting and detection correct	counting errors	Detection errors
no adhesion	7	6	9
slight adhesion	8	5	12
severe adhesion	2	12	11

YOLOv5-BiFPN and YOLOv8-BiFPN partial counting results are shown in Figure 10. In the first image of Figure 10(a), YOLOv5-BiFPN redundantly detects wheat stems and misses wheat husks, while in the first image of Figure 10(b), YOLOv8-BiFPN falsely detects wheat husks as wheat grains and misses wheat stems. In the second image of Figure 10(a), YOLOv5-BiFPN redundantly detects wheat stems, whereas in the second image of Figure 10(b), YOLOv8-BiFPN detects and counts correctly. In the third image of Figure 10(a), YOLOv5 incorrectly detects wheat husks as wheat grains, misses wheat stems, and duplicates counts of wheat grains. However, in the third image of Figure 10(b), YOLOv8-BiFPN detects and counts accurately.



(a)YOLOv5-BiFPN



(b)YOLOv8-BiFPN

Fig. 10 - Partial detection and counting result images

Based on the analysis above, it can be seen that the improved algorithms perform poorly in counting and detecting wheat grains in severely adhesive scenarios compared to other situations. This indicates that wheat grains are difficult to identify in severe adhesion conditions. Meanwhile, comparing experimental data reveals that both models show improved detection and counting accuracy in mildly adhesive situations. When comparing the counting errors and detection errors of YOLOv5-BiFPN, it is evident that the detection error rate is higher than the counting error rate.

Analysis of the experimental result graphs shows that YOLOv5-BiFPN exhibits a higher detection error rate for wheat husks and stems. The graphs frequently show misidentifications of wheat husks as stems and issues with repeated detection of stems.

In conclusion, the detection and counting accuracy of YOLOv8-BiFPN are higher than those of YOLOv5-BiFPN. Therefore, selecting YOLOv8-BiFPN for wheat grain detection and counting is a scientifically reasonable approach.

CONCLUSIONS

This study successfully developed a lightweight wheat grain detection model based on the improved YOLOv8, named YOLOv8-BiFPN, aiming to improve the accuracy and speed of wheat grain counting.

(1) This study created three scenarios of wheat grain datasets: non-adherent, slightly adherent, and severely adherent.

(2) BiFPN was introduced into the neck network of YOLOv8 in this study to achieve lightweighting of the YOLOv8 model. The improved model achieved a wheat grain detection mAP of 94.7%, with a 12.3% reduction in GFLOPs. The inference time of the improved YOLOv8 model was only 9.34 ms, with a memory footprint of only 4.0 MB, meeting the real-time counting requirements for wheat grains.

(3) The improved model YOLOv8-BiFPN was compared with YOLOv3, YOLOv5, YOLOv8, and YOLOv5-BiFPN in this study, showing that YOLOv8-BiFPN was superior in both detection accuracy and speed.

(4) Using the improved YOLOv8-BiFPN for wheat grain counting in different scenarios, the counting and detection accuracy were 70% in the non-adherent scenario, 75% in the slightly adherent scenario, and 55% in the severely adherent scenario. The algorithm's performance in severely adherent scenarios was lower compared to other scenarios, indicating difficulties in wheat grain identification under severe adherence conditions.

The shortcomings in this experiment lie in the decreased detection accuracy of the improved YOLOv8-BiFPN compared to YOLOv8. Therefore, further research can be conducted to improve YOLOv8 in order to find a wheat grain detection algorithm that enhances accuracy while maintaining faster detection speed. Additionally, the current algorithm exhibits relatively lower detection and counting accuracy in severely cluttered scenes. Optimization of the model can be continued to enhance the accuracy of wheat grain detection in severely cluttered scenarios.

ACKNOWLEDGEMENT

This research, titled “Wheat grains automatic counting based on lightweight YOLOv8”, was funded by the Basic Research Program of Shanxi Province (202303021212115).

REFERENCES

[1] Attri I, Awasthi L K, Sharma T P, et al. (2023). A review of deep learning techniques used in agriculture [J]. *Ecological Informatics*, 102217.

[2] Chen J, Mai H S, Luo L, et al. (2021). Effective feature fusion network in BIFPN for small object detection[C]// *IEEE international conference on image processing (ICIP).* IEEE, 699-703.

[3] Gai R, Chen N, Yuan H. (2023). A detection algorithm for cherry fruits based on the improved YOLO-v4 model [J]. *Neural Computing and Applications*, 35(19): 13895-13906.

[4] He L, Wei H, Wang Q. (2023). A New Target Detection Method of Ferrography Wear Particle Images Based on ECAM-YOLOv5-BiFPN Network [J]. *Sensors*, 23(14): 6477.

[5] Hussain M. (2023). YOLO-v1 to YOLO-v8, the rise of YOLO and its complementary nature toward digital manufacturing and industrial defect detection [J]. *Machines*,11(7): 677.

[6] Jocher G, Chaurasia A., Stoken A., et al. (2022). Ultralytics/YOLOv5: v6. 2-YOLOv5 Classification models, Apple M1, Reproducibility, ClearML and Deci.ai integrations [J]. *Zenodo*,12-16.

[7] Lawal M.O. (2021). Tomato detection based on modified YOLOv3 framework [J]. *Scientific Reports*, 11(1): 1-11.

[8] Li C., Li L., Jiang H., et al. (2022). YOLOv6: A single-stage object detection framework for industrial applications [J]. *arxiv preprint arxiv:2209*.

[9] Liu X. (2022). *Research on Automatic Counting of wheat seed based on Image Processing*. Bachelor’s Thesis, Anhui Agriculture University, Hefei, China (In Chinese with English abstract).

[10] Ma N., Su Y., Yang L., et al. (2024). Wheat Seed Detection and Counting Method Based on Improved YOLOv8 Model [J]. *Sensors*, 24(5): 1654.
[11] Peng M., Zhang W., Li F., et al. (2023). Weed detection with Improved Yolov 7[J]. EAI Endorsed *Transactions on Internet of Things*, 9(3): 1-2.
[12] Saleem M.H., Potgieter J., Arif K.M. (2021). Automation in agriculture by machine and deep learning techniques: A review of recent developments [J]. *Precision Agriculture*, 22(6): 2053-2091.
[13] Su P., Li H., Wang X., et al. (2023). Improvement of the YOLOv5 Model in the Optimization of the Brown Spot Disease Recognition Algorithm of Kidney Bean [J]. *Plants*, 12(21): 3765.
[14] Sun J., Zhang L., Zhou, X., et al. (2021). Detection of rice seed vigor level by using deep feature of hyperspectral images. *Transactions of the CSAE*, 37(14): 171-178. (in Chinese with English abstract)
[15] Terven J., Córdova-Esparza D.M, Romero-González J.A. (2023). A comprehensive review of YOLO architectures in computer vision: From YOLOv1 to YOLOv8 and YOLO-nas [J]. *Machine Learning and Knowledge Extraction*, 5(4): 1680-1716.
[16] Wang L., Zhang Q., Feng T. et al. (2023). Research on wheat grain counting method based on YOLOv7-ST model [J]. Transactions of the Chinese Society for Agricultural Machinery, 54(10): 188-197.
[17] Wang S., Zhang Y. (2023). Grad-CAM: understanding AI models [J]. *Comput. Mater. Contin*, 76: 1321-1324.
[18] Wang Y., Wu M., Shen Y. (2024). Identifying the Growth Status of Hydroponic Lettuce Based on YOLO-EfficientNet [J]. *Plants*,13(3): 372.
[19] Xiang S., Wang S., Xu M. et al. (2023). YOLO POD: a fast and accurate multi-task model for dense Soybean Pod counting [J]. *Plant methods*, 19(1): 8.
[20] Yang G., Wang J., Nie Z., et al. (2023). A lightweight YOLOv8 tomato detection algorithm combining feature enhancement and attention [J]. *Agronomy*, 13(7): 1824.
[21] Yu C., Shin Y. (2024). SAR ship detection based on improved YOLOv5 and BiFPN [J]. *ICT Express*, 10(1): 28-33.
[22] Zhang H., Ji J., Ma H. et al. (2023). Wheat Seed Phenotype Detection Device and Its Application. *Agriculture*, 13, 706.
[23] Zhao J., Kaga A., Yamada T. et al. (2023). Improved field-based soybean seed counting and localization with feature level considered [J]. *Plant Phenomics*, 5: 0026.