Journal of Agricultural Science and Technology A 16 (2026) 68-78
doi: 10.17265/2161-6256/2026.02.002

Research on Variety Detection of Highland Barley Seeds Based on Improved YOLOv8

Yanbo WANG[1], Ruokui CHANG[1], Yongcheng JIANG[1], Xin WANG[1], Haifeng WANG[2], Shide LI[3], and Hua LIU[1]

1. College of Engineering and Technology, Tianjin Agricultural University, Tianjin 300392, China

2. Gannan Tibetan Autonomous Prefecture Academy of Agricultural, Forestry and Pasture Sciences, Gannan, Gansu 747000, China

3. Department of English Teaching, Tianjin Agricultural University, Tianjin, 300384, China

Abstract: As an important characteristic crop in the Qinghai-Tibet Plateau, highland barley variety detection technology plays a key role in breeding improvement, food processing, and germplasm resource management. In recent years, deep learning-based machine vision methods have provided new ideas for seed detection. However, their application to highland barley, a special crop, still faces challenges. Most existing models are designed for plain crops and struggle to adapt to the phenotypic characteristics of highland barley seeds in high-altitude environments. Additionally, there is a lack of systematic phenotypic datasets to support model training. In view of this, this paper proposes a variety detection model for highland barley grains. By introducing an attention mechanism module and conducting multi-scale feature optimization on the head network of YOLOv8, the small target detection performance is enhanced. Furthermore, a loss calculation method more suitable for the characteristics of small target detection is adopted to further improve the overall detection performance of the model. Experimental results demonstrate that the effectiveness of the improved algorithm in this paper is significant. Compared with the original YOLOv8 model, the mean average precision (mAP) is increased by 6.1 percentage points, reaching 95.1%.

Key words: Highland barley, variety detection, convolutional neural networks, YOLOv8, attention mechanisms.

1. Introduction

Highland barley [1], as an important cereal crop in the Qinghai-Tibet Plateau, exhibits strong adaptability to harsh environmental conditions, including drought resistance, cold tolerance, and low soil fertility requirements. It plays a crucial role in ensuring food security and supporting sustainable agricultural development in plateau regions. Therefore, accurate detection of highland barley seed varieties [2] is of great significance for breeding improvement, seed purity inspection, germplasm resource management, and intelligent seed sorting.

With the rapid development of artificial intelligence, deep learning-based machine vision methods have been widely applied in agricultural seed detection tasks. Compared with traditional manual identification methods, machine vision techniques offer advantages such as non-contact operation, high efficiency, and automation potential. Existing studies have demonstrated the effectiveness of deep learning in seed analysis tasks. Ding et al. [3] applied machine learning and deep learning methods for maize seed vigor detection, while Bai et al. [4] proposed a YOLOv5-based method for wheat seed germination detection, confirming the feasibility of object detection algorithms in seed analysis.

In recent years, YOLO-based models have been extensively studied in agricultural applications. For example, Wang et al. [6] proposed a YOLOv5 model improved by optimized CBAM for crop pest identification. Wang et al. [7] developed YOLOv6-ESG for lightweight seafood detection. Liu et al. [8] reported a rice grain detection method based on

Corresponding authors: Ruokui CHANG, M.E., associate professor, research fields: intelligent equipment and key technologies, agricultural automation systems; Yongcheng JIANG, Ph.D., professor, research fields: intelligent control technology of modern agricultural equipment, intelligent detection and control technology.

YOLOv7 fused with GhostNetV2. In addition, improved YOLO architectures have also been applied to corn silk detection [9], wheat seed health recognition [10], and dense agricultural object detection tasks [11]. Although these studies demonstrate the effectiveness of YOLO-based models in agricultural seed detection, they mainly focus on common crops such as maize, wheat, rice, and cotton, and their methods cannot be directly generalized to highland barley seed variety detection.

However, the application of deep learning models to highland barley seed variety detection remains limited. First, most existing models are designed for common crops grown in plain regions and may not effectively adapt to the morphological characteristics of highland barley seeds in plateau environments [5]. Second, publicly available phenotypic datasets for highland barley seed varieties are still lacking, which limits model training and evaluation. To address these issues, this study constructs a phenotypic image dataset of six highland barley seed varieties and proposes an improved YOLOv8-based detection model named YOLOv8-SDW. Specifically, the SimAM attention mechanism [13] is introduced to enhance feature extraction, DyHead [16-18] is adopted to improve multi-scale feature fusion and small-object detection performance, and the WIoU loss function [19, 20] is employed to optimize bounding box regression.

The proposed method aims to improve detection accuracy while maintaining high inference speed, thereby providing technical support for intelligent highland barley seed variety detection in plateau agriculture.

2. Materials and Methods

2.1 Sample Collection and Preprocessing

2.1.1 Sample Collection

In this study, six different varieties of highland barley seeds were selected, including the new strains 0628, 0349-1, Ganqing 4, Ganqing 6, Ganqing 8, and Ganqing 9 (Figure 1). The samples were obtained from the Agricultural Science Research Institute of Gannan Tibetan Autonomous Prefecture. During sample preparation, damaged, shriveled, and visually abnormal seeds were manually removed, and only plump seeds with intact morphology were retained. Subsequently, seed varieties were identified by agricultural experts, and a double-check verification procedure was conducted to ensure annotation accuracy and label consistency. Image acquisition was performed using a linear scanning camera under natural lighting conditions, with a resolution of 3840 × 2160 pixels, and images were stored in JPG format. To reduce the influence of complex backgrounds on recognition performance, all seeds were randomly tiled on a pure white background plate for imaging. It should be noted that the current image acquisition setting represents a relatively controlled laboratory environment, which differs from real-world seed sorting scenarios involving complex backgrounds and varying illumination conditions. Therefore, the generalization ability of the proposed model in practical deployment still requires further validation.

Considering that a single seed variety may be mixed with a small proportion of other varieties in practical production, two datasets were constructed: a single-variety dataset and a mixed-variety dataset. Each single-variety image contains 10-50 seeds, while each mixed-variety image contains 10-50 seeds of the dominant variety and 2-10 seeds from other varieties.

A total of 648 original images were collected, including 401 single-variety images and 247 mixed-variety images. Although the number of original images is limited, each image contains multiple seed targets, resulting in approximately 21368 annotated seed instances (bounding boxes) in total, which provides sufficient object-level samples for model training.

2.1.2 Data Preprocessing

To address the potential overfitting issue caused by the relatively limited size of the original dataset,

conventional image augmentation methods were first adopted to expand the dataset and improve model generalization ability, and then the Mosaic data augmentation [12] technique was used to enrich dataset diversity. The specific steps are as follows: first, four images are randomly selected from the dataset, which may contain objects of different categories and different backgrounds; then each image is randomly cropped, with the size and position of the cropped area determined randomly, and the cropped images may have different sizes; next, the four cropped images are spliced into a large image, usually in a 2 × 2 grid layout, that is, the four images are placed in the four quadrants of the large image respectively; according to the new coordinates after image splicing, the positions of the target bounding boxes in each image are adjusted to ensure that the bounding boxes still accurately enclose the target objects; finally, the spliced large image is resized to the input size required by the model.

Before model training, the collected original highland barley seed dataset was first divided into training, validation, and test sets using a stratified sampling strategy at an approximate ratio of 7:2:1 to ensure balanced category distribution across subsets. Subsequently, data augmentation was independently performed within each subset to expand sample size and improve dataset diversity. The augmented validation and test sets were used only to assess model stability under transformed conditions rather than to simulate strictly real-world unseen distributions.

It should be noted that data augmentation was performed independently within the training, validation, and test sets after dataset splitting. Since the original images in each subset were mutually independent, no overlap occurred among the datasets. This preprocessing strategy was adopted to increase sample diversity while maintaining balanced category distributions across all subsets.

Table 1 presents the detailed distribution of the highland barley seed dataset. Ultimately, the training set contains 1,080 single-variety images and 720 mixed-variety images, totaling 1,800 highland barley seed images. The validation set comprises 300 single-variety images and 180 mixed-variety images, with a total of 480 images. The test set includes 120 single-variety images and 80 mixed-variety images, amounting to 200 images in total.

a. New Product Series 0628 b. New Product Series 0349-1 c. Ganqing-4

d. Ganqing-6 f. Ganqing-8 g. Ganqing-9

Fig. 1 Samples of six highland barley varieties.

Table 1 Highland barley seed image dataset.

Dataset	Single-variety (before augmentation)	Single-variety (after augmentation)	Mixed-variety (before augmentation)	Mixed-variety (after augmentation)
Training set	311	1,080	189	720
Validation set	56	300	35	180
Test set	34	120	23	80
Total images	401	1,500	247	980

2.2 Model Optimization

2.2.1 YOLOv8-SDW Model

In the process of highland barley variety identification and detection, algorithm optimization was performed on the YOLOv8n model. First, a global attention mechanism was added to the backbone network of the YOLOv8n model, to enhance the model's ability to extract high-level semantic features of highland barley seeds. Subsequently, the Head part was replaced with DyHead. Through multi-scale network optimization of the detection head, the model's detection performance for small targets was improved. Finally, the loss function was optimized to accelerate the model's convergence speed, while also boosting the detection accuracy of small targets and the precision of bounding box regression. Based on the above improvements, this paper finally proposes the YOLOv8-SDW model for highland barley variety detection, where S, D, and W respectively denote SimAM, DyHead, and WIoU, as shown in Figure 2.

2.2.2 SimAM Attention Mechanism

The attention mechanism dynamically allocates computational resources, enabling the model to selectively focus on important regions and features within input data. It effectively filters irrelevant information and enhances the model's capability to extract key features. Due to the color similarity among different highland barley varieties, and the significant appearance variations under different lighting conditions, introducing an attention mechanism allows the model to deeply focus on the high-level semantic features of seeds. This improves the generalization ability and robustness of the model, and further boosts the accuracy of highland barley variety detection.

SimAM (Simple Attention Module) is a parameter-free attention mechanism proposed by Yang et al. [13]. Unlike conventional attention modules, SimAM evaluates the importance of neurons from both spatial and channel dimensions simultaneously by constructing an energy function derived from neuroscience theory, thereby achieving 3D attention weighting without introducing additional learnable parameters.

Figure 3 shows the comparison of attention mechanisms with different dimensional weights. One-dimensional attention mechanisms mainly model channel relationships, with SE [14] (Squeeze-and-Excitation) as a representative example. Two-dimensional attention mechanisms further incorporate spatial attention, among which CBAM [15] (Convolutional Block Attention Module) is a typical approach. However, these methods usually model channel or spatial information separately, which may limit the interaction between different feature dimensions. In contrast, SimAM adopts a 3D attention strategy to jointly model channel and spatial information, enabling more comprehensive feature representation and more balanced feature weighting.

SimAM evaluates the importance of each neuron by constructing an energy function and generates attention weights accordingly [13]. Assume that the input feature map is $X \in R^{C\times H\times W}$, where C, H, and W denote the number of channels, height, and width, respectively. For each neuron t in the feature map, the energy function is defined as Equation (1):

$$e_t^* = \frac{4(\hat{\sigma}^2 + \lambda)}{(t-\hat{\mu})^2 + 2\hat{\sigma}^2 + 2\lambda} \tag{1}$$

where $\hat{\mu}$ and $\hat{\sigma}^2$ denote the mean and variance of neurons within the same channel, respectively, and λ is a small regularization constant used to ensure numerical stability.

The attention weight is then obtained through the sigmoid activation function, as shown in Equation (2):

$$A_t = Sigmoid\left(\frac{1}{e_t^*}\right) \tag{2}$$

where A_t represents the attention weight assigned to neuron t.

Finally, the attention weights are multiplied element-wise with the input feature map to obtain the enhanced feature map Y, as expressed in Equation (3):

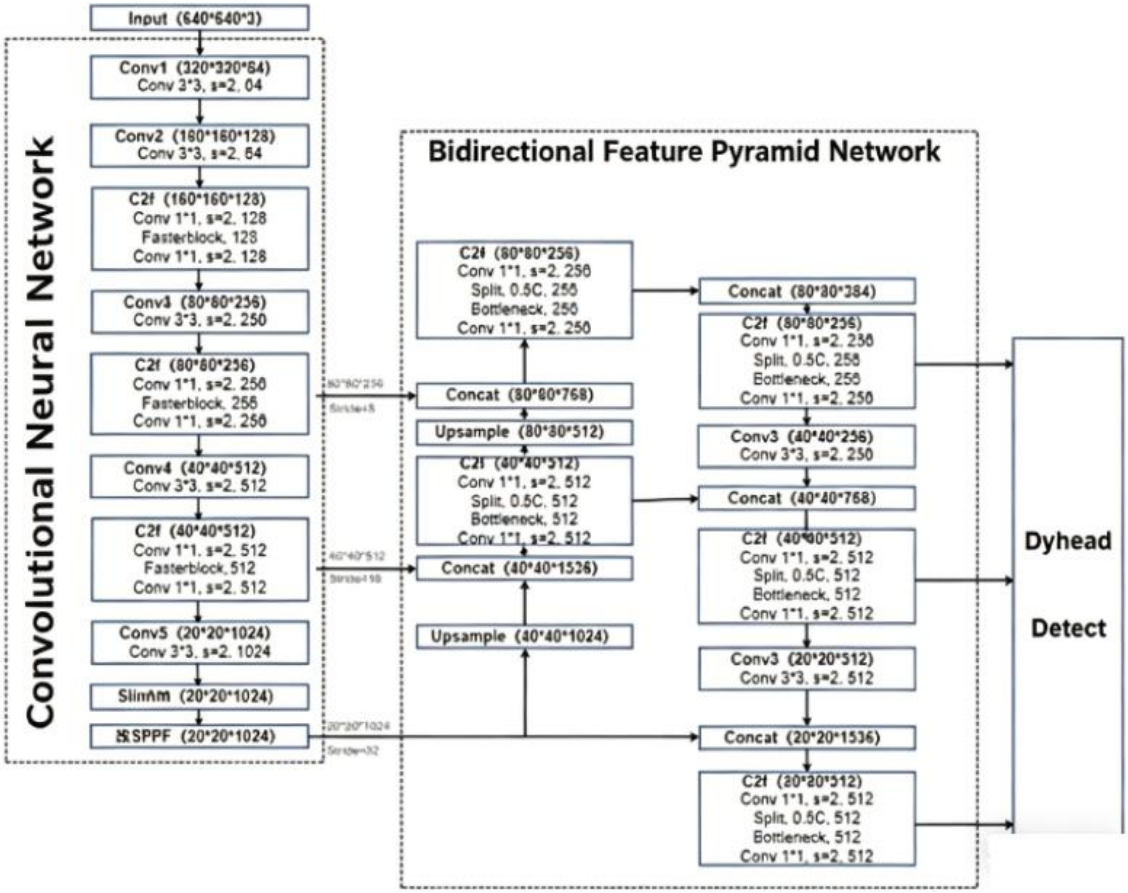
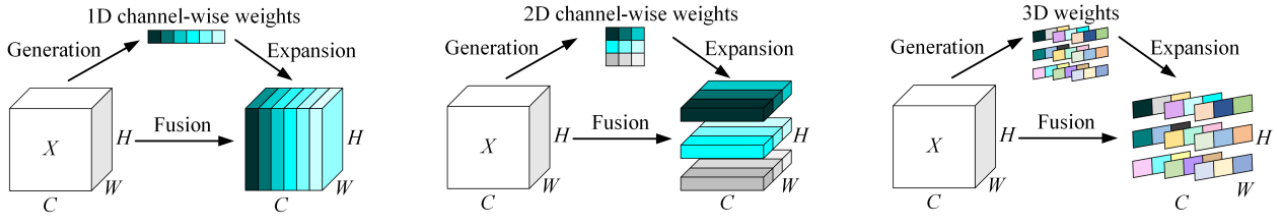


Fig. 2 YOLOv8-SDW architecture diagram.



(a) 1D Attention Mechanism (b) 2D Attention Mechanism (c) 3D Attention Mechanism

Fig. 3 Attention mechanisms in different dimensions.

$$Y = X \odot A \tag{3}$$

where $\odot$ denotes element-wise multiplication. In this way, SimAM can adaptively assign attention weights across both channel and spatial dimensions, enhancing discriminative features while suppressing redundant information.

2.2.3 Dynamic Detection Head Based on Attention Mechanism

To address the problems of small individual sizes and limited feature information in some highland barley samples, as well as the potential feature loss caused by network down sampling, this paper introduces DyHead [16], a dynamic detection head based on the attention mechanism. By integrating dynamic convolution and attention mechanisms, this module adaptively adjusts the response of the detection head to features at different scales and positions, thereby significantly enhancing the model's feature extraction capability and detection performance.

The implementation of DyHead mainly consists of

three key components: the scale-aware module, the spatial-aware module, and the task-aware module [17]. The scale-aware module is responsible for modeling feature interactions across different feature pyramid levels, thereby enhancing multi-scale feature fusion capability. The spatial-aware module dynamically focuses on important spatial regions to improve localization performance, while the task-aware module adaptively allocates feature responses according to different task requirements, such as classification and localization [18].

Finally, DyHead adaptively fuses the scale-aware feature F_s, spatial-aware feature F_{sp}, and task-aware feature F_{tp} to generate the final output feature F_{out}. The feature aggregation process can be generally expressed as Equation (4):

$$F_{out} = f(F_s, F_{sp}, F_{tp}) \tag{4}$$

where $f(\cdot)$ denotes the adaptive attention-based fusion function in DyHead.

The above three distinct attention mechanism modules can be unified into an efficient attention-learning dynamic detection head. The original DyHead structure was adopted in this study without modification to its core attention fusion mechanism. The overall structure of DyHead is shown in Figure 4.

2.2.4 WIoU Loss Function

As a core indicator in model training, the loss function is mainly used to quantify the deviation between predicted results and ground-truth labels, and serves as the optimization objective to guide the convergence of the algorithm.

Traditional IoU-based loss functions mainly evaluate the overlap ratio between predicted bounding boxes and ground-truth boxes. However, they are often sensitive to low-quality samples and may generate unstable gradients during training, especially in small-object detection tasks. To address these limitations, this study adopts WIoU [19] (Wise Intersection over Union) as the bounding box regression loss function. WIoU introduces a dynamic focusing mechanism on the basis of traditional IoU and evaluates anchor box quality using outlier degree, thereby adaptively adjusting the contribution of different samples during training. Compared with conventional IoU-based losses, WIoU can effectively suppress harmful gradients from low-quality samples while emphasizing ordinary-quality anchor boxes, which contributes to more stable bounding box regression and better localization performance [20].

The WIoU loss function is defined as Equation (5), which effectively balances the optimization contribution of samples with different localization qualities and improves the overall performance of the model.

$$\mathcal{L}_{\mathrm{WIoU}} = \mathcal{R}_{\mathrm{WIoU}}\mathcal{L}_{\mathrm{IoU}} \tag{5}$$

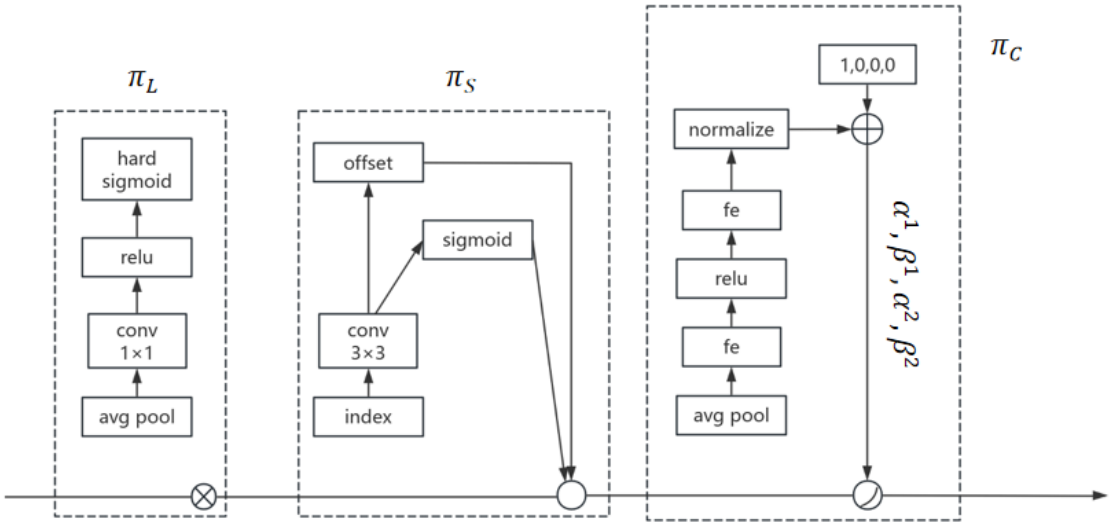


Fig. 4 DyHead structure diagram.

where $\mathcal{L}_{WIoU}$ denotes the WIoU bounding box regression loss, $\mathcal{L}_{IoU}$ denotes the IoU loss, and $\mathcal{R}_{WIoU}$ denotes the dynamic reweighting term defined in the original WIoU formulation, which is used to adaptively adjust the gradient contribution of samples with different localization qualities.

By dynamically reallocating optimization focus during training, WIoU reduces the influence of low-quality samples and improves the robustness of bounding box regression. This is particularly beneficial for highland barley seed detection, where target objects are relatively small and inter-class appearance differences are subtle.

2.3 Experimental Environment and Parameters

In the experiments, the input resolution of highland barley seed images was uniformly resized to 640 × 640 pixels. Model training and testing were implemented based on the PyTorch deep learning framework under Python 3.9. The experiments were conducted on a workstation equipped with an NVIDIA RTX 3090 GPU with 24 GB memory. The training epochs were set to 300, and the batch size was set to 16, the SGD optimizer was adopted, with an initial learning rate of 0.01, momentum of 0.937, and weight decay of 0.0005. In addition, a cosine annealing learning rate scheduling strategy was employed during training to dynamically adjust the learning rate and improve model convergence stability.

3. Results and Discussion

3.1 Ablation Experiments

3.1.1 Attention Mechanism Improvement Experiment

Table 2 shows the detection performance of four models: YOLOv8, YOLOv8-SE, YOLOv8-CBAM and YOLOv8-SimAM. As can be seen from the table, compared with the original YOLOv8 model, the YOLOv8-SimAM algorithm improves precision by 2.3%, by 1.2% compared with YOLOv8-SE, and by 0.3% compared with YOLOv8-CBAM. In terms of recall, YOLOv8-SimAM increases by 1% over the original YOLOv8, by 2.2% over YOLOv8-SE, and is equal to that of YOLOv8-CBAM. In terms of mean average precision (mAP), YOLOv8-SimAM is improved by 2.7% compared with YOLOv8, by 0.7% compared with YOLOv8-SE, and by 0.2% compared with YOLOv8-CBAM. Overall, the detection capability of all models is enhanced after introducing an attention mechanism. However, different attention modules differ significantly in the degree of increased computational complexity. The SE module leads to an increase of approximately 0.1 GFLOPs, the CBAM module increases by 0.5 GFLOPs, and the SimAM module increases by 0.4 GFLOPs.

In terms of model performance and computational efficiency, SimAM achieves higher detection accuracy with only a slight increase in computational cost compared with the original YOLOv8 model. Compared with CBAM, SimAM introduces lower computational overhead while maintaining competitive detection performance. Therefore, SimAM was selected as the attention mechanism for subsequent experiments.

3.1.2 Dynamic Detection Head Improvement Experiment

Table 3 compares the detection performance of the YOLOv8-SimAM model and the YOLOv8-SimAM-DyHead model. After introducing the DyHead structure, the improved model achieves increases of 3.7 percentage points, 1 percentage point, and 1.9 percentage points in precision, recall, and mAP, respectively. These results indicate that DyHead effectively enhances multi-scale feature representation and further improves overall detection performance on the highland barley dataset.

3.1.3 Loss Function Improvement Experiment

Table 4 presents the test results of the YOLOv8-SimAM-DyHead model and the YOLOv8-SDW model. According to the experimental evaluation metrics, the YOLOv8-SDW model optimized with the WIoU loss function achieves a significant performance improvement compared with the YOLOv8-SimAM-DyHead baseline model.

Table 2 Comparison results of different attention mechanisms.

Model	Precision	Recall	mAP @0.5	Params	FLOPs
YOLOv8	83%	88%	89%	3.2	8.7
YOLOv8-SE	84.1%	86.8%	91%	3.5	8.8
YOLOv8-CBAM	85%	89%	91.5%	3.3	9.2
YOLOv8-SimAM	85.3%	89%	91.7%	3.2	9.1

Table 3 Comparison results of different detection heads.

Model	Precision	Recall	mAP @0.5
YOLOv8-SimAM	85.3%	89%	91.7%
YOLOv8-SimAM-DyHead	89%	90%	93.6%

Table 4 Comparison result.

Model	Precision	Recall	mAP @0.5
YOLOv8-SimAM-DyHead	89%	90%	93.6%
YOLOv8-SDW	91%	90%	95.1%

Specifically, the optimized model is enhanced by 2 percentage points and 1.5 percentage points in the two key indicators of detection precision and mAP, respectively, while the recall rate remains unchanged. Although the recall rate does not change, the simultaneous improvement in precision and mAP fully demonstrates that the improved model has achieved a substantial enhancement in overall detection capability. These experimental results indicate that the adopted loss function optimization strategy effectively strengthens the comprehensive detection performance of the model, especially in improving accuracy while maintaining the original recall ability.

3.2 Comparative Experiments

To verify the superiority of the network proposed in this chapter for highland barley variety detection, the proposed algorithm is compared and analyzed in detail with the two-stage object detection algorithm Faster R-CNN [21] and the one-stage object detection algorithms YOLOv5n and YOLOv8n. All control experiments are conducted under a unified software and hardware platform, and the division of the training set and validation set remains unchanged to ensure the reliability and comparability of the experimental results.

As shown in Figure 5, the iteration curves of mean average precision (mAP) of the four models on the validation set are presented. It can be observed from the data that compared with the other three models, the algorithm proposed in this paper has a significantly faster convergence speed on the mAP curve, and its mean precision obviously exceeds that of the YOLOv5n and YOLOv8n models. However, in the later stage of training, the final stable mAP of the Faster R-CNN model slightly surpasses that of the proposed algorithm.

The two-stage object detection algorithm Faster R-CNN places particular emphasis on accuracy and stability in its design [22]. By integrating techniques such as a Region Proposal Network (RPN), precise position adjustment, and multi-scale feature extraction, Faster R-CNN tends to achieve higher average precision in object detection tasks [23]. Nevertheless, this performance-oriented design also incurs high computational overhead, resulting in low model inference efficiency, which makes it difficult to meet the requirements of application scenarios with high real-time demands. Through comprehensive comparative analysis, the proposed algorithm not only satisfies the real-time detection requirements of highland barley varieties, but also maintains competitive detection accuracy. Therefore, it achieves a favorable balance between detection accuracy and inference efficiency.

Although its mean average precision is slightly lower than that of Faster R-CNN, it exhibits superior detection speed and real-time performance, which can better satisfy the actual requirements of variety detection in highland barley seed classification.

Table 5 shows the comparison results between the two-stage object detection algorithms, one-stage object detection algorithms, and the proposed algorithm. As can be seen from the results, the recall of the proposed algorithm matches that of Faster R-CNN and is significantly higher than the other two one-stage object detection methods. The precision of the proposed algorithm reaches 91%, and the mean average precision is 95.1%, both of which exceed those of YOLOv5n and YOLOv8n but are slightly lower than those of Faster R-CNN. The proposed algorithm achieves a frame rate of 89.0 FPS, which is higher than that of Faster R-CNN and YOLOv5n but slightly lower than that of YOLOv8n. Through comprehensive comparative analysis, the algorithm proposed in this chapter not only meets the real-time detection requirements of highland barley varieties but also performs excellently in detection accuracy, achieving a favorable balance between detection accuracy and inference efficiency.

3.3 Ablation Experiment

To verify the comprehensive effects of the three improvement strategies based on the YOLOv8n framework on the highland barley seed dataset, an ablation experiment was designed and conducted in this study. All experiments were performed on a unified software and hardware platform, and the division of the training set and test set remained unchanged to ensure the reliability and comparability of the experimental results. Figure 6 shows the training results of the four models. From the mean average precision (mAP) curves on the validation set, it can be observed that the original YOLOv8 model achieves the lowest mAP, while the optimized YOLOv8-SDW model achieves the highest mAP. This result confirms the significant effectiveness of the proposed algorithm optimization scheme.

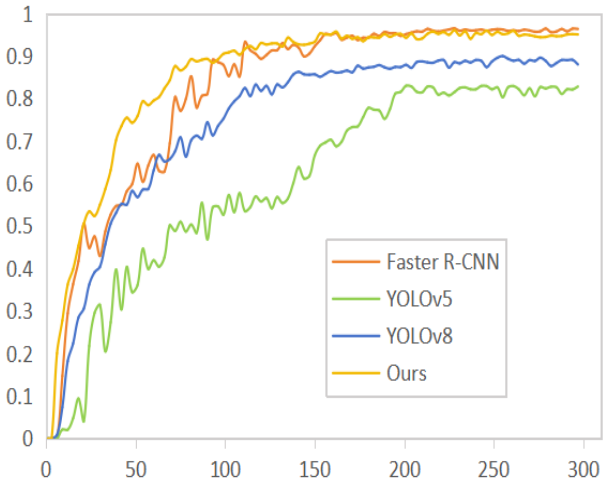


Fig. 5 Comparison of mean average precision curves for different model versions.

Table 5 Comparison of results of different algorithms.

Model	Precision	Recall	mAP @0.5	FPS
Faster R-CNN	91.5%	90%	96.4%	29.1
YOLOv5n	79%	80%	83%	68.6
YOLOv8n	83%	88%	89%	91.2
YOLOv8-SDW	91%	90%	95.1%	89.0

Table 6 Ablation experiment.

No.	SimAM	DyHead	WIoU	Precision	Recall	mAP @0.5	Size MB^{-1}	FPS
1	×	×	×	83%	88%	89%	5.95	91.2
2	√	×	×	85.3%	89%	91.7%	5.96	91.0
3	√	√	×	89%	90%	93.6%	6.89	89.0
4	√	√	√	91%	90%	95.1%	6.89	89.0

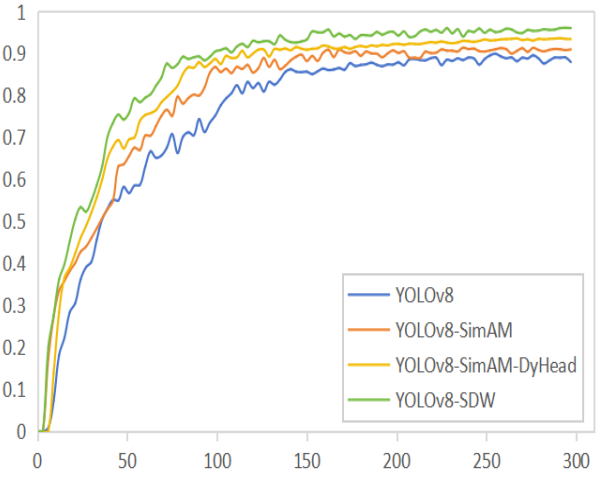


Fig. 6 Comparison of mean average precision curves for different improved YOLOv8 versions.

This study systematically evaluated the performance of the four models on key performance indicators, and the ablation results are shown in Table 6. The experimental data show that after introducing the SimAM attention mechanism, the model performance is significantly improved while maintaining stable parameters (+0.01 MB) and inference speed (-0.2 FPS). Specifically, the mean average precision (mAP) is increased by 2.7 percentage points to 91.7%. This result verifies the effectiveness of the SimAM module in enhancing the extraction of high-level semantic features.

4. Conclusion and Future Perspective

This study focuses on the field of intelligent crop seed detection and conducts systematic research on the variety identification of highland barley, a characteristic crop on the Qinghai-Tibet Plateau. Based on the theoretical foundation of the YOLO series algorithms and drawing on research achievements in other crop seed detection fields, an improved YOLOv8 model integrated with the SimAM attention mechanism is proposed. First, a specialized dataset is constructed according to the morphological characteristics of highland barley seeds, and then an improved YOLOv8n object detection model is developed. Through innovative methods such as introducing an attention mechanism, optimizing the network structure, and improving the loss function, the model achieves a mean average precision of 95.1% in highland barley seed detection tasks while maintaining a real-time processing speed of 89 FPS.

Although this study has achieved promising results in highland barley seed variety detection, the proposed method is still at the laboratory validation stage and cannot yet be directly applied to large-scale practical seed sorting scenarios. The highland barley dataset used in this study was independently collected on an experimental platform, with a relatively limited sample size, controlled imaging conditions, and a pure background setting, which differ from real-world environments involving varying illumination, occlusion, seed overlap, and complex backgrounds. These factors may affect the robustness and generalization ability of the model in practical deployment.

In future work, highland barley seed image data will be collected directly from seed sorting production lines and real agricultural environments to construct a larger-scale and more representative dataset. In addition, although the YOLOv8-SDW model performs well on the current six highland barley varieties, practical applications involve more diverse varieties and more complicated seed conditions. Therefore, future studies will further expand variety categories, optimize model robustness under complex scenarios, and improve the practical applicability of the proposed method in intelligent seed sorting systems.

Funding Statement

This work was supported by the Tianjin-Gansu Collaboration Key Research and Development Project of Gansu Province (Grant No. 24CXNP003). The funding body had no involvement in the design of the study, collection, analysis, and interpretation of data, writing of the manuscript, or the decision to publish the results.

References

[1] Xu, P., and Zhang, Z. B. 2023. "Industrialization of functional nutritious color wheat." *China Rural Science & Technology* 332 (01).

[2] Xia, X. Y., Li, F., and Xie, K. et al. 2024. "Research progress on non-destructive detection technology of crop seed vigor." *Transactions of Chinese Society of Agricultural Engineering* 43 (10): 64-73.

[3] Ding, Z. Y., Yue, X. J., and Zeng, F. G. et al. 2023. "Spectral detection of maize seed vigor based on machine learning and deep learning." *Journal of Huazhong Agricultural University* 42 (3): 230-240.

[4] Bai, W. W., Zhao, X. N., and Luo, B. et al. 2023. "Research on wheat seed germination detection method based on YOLOv5." *Acta Agriculturae Zhejiangensis* 35 (2): 445-454.

[5] Mukasa, C., Wakholi, M. A., and Faqeerzada et al. 2022. "Nondestructive discrimination of seedless from seeded watermelon seeds by using multivariate and deep learning image analysis." *Computers and Electronics in Agriculture*, article id. 106799.

[6] Wang, X., Dong, Q., and Yang, G. 2023. "YOLOv5 improved by optimized CBAM for crop pest identification." *Comput. Syst. Appl.* 32 (7): 261-268.

[7] Wang, J., Li, Q., and Fang, Z. et al. 2023. "YOLOv6-ESG: A lightweight seafood detection method." *Journal of Marine Science and Engineering* 11 (8): 1623.

[8] Liu, Q. H., Yang, X. Y., and Jie, H. et al. 2023. "Rice grain detection based on YOLO v7 fusing of GhostNetV2." *Transactions of the Chinese Society for Agricultural Machinery* 54 (12): 253-260, 299.

[9] Zhu, D., Wen, R., and Xiong, J. 2023. "Lightweight corn silk detection network incorporating with coordinate attention mechanism." *Trans. CSAE* 39: 145-153.

[10] Dong, W., and Liang, X. Y. 2024. "An algorithmic model for recognizing healthy wheat seeds based on YOLOv8." *Third International Conference on Electronic Information Engineering, Big Data, and Computer Technology (EIBDCT 2024)*, Vol. 13181. SPIE, 2024.

[11] Wang, J. P., Meng, H., and Zhen, Q. G. et al. 2024. "Camellia oleifera fruit static and dynamic detection counting based on improved COF-YOLO v8n." *Transactions of the Chinese Society for Agricultural Machinery* 55 (4): 193-203.

[12] Chen, C. Q., Fan, Y. C., and Wang, L. 2022. "Logo detection based on improved Mosaic data augmentation and feature fusion." *Computer Measurement & Control* 30 (10): 188-194, 201.

[13] Yang, L. X. et al. 2021. "Simam: A simple, parameter-free attention module for convolutional neural networks." *International Conference on Machine Learning.*

[14] Hu, J., Li, S., and Gang, S. 2018. "Squeeze-and-excitation networks." Proceedings of *the IEEE Conference on Computer Vision and Pattern Recognition*.

[15] Woo, S. et al. 2018. "CBAM: Convolutional block attention module." Proceedings of *the European Conference on Computer Vision (ECCV)*.

[16] Dai, X., Chen, Y., and Xiao, B. et al. 2021. "Dynamic head: Unifying object detection heads with attentions." *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 7373-7382.

[17] Yan, X., and Liu, F. 2023. "Dyhead-yolov5 based on improved object detection heads with attentions." *International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD). IEEE*, pp. 1-6.

[18] Jiang, Z., Han, Y., and Cheng, Y. et al. 2025. "An improved YOLOv8-dyhead-WiseIoU model for positioning and counting detection of grouting sleeves in a prefabricated wall." *Journal of Construction Engineering and Management* 151 (4): 04025016.

[19] Tong, Z., Chen, Y., and Xu, Z. et al. 2023. "Wise-IoU: bounding box regression loss with dynamic focusing mechanism." *arXiv preprint arXiv*: 2301.10051.

[20] Hu, D., Yu, M., and Wu, X. et al. 2024. "DGW-YOLOv8: A small insulator target detection algorithm based on deformable attention backbone and WIoU loss function." *IET image processing* 18 (4): 1096-1108.

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

[22] Zhao, Y. Q., Rao, Y., and Dong, S. P. et al. 2020. "A review of deep learning object detection methods." *Journal of Image and Graphics* 25 (4): 629-654.

[23] Xu, D. G., Wang, L., and Li, F. 2021. "A review of typical deep learning-based object detection algorithms." *Computer Engineering and Applications* 57 (8): 10-25.