Data segmentation method based on seedling density to improve the accuracy of rapeseed seedling recognition

Qi Wang, Liqing Chen, Quan Zheng, Lichao Liu*

(*School of Mechanical and Vehicle Engineering, Anhui Agricultural University, Hefei 230036, China*)

Abstract: The accurate recognition of rapeseed seedlings is the premise of achieving accurate counting. Traditional target detection algorithms and current improved algorithms cannot balance the counting accuracy and seedling recognition at different densities. Thus, this study proposed improved YOLOv7 recognition models by automatically dividing dense and non-dense datasets. For the non-dense dataset, the attention mechanism CA is introduced, and the loss function DIoU is substituted; for the dense dataset, the attention mechanism CBAM is introduced, and the loss function GIoU is substituted, so that the network focuses on the target object to improve the accuracy of the model. The results show that the improved accuracy of the non-dense area reaches 97.78%, mAP@0.5 reaches 98.32%, and mAP@0.5:0.95 reaches 66.86%. The accuracy of the improved dense area reaches 97.85%, mAP@0.5 reaches 95.91%, and mAP@0.5:0.95 reaches 69.9%. In addition, this paper compared SSD, YOLOv5, YOLOv7, YOLOv7-tiny, and the improved YOLOv7 and validates the feasibility of the improved YOLOv7 model. The improved YOLOv7 model proposed in this paper can provide a useful reference for future research directions of seedling recognition.

Keywords: rapeseed seedling, dense area, YOLOv7, object detection, counting

DOI: 10.25165/j.ijabe.20261904.9572

Citation: Wang Q, Chen L Q, Zheng Q, Liu L C. Data segmentation method based on seedling density to improve the accuracy of rapeseed seedling recognition. Int J Agric & Biol Eng, 2026; 19(4): 282–291.

1 Introduction

As the main oilseed crop in China, rape plays a vital role in China's edible oil market[1-3]. Anhui Province is a large oilseed rape production province, where this crop has a long planting history and is widely distributed. It is one of China's most important production areas of winter oilseed rape. Thus, improving the yield of oilseed rape is of great significance for China's oilseed rape industry development and edible oil safety[4]. Seedling density is an important indicator for assessing crop yield, but high-density planting can lead to competition among seedlings, which can affect their yield. The most common method to quantify plant density is to recognize and count the seedlings after emergence[5]. Traditional rape recognition usually relies on manual operation, which is not only inaccurate, but also inefficient and limited[6,7], which affects subsequent agricultural production management. Therefore, it is very important to take appropriate means to improve the accuracy of plant recognition and counting.

In recent years, precision agriculture technology has developed rapidly[8]. The combination of agriculture and computer vision is becoming more and more common, and deep learning is widely used in crop recognition and counting[9]. However, when the background is complex or dense, it easily causes false detection and missing detection, thus reducing the counting accuracy[10]. YOLOv7 is more advanced, more accurate, and faster than other versions of the model. Therefore, in this paper, YOLOv7 is selected as the base network model and improved to increase the accuracy and efficiency of counting and provide better support for agricultural production[11]. In the field of deep learning, Ma et al.[12] used the improved YOLOv5s lightweight to identify plants; the model memory occupied 8.9 MB, with an accuracy of 95.1%. Zhang et al.[13] proposed a YOLO-VOLO-LS recognition method, which obtained excellent results of 93.452%, 96.059%, 96.014%, and 96.039% in Val-acc, Test-acc, Recall, and Precision, respectively. Zhu et al.[14] identified duck eggs in a complex environment based on the YOLOv7 model, and the F_1 reached 95.5%. Zhao et al.[15] proposed a lightweight model, LW-YOLOv7, which can efficiently calculate the number of maize seedlings. Li et al.[16] proposed an improved YC-YOLOv7 model by incorporating optimizations such as depthwise separable convolution, CBAM attention mechanism, and WIoU loss function. The enhanced model achieved a mean average precision of 94.0%, a precision of 89.8%, and a recall rate of 91.2%. Li et al.[17] proposed a multi-object detection method based on FEPW R-CNN, which was validated on a rapeseed seedling dataset. The proposed method achieved 85.42% mAP@0.5, providing an efficient and high-precision solution for multi-object detection in complex scenarios. All the above studies have achieved good results, but the compatibility of different improved algorithms with recognizing objects is insufficient in complex backgrounds or high-density environments.

Therefore, aiming at this problem, this paper intends to study the target recognition and counting model of dense rapeseed seedlings and non-dense rapeseed seedlings, respectively. Using YOLOv7 as the base network model, for non-dense areas, the attention mechanism CA was introduced to effectively improve the feature extraction ability of rape at seedling stage, and the loss function DIoU was introduced to improve the overall performance of the detector. For dense areas, CBAM was introduced to improve

Received date: 2024-12-24 **Accepted date:** 2025-10-25

Biographies: Qi Wang, MS, research interest: crop phenotype detection and intelligent agricultural equipment, Email: 12513078@zju.edu.cn; **Liqing Chen**, PhD, Professor, research interest: intelligent agricultural machinery and equipment, modern vehicle design methods, Email: lqchen@ahau.edu.cn; **Quan Zheng**, Professor, research interest: design of tractor power chassis, resource utilization of agricultural and forestry residues, Email: zhengquan@ahau.edu.cn.

***Corresponding author: Lichao Liu**, PhD, Associate Professor, research interest: intelligent connected vehicle and agricultural machinery and equipment. No. 130, Changjiang West Road, Hefei 230036, China. Tel: +86-15156691929, Email: llchao@ahau.edu.cn.

the ability to identify dense areas, and the loss function GIoU was substituted. The target information of seedlings was enhanced, and the models were analyzed and compared.

2 Materials and methods

2.1 Overall scheme

Figure 1 is the overall flow chart of this paper, which mainly consists of dataset preparation, improved recognition model, and data integration. The dataset preparation mainly involves image acquisition, image preprocessing, image annotation, automatic segmentation of datasets, and data enhancement. The rapeseed seedling recognition model was improved on the YOLOv7 model, trained and verified on the dense dataset and the non-dense dataset respectively, and the detection results of rapeseed seedlings were obtained. Data integration was carried out to integrate the obtained non-dense detection results and dense detection results according to location information, and to count and compare the integrated results.

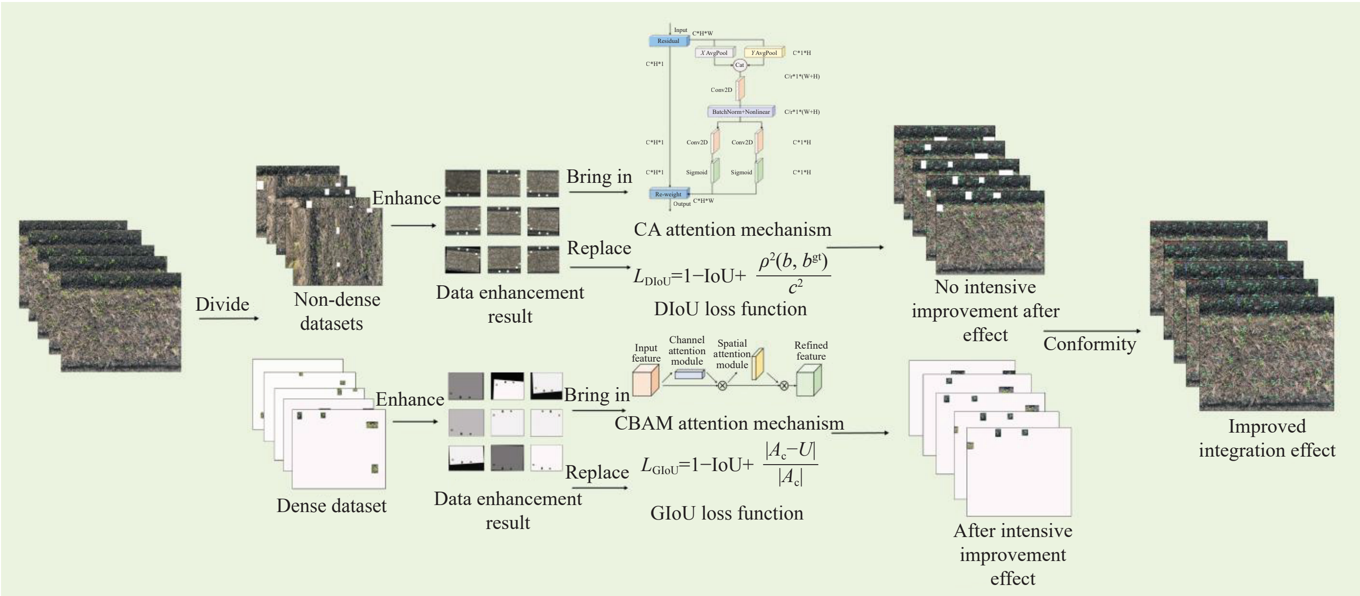


Figure 1 Overall flow chart

2.2 Dataset sample collection and production

In this study, images of rapeseed at seedling stage were collected from the rapeseed planting experimental field in Hanshan County, Maanshan City, Anhui Province. The rapeseed sowing method was broadcast sowing, the sowing time was October 20, 2022, and the collection time was December 6, 2022. The data were collected by DJI Yu2 UAV at a height of 1.5 m above the ground.

After emergence, for dense areas, there may be poor manual recognition effect and inaccurate counting, which needs to be solved by computer vision technology. To prepare the sample datasets, firstly, the dense area was manually marked by the annotation software Labelimg[18]. To achieve the automatic segmentation of dense and non-dense datasets, an automatic segmentation algorithm is designed: the manually annotated labels of dense regions are masked with white rectangles, and the remaining unmasked areas are considered as non-dense regions. Conversely, by applying this algorithm to fill all areas outside the dense region labels with white, the remaining unmasked portions then represent the dense regions. Then, the rapeseed seedlings in dense and non-dense areas were labeled, as shown in Figure 2.

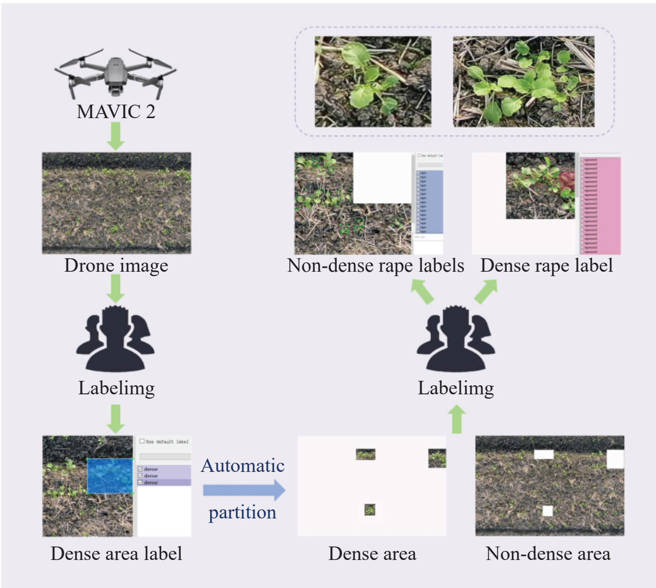


Figure 2 Dataset collection and production

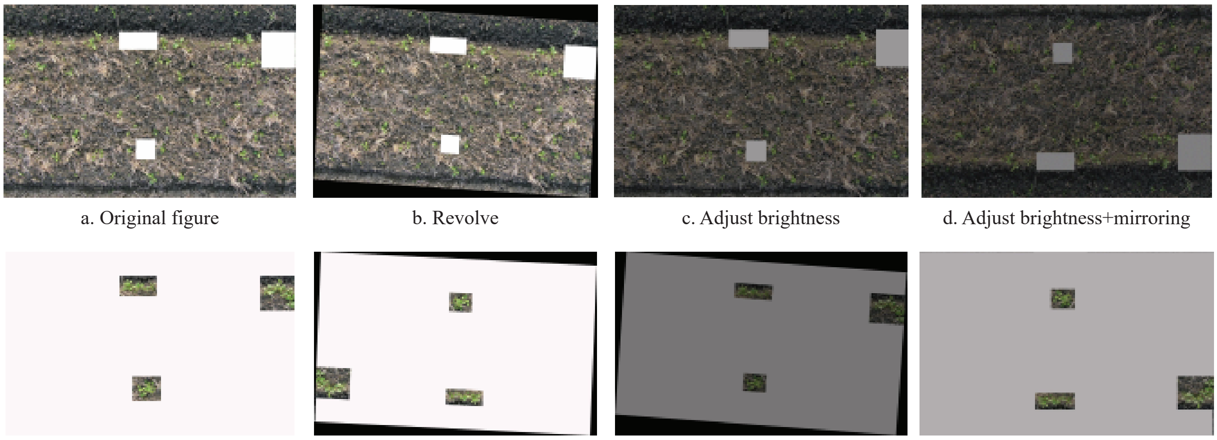
2.3 Data enhancement

Data enhancement is a common deep learning technique used to increase the diversity and amount of training data, thereby improving the robustness and generalization ability of the model[19]. The training samples selected in this paper are relatively small, which will affect the training results. Therefore, the data of rapeseed seedling stage undergo clipping, translation, brightness change, noise addition, angle rotation, and mirroring to enhance the image data[20] and thus increase data diversity. The data enhancement results are shown in Figure 3.

3 Methods of rape detection and recognition

3.1 YOLOv7 model

YOLO has been a favorite among researchers in the field of target detection. YOLOv7 was used in this paper. Compared with other YOLO models, YOLOv7 showed the best detection accuracy in the rapeseed seedling dataset. The network architecture of YOLOv7 consists of three parts: input, backbone, and head. The specific components are as follows[21]:



a. Original figure b. Revolve c. Adjust brightness d. Adjust brightness+mirroring

e. Original figure f. Image+rotation g. Adjust brightness+revolve h. Adjust brightness+mirroring

Figure 3 Data enhancement results

Input is the input link of rapeseed seedling stage image. The image of rapeseed seedling stage is pre-processed through screening and data enhancement.

Backbone layer includes several BConv layers, ELAN layers, and MP layers; its main function is to achieve channel doubling and feature extraction. The BConv, ELAN, and MP modules in the YOLOv7 backbone work together to extract features from input images for precise recognition.

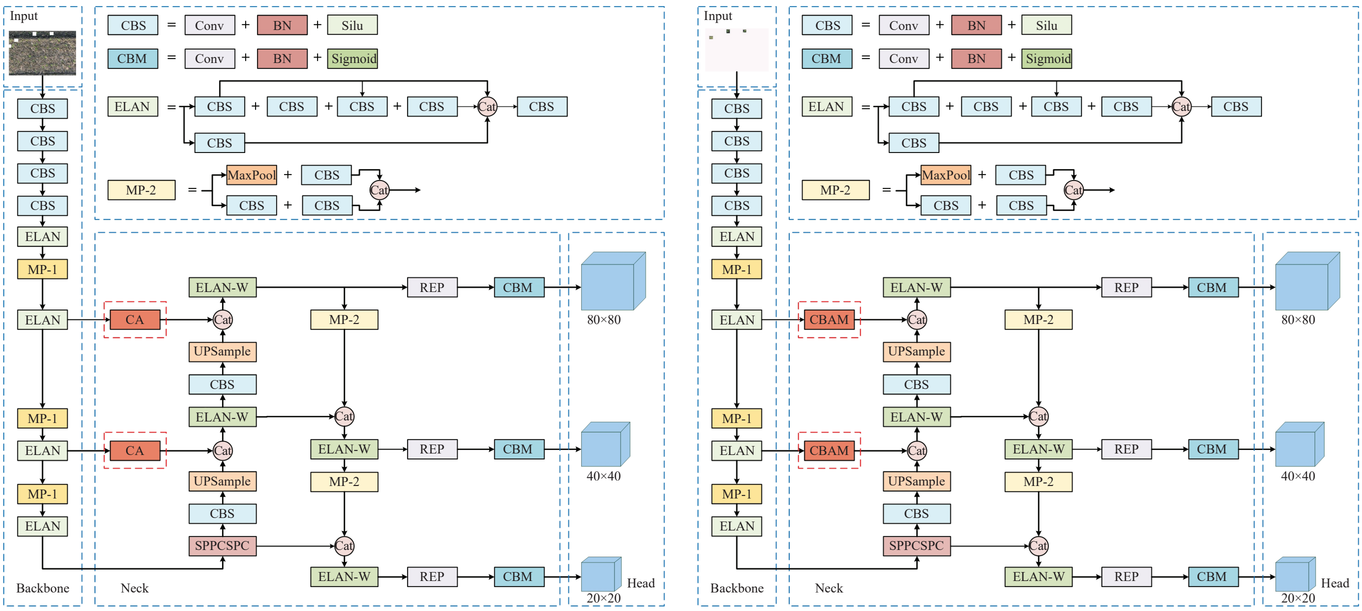
The head layer consists of the SPPCPC layer, several Conv layers, several MP layers, several Catconv layers, and the RepVGG block layer that outputs three heads[22], to realize the prediction of rapeseed seedling stage.

3.2 Improvement of recognition model of rapeseed seedling stage of YOLOv7

The research object of this paper is rape at the seedling stage. Due to the sowing method, some dense areas will be blocked at the seedling stage, which will affect the result. Therefore, in order to achieve an efficient and accurate recognition model for rapeseed seedlings, YOLOv7 was used as the basis model. For non-dense areas, CA attention mechanism was inserted into the backbone network and DIoU loss function was adopted to improve accuracy and detection effect. For dense areas, CBAM attention mechanism is inserted and GIoU loss function is used as a replacement to improve the precision of dense area recognition. The improved model structure is shown in Figure 4.

3.3 CA attention mechanism

Due to the complex background of rapeseed seedling images, the hidden layer of the backbone network of YOLOv7's network model is prone to be ignored by the network. The CA attention mechanism module can enhance this part of attention feature information to extract more comprehensive and rich feature information. Thus, the effective separation of the target area and the background in the image is realized. The CA structure diagram is shown in Figure 5.



a. Improve the network structure of YOLOv7 in non-dense areas b. Dense areas improve the network structure of YOLOv7

Figure 4 Improved network structure of YOLOv7

Input feature maps are globally average pooled in the direction of width and height respectively[23]; H and W are the height and width of the input feature map respectively, and the expression of the c channel with height h is as follows:

$$z_c^h(h) = \frac{1}{W}\sum_{0\le i<W} x_c(h,i) \tag{1}$$

where, $z_c^h(h)$ is output of the cth channel at height values of h; $x_c(h,i)$ is the eigenvector of the first row.

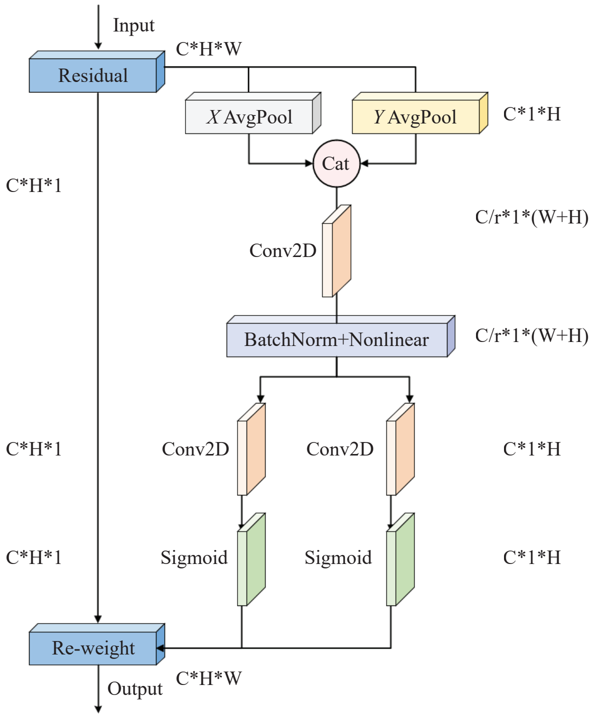


Figure 5 Structure of CA attention mechanism

Similarly, the expression for the c channel with width w is as follows:

$$z_c^w(w)=\frac{1}{H}\sum_{0\le j<H}x_c(j,w) \tag{2}$$

where, $z_c^w(w)$ is the output of the c channel in height; $x_c(j,w)$ is the eigenvector in the first row.

The feature maps in width and height directions are obtained respectively, and the features in the two spatial directions are aggregated through one-dimensional global pooling operation, so that the attention module can capture the correlation along one spatial direction and retain the position information along the other spatial direction, thus helping the network to capture the target of interest more accurately[24]; then 1×1 convolution is used to perform the transformation operation. The expression is as follows:

$$f=\delta(F_1([z^h,z^w])) \tag{3}$$

where, f is an intermediate feature map that encodes spatial information in horizontal and vertical directions; δ is a nonlinear activation function; F_1 is 1×1 convolution transformation; z^h, z^w is coded output.

Then f is divided into two separate tensors, and 1×1 convolution is used respectively to transform them to have the same number of channels as X. Finally, the attention weight matrix is obtained by Sigmoid activation function. The expressions are as follows:

$$\begin{aligned} g^h&=\sigma(F_h(f^h))\\ g^w&=\sigma(F_w(f^w)) \end{aligned} \tag{4}$$

where, g^h, g^w are characteristic tensors along the horizontal and vertical directions; F_h, F_w are 1×1 convolution transformation; σ is Sigmoid activation function.

The expression for the output through the coordinate attention module is as follows:

$$y_c(i,j)=X_c(i,j)\times g_c^h(i)\times g_c^w(j) \tag{5}$$

where, $y_c(i,j)$ is the final output; $X_c(i,j)$ is the input feature map; $g_c^h(i)$, $g_c^w(j)$ is the weight of attention in the horizontal and vertical directions.

3.4 CBAM attention mechanism

CBAM attention mechanism is composed of channel attention mechanism and spatial attention mechanism. Starting from the scope of channel and spatial, CBAM introduces spatial attention and channel attention to realize the sequential attention structure from channel to space[25]. Spatial attention can make the neural network pay more attention to the pixel regions that play a decisive role in the classification of the image and ignore the unimportant regions. Channel attention is used to deal with the distribution relationship of the channel in the feature map. At the same time, the attention allocation of the two dimensions enhances the performance improvement of the attention mechanism[26]. The CBAM structure diagram is shown in Figure 6.

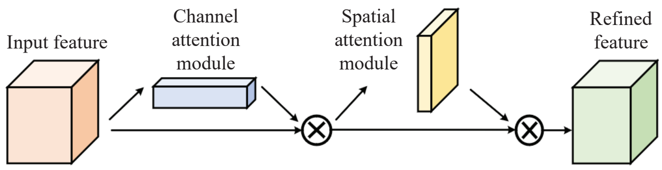


Figure 6 CBAM attention mechanism structure

The input feature graph F is globally averaged and globally maximally pooled to obtain one-dimensional feature vectors respectively. The two feature vectors pass through an MLP with shared weights, then the weights are added, and finally the Sigmoid activation function is scored to get the channel attention mechanism. The expression is as follows:

$$\begin{aligned} M_c(F)=&\sigma(\mathrm{MLP}(\mathrm{AvgPool}(F))+\mathrm{MLP}(\mathrm{MaxPool}(F)))=\\ &\sigma(W_1(W_0(F_{\mathrm{avg}}^c))W_1(W_0(F_{\max}^c))) \end{aligned} \tag{6}$$

where, $M_c(F)$ is channel attention vector; σ is Sigmoid activation function.

The input feature map F' is globally average pooled and globally maximum pooled in channel dimension to obtain the feature map with the same size as F, and the two feature maps are combined. Then a 7×7 convolution kernel is used for the convolution operation, and finally the spatial attention vector M_s is obtained through the Sigmoid activation function. The expression is as follows:

$$\begin{aligned} M_s(F)=&\sigma(f^{7\times7}([\mathrm{AvgPool}(F);\mathrm{MaxPool}(F)]))=\\ &\sigma(f^{7\times7}([F_{\mathrm{avg}}^s;F_{\max}^s])) \end{aligned} \tag{7}$$

where, $M_s(F)$ is a spatial attention vector; σ is Sigmoid activation function.

Finally, F' is multiplied by the attention vector M_s to obtain the final feature graph. The expression is as follows:

$$\text{Refine Feature}=F'\times M_s \tag{8}$$

where, Refine Feature is the final feature map; M_s is a spatial attention vector.

3.5 DIoU loss function

YOLOv7 adopts the CIoU loss function calculation method, but the CIoU loss function is largely dependent on the aggregated bounding box regression indicator and does not take into account the desired mismatch direction between the truth box and the prediction box. DIoU loss function adds the detection frame, including both the real frame and the prediction frame, and takes into account the distance, overlap rate, and scale between the target and anchor, so that the target frame regression becomes more stable and does not have problems such as divergence in the training process like IoU. DIoU loss function was calculated as follows:

$$L_{\mathrm{DIoU}} = 1 - \mathrm{IoU} + \frac{\rho^2(b, b^{gt})}{c^2} \tag{9}$$

where, c is the diagonal length of the smallest enclosing box that covers both boxes; $\rho^2(b, b^{gt})$ is the distance between the centers of the two boxes.

3.6 GIoU loss function

The CIoU loss function demonstrates certain limitations in handling non-overlapping bounding boxes, as it fails to provide effective optimization when the predicted box and target box do not intersect. Moreover, when the dimensions of both boxes remain constant, the CIoU loss value remains unchanged for any configuration that yields the same intersection value, meaning it cannot distinguish between different spatial relationships that produce identical intersection areas. To overcome these drawbacks, GIoU introduces the concept of the minimum enclosing rectangle that contains both the predicted and ground truth boxes. This approach preserves the essential properties of IoU while effectively addressing the identified shortcomings of CIoU. The calculation of GIoU is defined as follows:

$$\mathrm{GIoU} = \mathrm{IoU} - \frac{A_c - U}{A_c} \tag{10}$$

where, A_c is the area of two rectangles; U is a union of two rectangular boxes.

Then the expression of the GIoU loss function is as follows:

$$L_{\mathrm{GIoU}} = 1 - \mathrm{GIoU} \tag{11}$$

3.7 Model training and evaluation index

3.7.1 Training environment details

This test environment is based on the deep learning framework built by Windows 10, Python3.9, Pytorch, and Cuda[27], and the graphics card of the test hardware environment is NVIDIA GeForce RTX 3090. The number of training iterations is 300, the training batch size is 32, the learning rate is 1e-6, and the optimizer is SGD.

3.7.2 Evaluation index

In order to verify the feasibility of the rape seedlings model, this paper selected accuracy P (Precision), Recall R (Recall), and mAP (mean Average Precision) as evaluation metrics used in the evaluation of the rapeseed seedling models[28]. The specific definitions of these evaluation metrics are as follows:

$$P = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{12}$$

$$R = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{13}$$

If an object is correctly predicted and the IoU threshold exceeds the value we set, then the object is considered a positive sample. Otherwise the other objects are considered a negative sample, where TP (True Positives) represent the number of correctly classified rapeseed seedlings where both the detection result and the ground truth are rapeseed seedlings; FP (False Positives) represent the number of incorrectly classified where the detection result is rapeseed seedling and the ground truth is background; FN (False Negatives) represent the number of incorrectly classified where the detection result is the background and the ground truth is rapeseed seedling.

AP represents the area enclosed by precision and recall rates. The specific calculation method is shown in Equation (14); P(R)dR is a function with recall rate as the horizontal coordinate and precision as the vertical coordinate. In Equation (15), mAP refers to the average value of AP in all categories, where n represents the number of categories in the dataset. For the purposes of this article, there is only one category to be detected, so mAP can be regarded as AP. It should be noted that mAP (mean Average Precision) is an important metric for evaluating the overall performance of model detection results, which specifically includes mAP@0.5 and mAP@0.5:0.95. mAP@0.5 represents the average detection accuracy at an Intersection over Union (IoU) threshold of 0.5, reflecting the model's performance under basic localization requirements. mAP@0.5:0.95 comprehensively measures the robustness of the model under different localization accuracy requirements by calculating the average mAP across multiple IoU thresholds from 0.5 to 0.95 with a step size of 0.05.

$$\mathrm{AP} = \int_0^1 \mathrm{P}(R)\mathrm{d}R \times 100\% \tag{14}$$

$$\mathrm{mAP} = \frac{1}{n}\sum_{i=1}^{n-1} \mathrm{AP}_i \times 100\% \tag{15}$$

4 Results

4.1 Detection effect of undivided dataset

Before dividing the dense and non-dense datasets, the YOLOv7 model was used to detect the rapeseed seedling, and the detection results are shown in Figure 7. As shown in Figure 7a and Figure 7b, false detection and missing detection of overlapping detection boxes occurred, which are especially significant in the dense region of rapeseed seedlings. Therefore, the dense and non-dense areas are divided, and the detection algorithms of rapeseed seedlings are improved and analyzed respectively to improve the accuracy and robustness of the model in dense scenarios.

a. False detection of rapeseed seedling b. Missing detection of rapeseed seedling

Figure 7 Effectiveness of rapeseed seedling detection

4.2 Experimental results of different attention mechanisms

After testing the seedlings, it is found that there are problems of false detection and missing detection. In view of these two kinds of problems, attention modules were added to the backbone of YOLOv7 respectively, and comparative tests were conducted. The test results are listed in Tables 1 and 2. For the non-dense dataset, compared with the addition of SE and CBAM attention modules, accuracy was improved up to 95.23%, and it had the highest mean improvement in average accuracy. For the dense dataset, after adding CBAM attention mechanism, the accuracy improves to 94.57%.

Table 1 Comparison of results of different attention mechanisms in non-dense dataset

Method	P/%	R/%	mAP@0.5/%	mAP@0.5:0.95/%
YOLOv7	95.10	93.54	96.37	58.58
SE	94.26	92.63	95.58	55.46
CBAM	95.16	93.38	96.21	57.72
CA	95.23	93.20	96.46	58.58

Table 2 Comparison of results of different attention mechanisms in dense dataset

Method	*P*/%	*R*/%	mAP@0.5/%	mAP@0.5:0.95/%
YOLOv7	93.67	93.47	95.67	55.62
SE	92.63	92.43	94.58	53.79
CBAM	94.57	92.21	95.63	61.88
CA	93.02	92.72	95.10	54.80

4.3 Test results of different loss functions

For the non-dense dataset, improvements were made on the basis of adding CA attention mechanism, and different loss functions were substituted for comparison. The comparison results are listed in Table 3. It can be seen that, after substituting the DIoU loss function, the overall accuracy increases by 2.55%, mAP@0.5 increases by 1.86%, and mAP@0.5:0.95 increases by 8.28%.

Table 3 Comparison of loss functions of different IoU in non-dense dataset

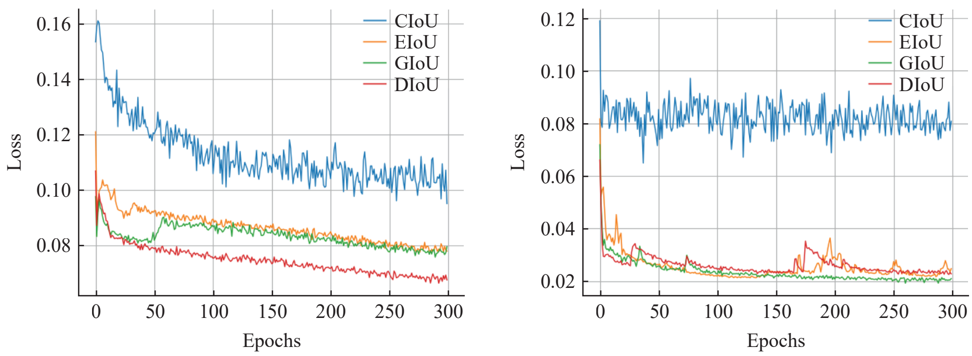
Method	*P*/%	*R*/%	mAP@0.5/%	mAP@0.5:0.95/%
CIoU	95.23	93.2	96.46	58.58
EIoU	97.42	96.00	98.22	66.13
DIoU	97.78	96.08	98.32	66.86
GIoU	97.69	96.57	98.50	67.75

For the dense dataset, the CBAM attention mechanism is added and improved, and different loss functions are substituted for comparison. The comparison results are listed in Table 4. As shown, after substituting the GIoU loss function, the overall accuracy is increased by 3.28%, mAP@0.5 is increased by 1.28%, and mAP@0.5:0.95 is increased by 8.02%.

Table 4 Comparison of loss functions of different IoU in dense dataset

Method	*P*/%	*R*/%	mAP@0.5/%	mAP@0.5:0.95/%
CIoU	94.57	92.21	94.63	61.88
EIoU	95.71	89.61	94.49	62.30
DIoU	96.12	87.96	92.72	60.34
GIoU	97.85	94.53	95.91	69.90

The convergence performance of different loss functions is shown in Figure 8. As can be seen from Figure 8a, EIoU, GIoU, and DIoU are significantly better than the model without a replacement function. The convergence speeds of EIoU and GIoU are similar, and the overall fluctuation of DIoU is small compared with other loss functions. As can be seen from Figure 8b, the convergence rates of EIoU, GIoU, and DIoU are very similar, and GIoU has the best performance.



a. Non-dense dataset with different IoU loss function curves b. Dense dataset with different IoU loss function curves

Figure 8 Loss function curves of different IoU

4.4 Test results before and after YOLOv7 improvement

By integrating the CA attention mechanism into the backbone network of the non-dense dataset, the capability of the model to extract features of rapeseed seedlings is significantly enhanced. Additionally, the DIoU loss function is adopted, which offers notable advantages in accelerating convergence speed and improving positioning accuracy. The detailed comparison of model performance before and after these improvements is presented in Table 5.

Table 5 Comparison of improvements in non-dense datasets

Method	*P*/%	*R*/%	mAP@0.5/%	mAP@0.5:0.95/%
YOLOv7	95.1	93.54	96.37	58.58
Improved YOLOv7	97.78	96.08	98.32	66.86

By introducing CBAM attention mechanism, the dense dataset can improve the generalization ability of the model and use the GIoU loss function as a replacement, which helps to improve the accuracy of the rapeseed seedling detection model in the boundary box regression task. The comparison before and after the improvement is listed in Table 6.

As can be seen from Table 5, after the improvement of the non-dense dataset, the accuracy rate of YOLOv7 model is improved by 2.68%, recall rate by 2.54%, mAP@0.5 by 1.95%, and mAP@0.5:0.95 by 8.28%. As can be seen from Table 6, the accuracy rate of the improved model with the dense dataset increases by 4.18%, recall rate by 1.06%, mAP@0.5 by 0.24%, and mAP@0.5:0.95 by 14.28%. In order to better verify the effectiveness of the improved method, the non-dense and dense datasets were respectively tested before and after the improvement. The result is shown in Figure 9, where the red circles indicate areas with missed or false detections. All three pre-improvement datasets exhibit instances of missed or false detections, while all three post-improvement models achieve accurate identification.

Table 6 Comparison of dense dataset improvements

Method	*P*/%	*R*/%	mAP@0.5/%	mAP@0.5:0.95/%
YOLOv7	93.67	93.47	95.67	55.62
Improved YOLOv7	97.85	94.53	95.91	69.90

4.5 Ablation experiment

In order to verify whether the improved YOLOv7 model is reasonable and effective in practical application, ablation experiments are carefully conducted on the improved dense and non-dense sub-models respectively. For the non-dense sub-models, the CA attention mechanism and DIoU loss function are substituted step by step on the basis of the original YOLOv7 baseline model. For the dense sub-model, the CBAM attention mechanism and GIoU loss function are substituted into the backbone network layer by layer on the basis of the original YOLOv7 baseline model, so as to systematically verify the effectiveness of each key improvement point. The ablation experiment results are listed in Table 7 and Table 8, where "√" indicates that the model included the given item and "-" indicates that the model did not include it.

In Table 7, Experiment 1 is the original YOLOv7 model; under the condition that CIoU loss function is used, mAP@0.5 value is 96.37%. Experiment 2 introduces the CA attention mechanism, and the *P* value and mAP@0.5 increase slightly. Experiment 3 replaces CIoU loss function with DIoU loss function, and both *P* and mAP values increase. Experiment 4 replaces on the basis of the introduction of CA attention mechanism. Compared with the original data, the *P* value increases by 2.68%, mAP@0.5 value increases by 1.95%, and mAP@0.5:0.95 value increases by 8.28%.

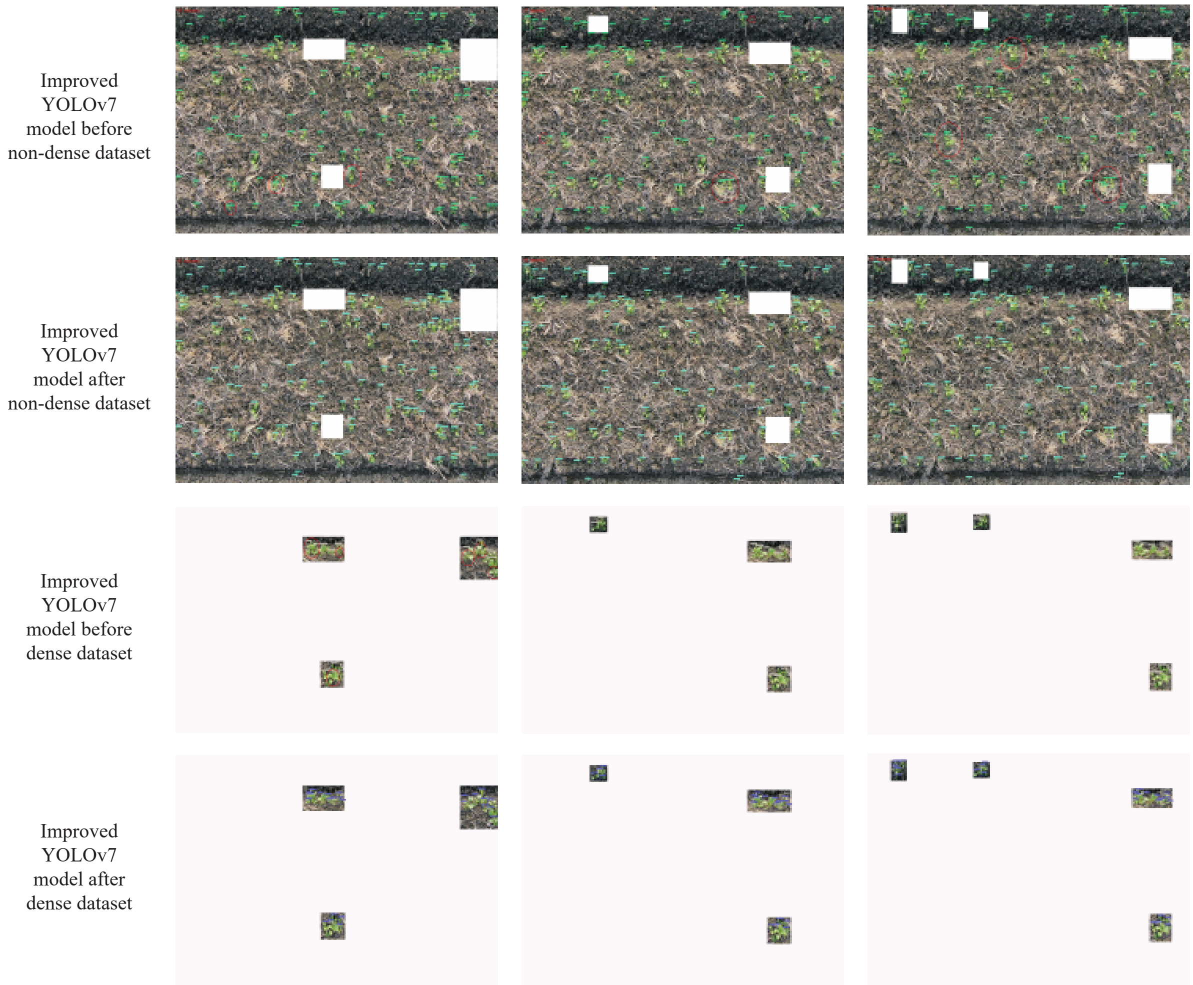


Figure 9 Improved pre- and post-test results

Table 7 Experimental results of non-dense dataset

Test No.	CA	DIoU	*P*/%	*R*/%	mAP@0.5/%	mAP@0.5:0.95/%
1	-	-	95.1	93.54	96.37	58.58
2	√	-	95.23	93.20	96.46	58.58
3	-	√	95.84	94.60	96.94	60.83
4	√	√	97.78	96.08	98.32	66.86

Table 8 Experimental results of dense dataset

Test No.	CBAM	GIoU	*P*/%	*R*/%	mAP@0.5/%	mAP@0.5:0.95/%
1	-	-	93.67	93.47	95.67	55.62
2	√	-	94.57	92.21	95.63	61.88
3	-	√	95.71	94.02	95.58	66.83
4	√	√	97.85	94.53	95.91	69.9

Similarly, in Table 8, Experiment 1 is the original model. Experiment 2 introduces CBAM attention mechanism, and *P* and mAP@0.5:0.95 are improved. Experiment 3 replaces the original loss function with GIoU; the precision value increases by 2.04%, while the mAP@0.5:0.95 value increases by 11.21%. Experiment 4 introduces CBAM attention module and uses the GIoU loss function to replace the original one. Compared with the original, the precision value increases by 4.18%, mAP@0.5 value increases by 0.24%, and mAP@0.5:0.95 value increases by 14.28%.

4.6 Comparison of recognition effect of different algorithms

In order to verify the effectiveness of the rapeseed seedling counting algorithm proposed in this paper, four different models including SSD, YOLOv5, YOLOv7, and YOLOv7-tiny were selected to identify rapeseed seedlings respectively. The recognition effects are shown in Figure 10, where red circles indicate areas with missed or false detections. While the SSD, YOLOv5, YOLOv7, and YOLOv7-tiny models all exhibit varying degrees of missed or false detections, the improved YOLOv7 model achieves accurate identification in all comparative cases. Experimental results demonstrate that the proposed enhanced model effectively improves detection reliability for rapeseed seedlings in dense environments, significantly reducing the occurrence of missed and false detections.

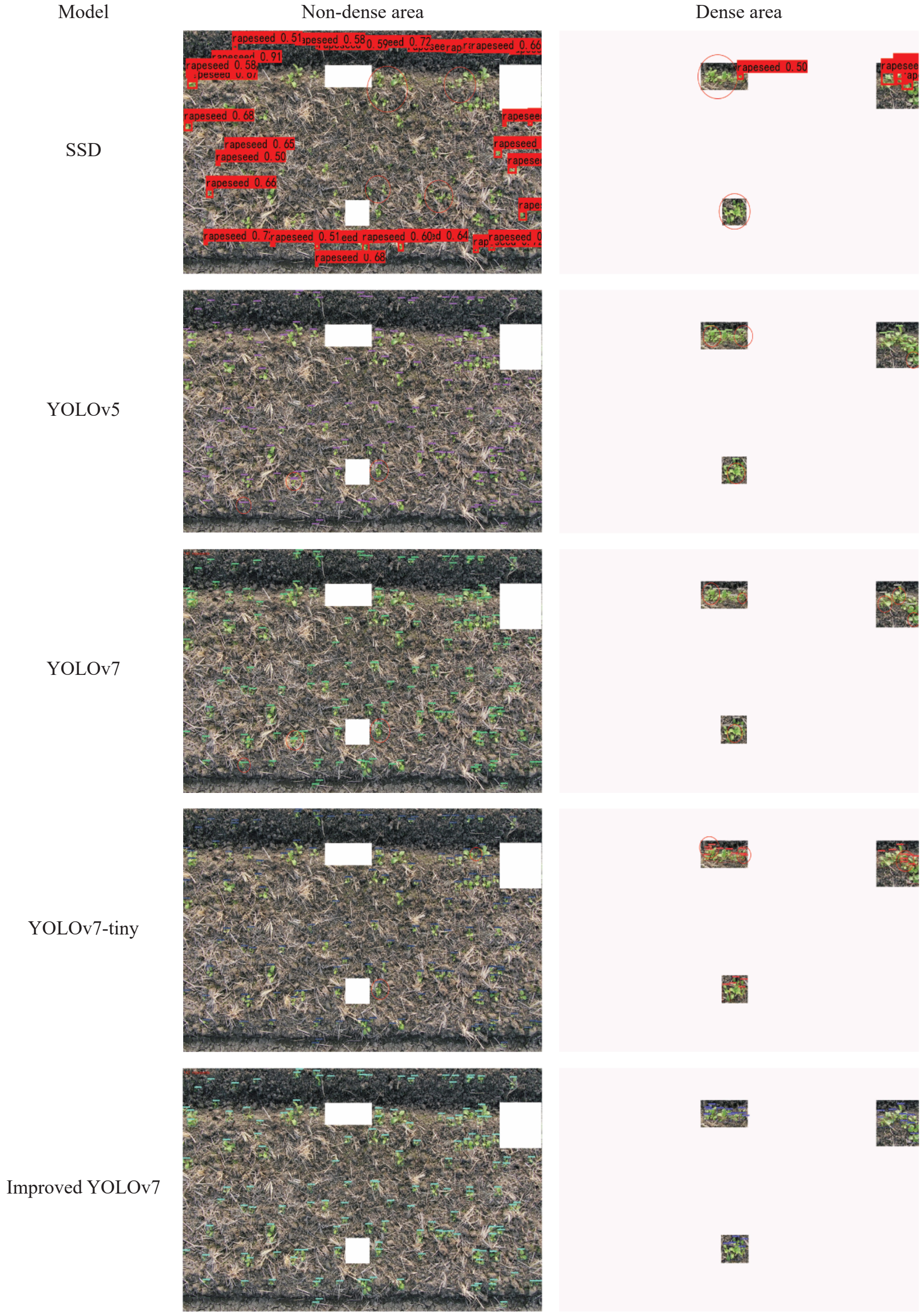


Figure 10 Effects of recognition with different algorithms

4.7 Improved before and after counting comparison

To validate the effectiveness of the proposed improvements to the YOLOv7 model, comprehensive ablation experiments were systematically conducted on both the enhanced dense and non-dense model variants. For the non-dense model configuration, the CA mechanism was integrated and the DIoU loss function was substituted into the baseline YOLOv7 architecture. Correspondingly, for the dense model variant, the CBAM was incorporated alongside the replacement with the GIoU loss function within the backbone network. These controlled experimental configurations were specifically designed to isolate and evaluate the individual contributions of each modification, thereby providing empirical evidence for the rationality of the proposed enhancements.

In Figure 11, the purple circle indicates misdetection, the yellow circle indicates missed detection, and the red rectangular box indicates dense area. The counting results are listed in Table 9. It can be seen from the results that the number of improved test counts are all increased, and there are both false detection and missing detection in groups 1 and 2, but the number of missing detections is significantly more than the number of false detections. The counting results also show an increase, which further verifies the feasibility of the proposed method.

Table 9 Counting test results

Test number	Undifferentiated count	Integration counting	Increased number
1	130	131	1
2	147	159	12
3	168	192	24

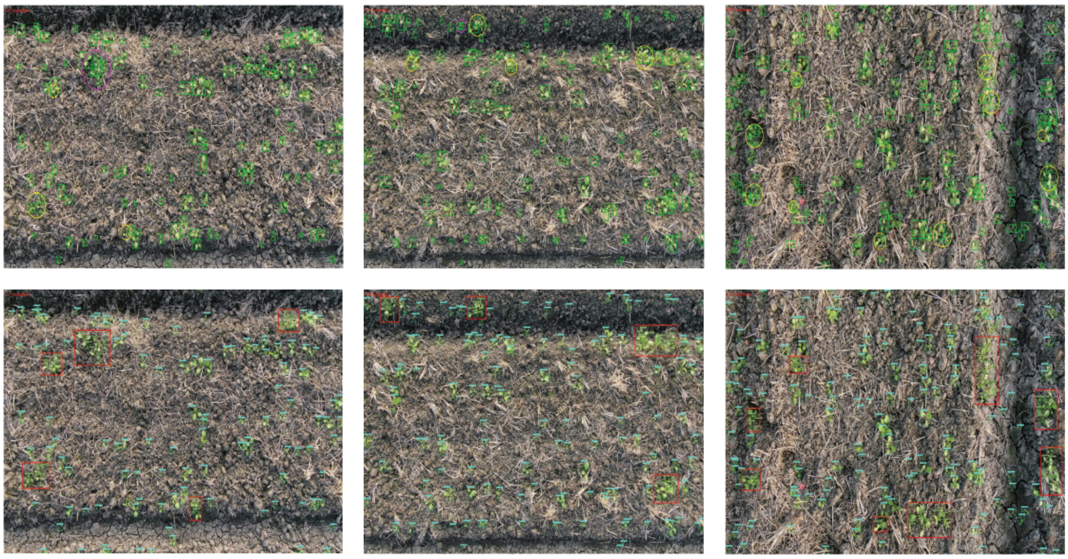


Figure 11 Improved before and after counting comparison

5 Conclusions

In this study, based on a novel dataset-making method and an improved target detection algorithm, a solution for accurate seedling recognition under different densities is constructed. The dataset production method that automatically divides dense and non-dense datasets and the oilseed rape seedling recognition model based on improved YOLOv7 are proposed. The main conclusions are as follows:

In order to focus more on the more critical parts of the image and extract more comprehensive and richer feature information, the CA attention mechanism and CBAM attention mechanism are added to the backbone network, respectively. The results show that, for the non-dense dataset, compared with the addition of the SE and CBAM attention modules, the accuracy of the model is improved to 95.23% after the addition of the CA attention module, and the average mean accuracy improvement is the highest. For the dense dataset, the accuracy improves to 94.57% after adding the CBAM attention mechanism.

On the basis of introducing the attention mechanism, loss functions DIoU and GIoU are used to replace the original loss functions. The results show that for non-dense datasets, the overall accuracy is improved by 2.68%, mAP@0.5 by 1.86%, and mAP@0.5:0.95 by 8.28% after replacing with the DIoU loss function. For dense areas, the overall accuracy is improved by 3.28%, mAP@0.5 by 0.24%, and mAP@0.5:0.95 by 8.02% after replacing with the GIoU loss function.

The improved YOLOv7 model in this paper is compared with the SSD, YOLOv5, YOLOv7, and YOLOv7-tiny network models. The results show that the improved YOLOv7 model has the best performance in recognition effect and counting, and the number of counting has increased in different degrees in the test. Compared to the original YOLOv7 model, the accuracy of the improved non-dense area is increased by 2.68%, mAP@0.5 is increased by 1.95%, and mAP@0.5:0.95 is increased by 8.28%. The accuracy of the improved dense area is increased by 4.18%, mAP@0.5 is increased by 0.24%, and mAP@0.5:0.95 is increased by 14.28%.

Acknowledgements

This study was supported by the National Natural Science Foundation of China (Grant No. 52305241) and the National Key Research and Development Program of China (Grant No. 2022YFD2301402).

[References]

[1] Wang Q. Development status and countermeasures of oilseed rape industry in Anhui Province. Modern Agricultural Science and Technology, 2023(10): 213–216. (in Chinese) DOI: 10.3969/j.issn.1007-5739.2023.10.054.

[2] Zhao D H, Cao Z W, Chen L J, Zhang G X, Zhu Y Z, Han J. Optimising the effect of nitrogen on winter oilseed rape grain yield in China: A meta-analysis. European Journal of Agronomy, 2023; 144: 126755.

[3] Xu X, Zhao Y. A Band Math-ROC operation for early differentiation between *Sclerotinia sclerotiorum* and *Botrytis cinerea* in oilseed rape. Computers and Electronics in Agriculture, 2015; 118: 116–123.

[4] Wang H, Han S, Zhou X, Bu R, Wu J. Fertilizer application technology for mechanically sown winter rapeseed in the river basin of Anhui Province. China Agricultural Technology Extension, 2019; S1: 91–93.

[5] Yang H B, Lan Y B, Lu L Q, Gong D C, Miao J C, Zhao J. New method for cotton fractional vegetation cover extraction based on UAV RGB images. Int J Agric & Biol Eng, 2022; 15(4): 172–180.

[6] Zhuang L H, Wang C Y, Hao H Y, Li J H, Xu L Q, Liu S Y, et al. Maize emergence rate and leaf emergence speed estimation via image detection under field rail-based phenotyping platform. Computers and Electronics in Agriculture, 2024; 220: 108838.

[7] Hu W Z, Wane S O, Zhu J K, Li D S, Zhang Q, Bie X T, et al. Review of deep learning-based weed identification in crop fields. Int J Agric & Biol Eng, 2023; 16(4): 1–10.

[8] Lin H B, Lu Y D, Ding R C, Xiu Y F, Yang F Z. Detection of wheat seedling lines in the complex environment via deep learning. Int J Agric & Biol Eng, 2024; 17(5): 255–265.

[9] Rai N, Zhang Y, Ram B G, Schumacher L, Yellavajjala R K, Bajwa S, et al. Applications of deep learning in precision weed management: A review. Computers and Electronics in Agriculture, 2023; 206: 107698.

[10] Meng X P, Li C C, Li J B, Li X Y, Guo F C, Xiao Z. YOLOv7-MA: Improved YOLOv7-based wheat head detection and counting. Remote Sensing, 2023; 15(15): 3770.

[11] Wang S Y, Wu D S, Zheng X Y. TBC-YOLOv7: A refined YOLOv7-based algorithm for tea bud grading detection. Frontiers in Plant Science, 2023; 14: 1223410.

[12] Ma H X, Dong K B, Wang Y F, Wei S H, Huang W G, Gou J P. Research on lightweight plant recognition model based on improved YOLOv5s. Transactions of the CSAM, 2023; 54(8): 267–276. (in Chinese)

[13] Zhang P, Li D. YOLO-VOLO-LS: A novel method for variety identification of early lettuce seedlings. Frontiers in Plant Science, 2022; 13: 806878.

[14] Zhu Z, He Y, Li W, Cai Z, Wang Q, Ma M. Improved YOLOv7 model for duck egg recognition and localization in complex environments. Transactions of the CSAE, 2023; 39(11): 274–285. (in Chinese)

[15] Zhao K, Zhao L, Zhao Y, Deng H. Study on lightweight model of maize

seedling object detection based on YOLOv7. Applied Sciences, 2023; 13(13): 7731.

[16] Li Z D, Zhang Y F, Wang Y H, Zhao Q H, Liu L C, Zhang T, et al. Identification of rapeseed seedling number based on YC-YOLOv7 model. Transactions of the CSAM, 2024; 55(12): 322–332. (in Chinese)

[17] Li L, Long C F, Yang Y J, Wang X, Liu X B. A detection method for rapeseed and weeds at the seedling stage based on FEPW R-CNN. Agriculture and Technology, 2025; 45(10): 35–40. (in Chinese)

[18] Yang H, Liu Y, Wang S, Qu H, Li N, Wu J, et al. Improved apple fruit target recognition method based on YOLOv7 model. Agriculture, 2023; 13: 1278.

[19] Wu W, Li X L, Hu Z H, Liu X Z. Ship detection and recognition based on improved YOLOv7. Comput Mater Contin, 2023; 76: 489–498.

[20] Pan Y Y, Zhu N Z, Ding L, Li X H, Goh H-H, Han C, et al. Identification and counting of sugarcane seedlings in the field using improved faster R-CNN. Remote Sens, 2022; 14: 5846.

[21] Zhao L L, Zhu M L. MS-YOLOv7: YOLOv7 based on multi-scale for object detection on UAV aerial photography. Drones, 2023; 7: 188.

[22] Li Q X, Ma W J, Li H, Zhang X D, Zhang R Y, Zhou W H. Cotton-YOLO: Improved YOLOv7 for rapid detection of foreign fibers in seed cotton. Computers and Electronics in Agriculture, 2024; 219: 108752.

[23] Meng J, Kang F, Wang Y, Tong S, Zhang C, Chen C. Tea buds detection in complex background based on improved YOLOv7. IEEE Access, 2023; 11: 88295–88304.

[24] Wang J B, Wu J, Wu J W, Wang J P, Wang J. YOLOv7 optimization model based on attention mechanism applied in dense scenes. Applied Sciences, 2023; 13: 9173.

[25] Liu Y, Wang H R, Liu Y H, Luo Y Y, Li H Y, Chen H F, et al. A trunk detection method for camellia oleifera fruit harvesting robot based on improved YOLOv7. Forests, 2023; 14: 1453.

[26] Jiang T Y, Li C, Yang M, Wang Z L. An improved YOLOv5s algorithm for object detection with an attention mechanism. Electronics, 2022; 11: 2494.

[27] Sozzi M, Cantalamessa S, Cogato A, Kayad A, Marinello F. Automatic bunch detection in white grape varieties using YOLOv3, YOLOv4, and YOLOv5 deep learning algorithms. Agronomy, 2022; 12: 319.

[28] Li H R, Yan E P, Jiang J W, Mo D K. Monitoring of key *Camellia Oleifera* phenology features using field cameras and deep learning. Computers and Electronics in Agriculture, 2024; 219: 108748.