Dense strawberry maturity recognition neural network based on multichannel attention mechanism

Zhaorui Cao[1], Chunwen Wang[1], Xiaojie Xin[1], Xinyao Chen[1], Jiahui Cai[2*]

(1. *College of Equipment Engineering, Shenyang Ligong University, Shenyang 110159, Liaoning, China*;
2. *College of Food Science and Engineering, Bohai University, Jinzhou 121013, Liaoning, China*)

Abstract: Color is an important indicator of strawberry maturity; therefore, identifying color changes during harvesting and delivery is important. The soft skin of strawberries can easily cause mechanical damage during harvesting and manual selection. While scholars have made significant progress in using machine vision technology for crop detection, detecting strawberries that are densely placed, differently sized, and in different growth states remains challenging. This study proposes a YOLOv8s-based dense strawberry maturity recognition convolutional neural network that integrates multiscale feature attention. The proposed model addresses several issues, including the overlapping of adjacent fruit features, fuzzy maturity criteria, and difficulty distinguishing individual fruits when they are densely placed. It utilizes a multichannel attention mechanism to fuse semantic features of strawberries at different scales, enhancing the independent summarization ability of image information of individual strawberries in dense and randomly placed environments. It also introduces wavelet down sampling convolution as the backbone of network layers to enhance the ability to capture detailed features of small strawberries. Furthermore, with the integration of the weighted intersection over union loss function, it optimizes the convergence effect and inference accuracy of network training. On a custom strawberry dataset, model accuracy, recall rate, mAP@0.5, and mAP@0.5:0.95 increased by 1.1%, 1.8%, 1.2%, and 0.8%, respectively, compared to the original YOLOv8s model, showing good recognition accuracy and stability when facing dense strawberry arrays with multiple sizes and arbitrary placement. The proposed model can quickly detect the maturity of individual strawberry fruits in random environments, improve sorting efficiency, and reduce post-harvest losses. This study provides new ideas and technical references for the application of machine vision technology in the field of dense crop maturity recognition.

Keywords: object detection, feature fusion, maturity testing, deep learning
DOI: 10.25165/j.ijabe.20261901.9931

Citation: Cao Z R, Wang C W, Xin X J, Chen X Y, Cai J H. Dense strawberry maturity recognition neural network based on multichannel attention mechanism. Int J Agric & Biol Eng, 2026; 19(1): 295–301.

1 Introduction

Strawberries (*Fragaria ananassa* Duch.) are perennial herbaceous plants with tender skin and bright colors[1,2]. Color is an important indicator of strawberry maturity[3], and traditional agriculture uses manual sorting to pick strawberries. However, the manual picking and post-harvest screening process leads to high cost and mechanical damage, causing fruit decay and deterioration, thereby resulting in economic losses[4-6]. In this context, the high precision and efficiency of machine vision technology, particularly target detection technology, have made it a potential solution to improve the automation of the crop picking process[7-9]. Combining target detection technology with robot technology and applying it to the production process of strawberries and other crops is expected to significantly reduce labor costs and improve picking efficiency.

Scholars have made significant achievements in strawberry image recognition, including monitoring strawberry growth state in complex environments. An et al.[10] proposed a 3 convolutional CSP Bottleneck with Horizontal and Bidirectional connections (C3HB) feature extraction module based on the You Only Look Once version X (YOLOX) model and introduced the Normalization-based Attention Module (NAM) attention mechanism and Scylla Intersection over Union (SIoU) loss function, which effectively improved the detection accuracy and recall rate of strawberry fruits in five growth states, and the model size remained relatively stable. Du et al.[11] proposed the DSW-YOLO network model to solve the problem of accurately detecting ripening strawberries under occlusion. The average detection accuracy was improved by 2.2% by integrating the DCNv3 model, SA attention mechanism, and wise intersection over union (WIoU) V3 loss function. It was successfully tested on a Jetson Xavier NX computing core to verify its rapid detection ability in a field environment. To distinguish the different growth stages of strawberries more accurately, Wang et al.[12] designed the DSE-YOLO network, extracted the multidimensional features of images by using the DSE module, and constructed the DEMSE loss function, which significantly improved the recognition accuracy of the model for maturity of strawberries. Shen et al.[13] proposed the RTF-YOLO network model to solve the problem of automatic harvesting and yield prediction of strawberries. The network enhanced the ability to extract strawberry features and improved detection accuracy by introducing an efficient convolution module and a triple attention mechanism. He et al.[14] proposed the use of KTD-YOLOv8 to enhance the accuracy and efficiency of disease detection in strawberry leaves.

Received date: 2025-05-22 **Accepted date:** 2026-01-04
Biographies: Zhaorui Cao, PhD, research interest: machine vision, Email: caozhaorui@163.com; **Chunwen Wang**, MS, research interest: machine vision, Email: 544024270@qq.com; **Xiaojie Xin**, MS, research interest: machine vision, Email: 1425332761@qq.com; **Xinyao Chen**, MS, research interest: machine vision, Email: 1340865646@qq.com.
***Corresponding author: Jiahui Cai**, PhD, research interest: post-harvest storage and preservation of fruits and vegetables. Address: College of Food Science and Engineering, Bohai University, Jinzhou, Liaoning 121013, China. Tel: +86-13654216566, Email: caijiahui@bhu.edu.cn.

Additionally, the Triplet Attention mechanism was integrated to better extract and fuse multiscale features, and a parameter-sharing diverse branch block (DBB) was added to enhance the model's ability to process targets at different spatial scales. This novel approach led to a 2.8% increase in the average accuracy and a 38.5% reduction in floating-point calculations compared with the original YOLOv8. Li et al.[15] proposed a lightweight multistage detection method based on YOLOv7-tiny to detect strawberries in complex environments. A coordinate attention module was introduced to improve the focus on the target area, and the WIoU loss function was integrated to enhance the detection of occluded strawberries. With the adoption of the Adan gradient descent algorithm and a dedicated detection head for small targets, this method achieved a 2.44% improvement in mAP@0.5, along with a reduction in GFLOPs and Params.

Despite significant progress in previous research, strawberries were mostly recognized and detected under conditions of neat arrangement and uniform size. However, in practical situations, individual strawberries are closely spaced, vary in size, and differ in maturity. Due to the attachment of semantic feature information in images, this can affect the final recognition results. Therefore, rapid recognition of strawberry maturity, considering dense arrangement, different sizes, and varying fruit states, remains a key issue that urgently needs to be addressed. Therefore, this study proposes a dense strawberry maturity recognition neural network (DSMRNN) based on a multichannel attention mechanism to improve strawberry maturity recognition efficiency under random and messy placement. The semantic features of independent strawberries of different sizes were deeply mined and isolated by constructing a feature extraction and fusion network with a wavelet downsampling convolution module and a multichannel attention mechanism, and the information series and maturity pollution of individuals with different sizes, qualities, and maturity were overcome. Simultaneously, WIoU was used to optimize the convergence efficiency and maturity detection accuracy of network training and reduce the impact of uneven proportions of strawberry datasets with different qualities on network training and detection results. The proposed method can be used to automatically detect the maturity of strawberries under arbitrary placement and provide strong technical support for the automation and intelligence of strawberry picking and quality analysis.

2 Algorithm structure

2.1 DSMRNN model structure

Since the algorithm is targeted at edge-based real-time strawberry detection computing platforms (such as portable computers, drones, autonomous vehicles, etc.), and the accuracy of YOLOv8s can meet the detection task requirements of this paper, a full-sized neural network with high computational overhead is not chosen. The comparison of YOLOv8 model parameters is shown in Table 1.

Table 1 Comparison of model parameters

Model	Params/M	FLOPs/GFLOPs	Time/ms
YOLOv8s	11.2	28.6	2.5
YOLOv8m	25.9	78.9	5.1
YOLOv8l	43.7	165.2	7.6
YOLOv8x	68.2	257.8	11.3

A DSMRNN network (Figure 1) was constructed to rapidly and accurately identify the maturity of multiple strawberry individuals in dense and random placements. The network is based on the YOLOv8 series. It constructs WTCM to preserve boundary details and separate conjoined strawberries; constructs SJIFM to enable the model to maintain small-scale maturity features in dense scenes; and introduces WIoU to enhance the robustness of imbalanced maturity samples. To elaborate, the wavelet transform convolution module (WTCM) is introduced to enhance the downsampling accuracy of the algorithm for multiscale targets and weaken the influence of image geometric factors, such as strawberry overlap and adhesion, on maturity detection accuracy under arbitrary placement. A CSP bottleneck with two convolutions (C2f) is introduced to aggregate and compress feature graphs through splicing and convolution operations, improving the expression ability and efficiency of the model[16-18]. A convolution block with Batch Normalization and SiLU (CBS) is introduced to realize feature extraction and nonlinear transformation of the input feature map[19-21]. Simultaneously, the space jump information fusion module (SJIFM) is designed to spread the multiscale feature map across layers to ensure that the semantic information of small strawberries that are not sufficiently mature will not be lost because of the deepening of the network. In addition, spatial pyramid pooling (SPPF) is introduced to extract features at different scales to enhance the scale invariance of the network, and multiscale feature extraction of the input feature map is realized by combining pooling operations at different scales, improving the generalizability and detection performance of the model[22-24]. DSMRNN generates three groups of strawberry feature maps (Output 1, Output 2, Output 3) with different scales for parallel detection after completing the feature extraction calculation to reduce the influence of the semantic features of small strawberries on large ones[25-27]. Finally, Wise-IoU is used to supervise the convergence of network training to reduce the influence of unbalanced strawberry datasets with different maturities on the network training results.

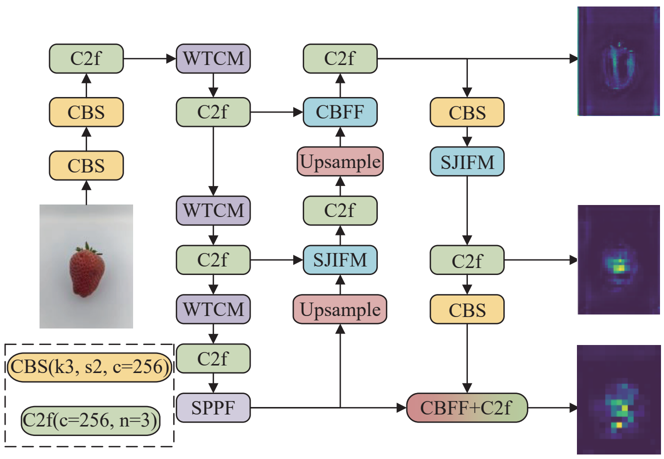


Figure 1 DSMRNN network structure

2.2 Wavelet transform convolution module

Since individual strawberries are prone to conglutination in random and dense placement environments, some semantic features are confused, further affecting the accuracy of judging individual maturity. To solve this problem, this study introduces WTCM into the network to reduce spatial information loss in the process of convolution pooling, improving the ability to extract and segment strawberry feature information in a dense environment. WTCM first decomposes the image features through a low-pass filter h_0 and high-pass filter h_1 to obtain four components: diagonal high-frequency component HH, vertical high-frequency component HL, low-frequency component LL, and horizontal low-frequency component LH[28-30]. Subsequently, the four components are connected by the CBS module to obtain a new feature in which the resolution is reduced to 1/2, and the number of channels is four. The above process is expressed as follows:

$$\mathrm{LL} = X * h_0 \cdot h_1 \tag{1}$$

$$\mathrm{LH} = X * h_0 \cdot h_1^T \tag{2}$$

$$\mathrm{HL} = X * h_1 \cdot h_0^T \tag{3}$$

$$\mathrm{HH} = X * h_1 \cdot h_1^T \tag{4}$$

$$Y = \mathrm{CBS}\left[\mathrm{LL,LH,HL,HH}\right] \tag{5}$$

where, X is the input characteristic diagram, and Y is the output characteristic diagram. Through this method, WTCM can characterize each individual strawberry with four features, namely HH, HL, LL, and LH, in a single network module through the strategy of parallel feature reservation. This approach ensures at least one feature dimension difference between different adjacent strawberries and enhances the spatial discrimination ability of DSMRNN for adhered individuals.

2.3 Space jump information fusion module

Owing to the contextual differences in semantic information between strawberry feature maps of different scales, if the feature extraction layer is excessively deepened to improve network detection accuracy, the semantic features of small strawberries will be lost easily in the deep network. This study introduces SJIFM, as shown in Figure 2, for cross-level feature fusion to solve the interference caused by different strawberry sizes, detection standards, and semantic information intensity in the natural environment. The module preserves the semantic information of small strawberries and ensures that it can be transmitted to the deep network. SJIFM first standardizes the sizes of the input images F_1 and F_2 through two 1×1 convolutions, adds the two output elements to obtain the superimposed features, then uses the convolutional block attention module (CBAM) to improve the perception ability of the model, and finally multiplies it with the input features by elements to obtain more strawberry features by multiplying the semantic context information of the images.

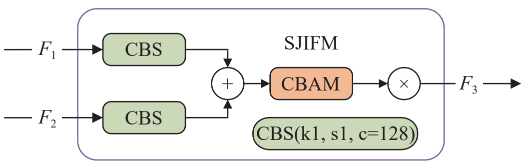


Figure 2 SJIFM feature fusion module

The definition of the SJIFM module is given in Equation (6), where F_1 and F_2 are the feature inputs of different scales, and F_3 is the output of the SJIFM module:

$$F_3 = \mathrm{CBAM}(\mathrm{CBS}(F_1) + \mathrm{CBS}(F_2)) \times F_2 \tag{6}$$

The CBAM module in SJIFM can overcome the problem of conventional convolution being insensitive to information of different scales, shapes, and directions[11,31,32]. Its structure is illustrated in Figure 3. The attention mechanism of CBAM introduces the channel attention module (CAM) and the spatial attention module (SAM). CAM can enhance the characteristics of different channels, whereas SAM can extract key information from different locations in space. CBAM multiplies the output elements of CAM and SAM to reduce irrelevant information and retain key information, thereby enhancing the network's perception of strawberries of different sizes.

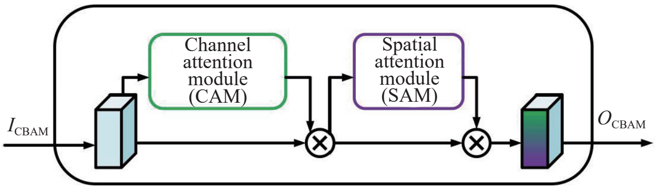


Figure 3 CBAM attention mechanism

The CBAM attention mechanism is defined as shown in Equation (7) in which Spatial represents spatial attention mechanism transformation, Channel represents channel attention mechanism transformation, and ICBAM and OCBAM are the input and output of the module, respectively:

$$F_2 = \mathrm{Spatial}(\mathrm{Channel}(F_1) \times F_1) \times (\mathrm{Channel}(F_1) \times F_1) \tag{7}$$

2.4 Wise-IoU loss function

For strawberry data samples collected from natural environments, achieving a complete balance of maturity samples for all individual strawberries is challenging. When the collection environment and conditions are good, strawberries with full size and high maturity account for more in the training set; however, in the case of an unsatisfactory collection environment, strawberries of average size and mixed quality constitute the majority. This data imbalance significantly affects the ability of the algorithm to perceive the maturity characteristics of strawberries, ultimately leading to difficulties in distinguishing between mature and immature individuals. Therefore, this study used the Wise-IoU loss function to balance the influence of asymmetric samples on the network training and reasoning process, thereby solving the issue.

Wise-IoU (WIoU) can weaken the interference of geometric factors on the accuracy of network training and reasoning when the predicted frame coincides well with the real frame[33,34]. In WIoU, L_{IoU} is used to balance the gradient gain of samples with different acquisition difficulties, making the network more focused on a common recognition frame. L_{IoU} in WIoU is defined as follows:

$$L_{\mathrm{IoU}} = 1 - \mathrm{IoU} \tag{8}$$

$$R_{\mathrm{WIoU}} = \exp\left(\frac{(x - x_{\mathrm{gt}})^2 + (y - y_{\mathrm{gt}})^2}{(W_g^2 + H_g^2)^n}\right) \tag{9}$$

$$L_{\mathrm{WIoUv1}} = R_{\mathrm{WIoU}} L_{\mathrm{IoU}} \tag{10}$$

To prevent the convergence speed from slowing down, L_{IoU} adds the monotonic focusing coefficient $L_{\mathrm{IoU}}^{\gamma}$ and uses the normalization factor L^* to accelerate the network convergence. The process is as follows:

$$L_{\mathrm{WIoUv2}} = L_{\mathrm{IoU}}^{\gamma^*} L_{\mathrm{WIoUv1}},\ \ \gamma > 0 \tag{11}$$

$$L_{\mathrm{WIoUv2}} = \left(\frac{L_{\mathrm{IoU}}^*}{L_{\mathrm{IoU}}}\right)^{\gamma} L_{\mathrm{WIoUv1}} \tag{12}$$

In addition, this study used the dynamic nonmonotonic focusing mechanism[34] to reduce the introduction of low-quality strawberry image samples, thereby improving the overall performance of the detector. The dynamic nonmonotonic focusing mechanism is defined as follows:

$$\beta = \frac{L_{\mathrm{IoU}}^*}{L_{\mathrm{IoU}}} \in [0, +\infty] \tag{13}$$

$$L_{\text{WIoUv3}} = rL_{\text{WIoUv1}}, \quad r = \frac{\beta}{\delta\alpha^{\beta-\delta}} \tag{14}$$

where, β is the outlier; α and δ are manually set hyperparameters (β=1.9, α=1.7, δ=0.5). Based on the above method, WIoU can effectively regulate the contribution of various strawberry image samples to network training to prevent high false detection rates of color break strawberries caused by training tendency deviation.

3 Experimental tests

3.1 Experimental environment

The experiments were conducted using Windows 10, AMD epyc7642, NVIDIA RTX4060 GPU; 16 GB memory, 12 GB main memory, Python 3.8, PyTorch 1.11, CUDA 11.3; and PyCharm 2023.1 (training software). The network training parameters were set as follows: epochs were set to 300 and the batch size was 8.

3.2 Training conditions

A strawberry plant dataset captured by the research team was adopted for training and testing, and the fruit was shot at 10 to 20 cm. The images in the dataset were collected from a greenhouse during the day and were divided into three categories: green strawberries (completely immature), color break strawberries (partially mature), and fully ripe strawberries. The identification criteria are as follows: Red: 220±50, Green: 0+10, Blue: 24±10. If the detection area is greater than 80%, it is fully cooked; if it is between 40% and 80%, it is partially cooked; if it is less than 40%, it is uncooked. A total of 3153 pictures were selected for training DSMRNN, including 2522 training sets, 631 verification sets, and 624 test images; the ratio of sample categories is 3.1:3.5:3.4. During the training process, strategies such as random horizontal flipping, scaling, translation, and color space jittering were employed. The input image resolution was adjusted to 640×640 to ensure a balance between detection accuracy and computational efficiency. Image normalization was performed before training, and corresponding annotations were adjusted synchronously based on geometric transformations. The loss function and accuracy convergence curves of the training results are shown in Figure 4. The total number of epochs was 300, model accuracy was 90.5%, recall rate was 93.9%, mAP@0.5 was 96.4%, and mAP@0.5 was 89.8%.

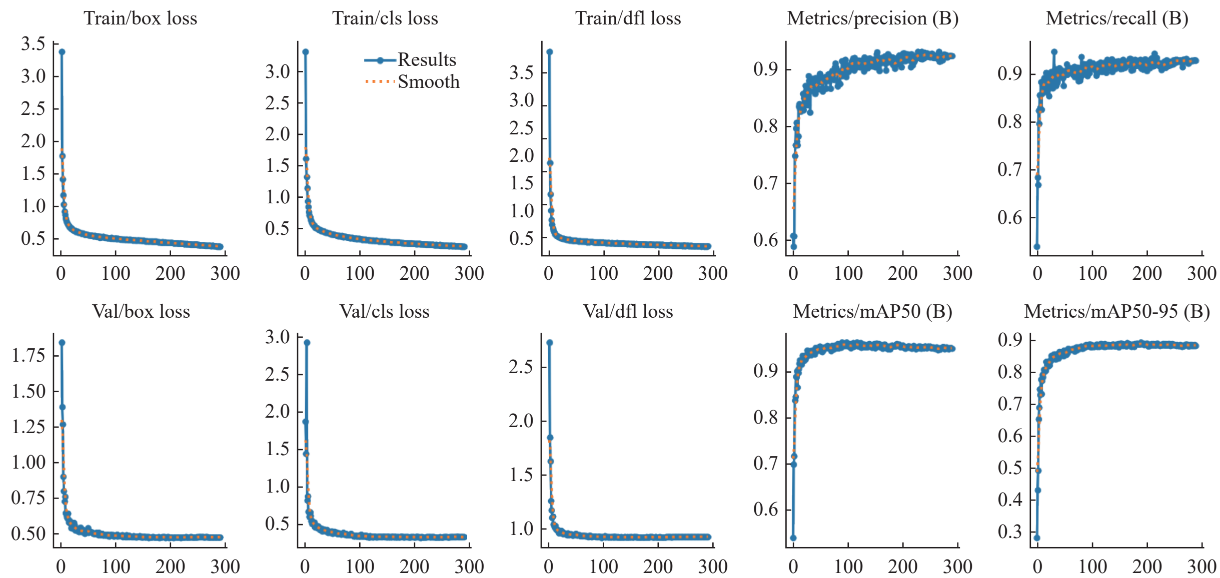


Figure 4 DSMRNN loss function and accuracy curve

3.3 Test results and analysis

Ablation experiments were performed on each module to verify the performance of the proposed model. Experiment 1 used the original network (YOLOv8s) without any modules, whereas Experiment 2 used the WTCM module for feature extraction. Experiment 3 added the SJIFM module, and Experiment 4 introduced the WIoU loss function. Evaluations were performed in terms of precision, recall, and average precision (mAP), and the results are listed in Table 2.

After using WTCM, the accuracy and recall rates of the model increased by 0.2%, 0.3%, 0.3%, and 0.2%, respectively. After the SJIFM module was introduced, all the indices increased to 1%, 1.8%, 0.8%, and 0.5%, respectively (Table 2). After the WIoU loss function was introduced, the accuracy and recall did not change noticeably; however, both mAP@0.5 and mAP@0.5:0.95 increased by 0.2%. From this experiment, it is observed that the WTCM, SJIFM, and WIoU proposed in this study can improve the computing performance of a network from different aspects.

Table 2 Results of ablation experiment

No.	WTCM	SJIFM	WIoU	P/%	R/%	mAP@0.5		Total	mAP@0.5:0.95		Total
						Packaged	Dense		Packaged	Dense	
Experiment 1	-	-	-	89.2	91.8	95.8	94.4	95.1	89.5	88.3	88.9
Experiment 2	√	-	-	89.4	92.1	96.1	94.7	95.4	89.7	88.5	89.1
Experiment 3	√	√	-	90.4	93.9	96.8	95.6	96.2	90.2	89.0	89.6
Experiment 4	√	√	√	90.5	93.9	97.0	95.8	96.4	90.6	89.0	89.8

This study carried out a horizontal comparison test to analyze the performance of the algorithm to further verify the effectiveness and superiority of DSMRNN. In this comparative experiment, the YOLOv3, YOLOv5s, YOLOv8s, YOLOv9s, and YOLOv10s models were used as ring comparison objects, and the same training environment and conditions were adopted. The results of the comparative experiment are listed in Table 3.

Experiments showed that under the same training conditions and environment, compared with the YOLOv8s network, the proposed DSMRNN model improves accuracy by 1.1%, recall by 1.8%, mAP@0.5 by 1.2%, and mAP@0.5:0.95 by 0.8%. Compared with YOLOv10s, the accuracy increased by 0.4%, recall increased by 0.7%, and mAP@0.5 and mAP@0.5:0.95 increased by 0.4%. In addition, compared with other classical models, all the indices were improved to different degrees, which proves the effectiveness of the algorithm and that it is more suitable for detecting the maturity recognition of randomly and densely placed strawberries.

Table 4 compares the effects of different IoU-based loss functions on the detection performance of DSMRNN.

WIoU achieves the best performance across all evaluation metrics. Compared with CIoU, WIoU improves precision, recall, mAP@0.5, and mAP@0.5:0.95 by 0.6%, 1.3%, 0.9%, and 0.7%, respectively.

To explain the detection effect of DSMRNN on strawberry maturity more intuitively, the above experimental test results are shown. In this experiment, according to the case shown in Figure 5, the maturity of strawberries was marked as fully ripened, half ripened, and green strawberries from left to right to distinguish individual strawberries with different maturities.

After determining the classification level of strawberry maturity, this study identified the maturity under common strawberry packaging and placement methods to verify the computing power of DSMRNN in an actual engineering environment. First, the maturity of the strawberries placed neatly in a packaging box was tested. The comparison network was composed of three models: DSMRNN, YOLOv8s, and YOLOv5s, and the experimental results are shown in Figure 6 (columns from left to right).

Table 3 Results of comparative experiment

Model	P/%	R/%	Green	Color break	Fully ripe	mAP@0.5		Total	mAP@0.5:0.95		Total
						Packaged	Dense		Packaged	Dense	
YOLOv3	88.9	91.2	93.4	94.6	95.6	95.2	93.8	94.5	88.7	87.5	88.1
YOLOv5s	89.2	91.8	94.2	95.0	96.1	96.0	94.2	95.1	89.5	88.3	88.9
YOLOv8s	89.4	92.1	94.4	95.2	96.0	96.1	94.3	95.2	89.6	88.4	89.0
YOLOv9s	89.8	92.5	94.8	95.6	96.7	96.4	95.0	95.7	89.7	88.5	89.1
YOLOv10s	90.1	93.2	95.1	96.0	96.9	96.6	95.4	96.0	90.0	88.8	89.4
CBFF-YOLO	90.5	93.9	95.8	96.6	97.4	97.0	95.8	96.4	90.6	89.0	89.8

Table 4 Results of comparative experiment

Name	P/%	R/%	mAP@0.5		Total	mAP@0.5:0.95		Total
			Packaged	Dense		Packaged	Dense	
CIoU	89.9	92.6	96.1	94.9	95.5	89.5	88.7	89.1
SIoU	90.1	93.2	96.5	95.7	96.1	90.1	88.9	89.5
WIoU	90.5	93.9	97.0	95.8	96.4	90.6	89.0	89.8

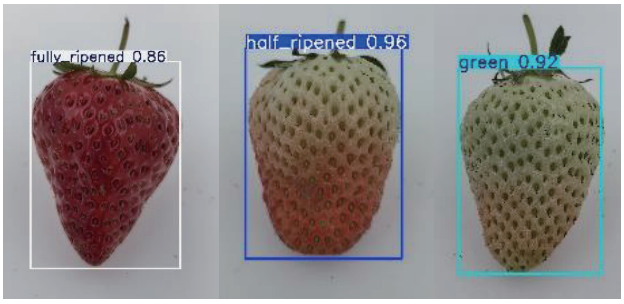


Figure 5 Classification of strawberry maturity

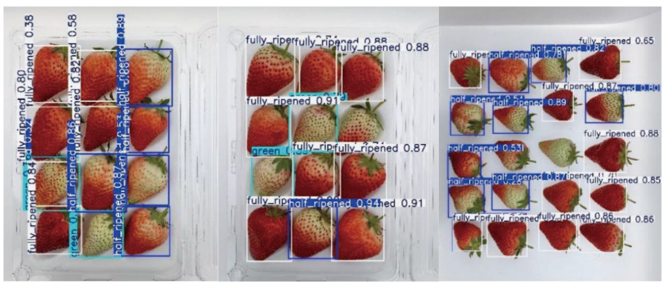
The false detection rate of the DSMRNN was lower than that of the other two classical models, and accuracy improved. In the second group of comparative experiments, from different perspectives, the false detection rate of DSMRNN was lower than that of YOLOv8s, and the missed detection situation was significantly reduced compared with YOLOv5s. In the third group, DSMRNN had significantly fewer missed detections than the other two models when the experimental targets were denser and more in number.

Subsequently, this study compared the detection results of strawberry maturity with those of DSMRNN, YOLOv8s, and YOLOv5s, and the results are illustrated in Figure 7.

YOLOv5s had many false and missed detections when detecting dense strawberries; YOLOv8s had relatively few missed detections, but false detections still occurred. Compared to

a. DSMRNN test results

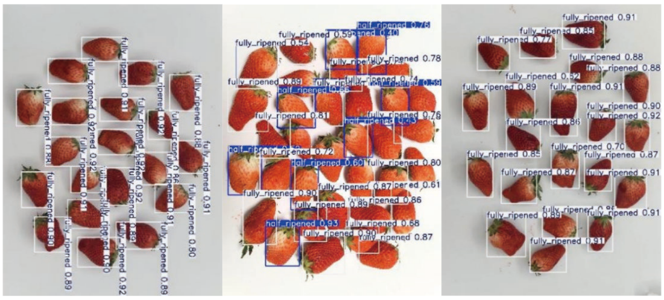
b. YOLOv8s test results



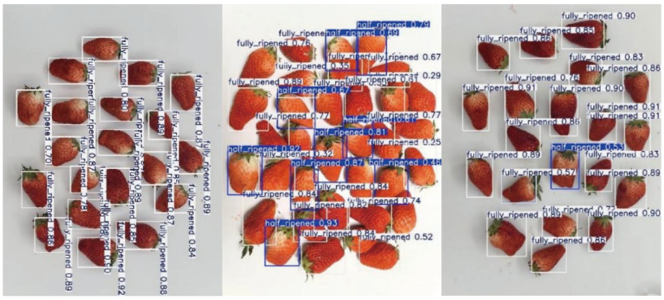
c. YOLOv5s test results

Figure 6 Maturity detection of packaged strawberries

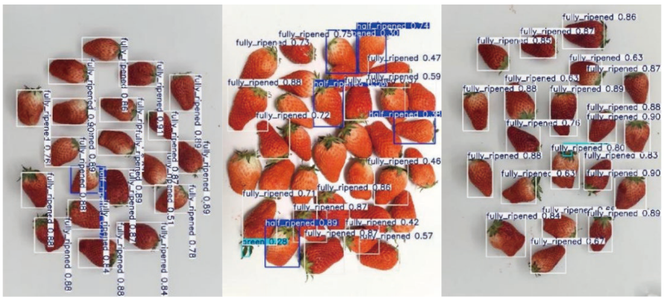
YOLOv8s and YOLOv5s, DSMRNN achieved a lower missed detection rate and more accurately detected strawberry maturity under dense conditions.



a. DSMRNN test results



b. YOLOv8s test results



c. YOLOv5s test results

Figure 7 Maturity detection of intensive strawberries

4 Conclusions

In this study, DSMRNN based on a multichannel attention mechanism was proposed to solve the problem of characteristic interference between fruits in strawberry maturity detection under random and dense placement. Through the convolution module of the wavelet transform, the algorithm improves multiscale target downsampling accuracy and effectively weakens the interference of geometric factors, such as overlap and adhesion, on the accuracy of maturity detection under arbitrary placement. SJIFM is used to carry out cross-layer fusion propagation of multiscale feature maps to ensure that the semantic information of small strawberries that are not sufficiently mature is not lost in the process of network deepening. WIoU is used to supervise the convergence of the network training process to reduce the negative impact of unbalanced strawberry datasets with different maturities on network training results. The experimental results showed that compared with the backbone network, the accuracy of DSMRNN increased by 1.1%, recall rate increased by 1.8%, mAP@0.5 increased by 1.2%, and mAP@0.5:0.95 increased by 0.8%. DSMRNN also outperformed other classical models, which proves that the DSMRNN can detect strawberry maturity when they are neatly packed and dispersed randomly. This algorithm provides a new method for the application of machine vision in strawberry post-harvest screening, improving efficiency and reducing post-harvest losses.

Acknowledgements

This work was supported by Basic Research Project of Liaoning Provincial Department of Education (Grant No. LJ212410144050) and High-level Talent Introduction Research Support Project of Shenyang Ligong University (Grant No. 1010147001254).

[References]

[1] Xu C, Zhang X Y, Liang J, Fu Y J, Wang J, Jiang M, et al. Cell wall and reactive oxygen metabolism responses of strawberry fruit during storage to low voltage electrostatic field treatment. Postharvest Biol Technol, 2022; 192: 112017.

[2] Zhang Y T, Li S L, Deng M Y, Gui R, Liu Y Q, Chen X P, et al. Blue light combined with salicylic acid treatment maintained the postharvest quality of strawberry fruit during refrigerated storage. Food Chem X, 2022; 15: 100384.

[3] Wang L, Luo Z S, Yang M Y, Liang Z, Qi M, Dong Y Y, et al. The action of RED light: Specific elevation of pelargonidin-based anthocyanin through ABA-related pathway in strawberry. Postharvest Biol Technol, 2022; 186: 111835.

[4] De Preter A, Anthonis J, De Baerdemaeker J. Development of a robot for harvesting strawberries. IFAC-PapersOnLine, 2018; 51(17): 14–19.

[5] Wang C L, Wang H M, Han Q Y, Zhang Z G, Kong D D, Zou X J. Strawberry detection and ripeness classification using YOLOv8+ model and image processing method. Agriculture, 2024; 14(5): 751.

[6] Yamamoto S, Hayashi S, Saito S, Ochiai Y, Yamashita T, Sugano S. Development of robotic strawberry harvester to approach target fruit from hanging bench side. IFAC Proceedings Volumes, 2010; 43(26): 95–100.

[7] Yu Y, Zhang K L, Liu H, Yang L, Zhang D X. Real-time visual localization of the picking points for a ridge-planting strawberry harvesting robot. IEEE Access, 2020; 8: 116556–116568.

[8] Yu Y, Zhang K L, Yang L, Zhang D X. Fruit detection for strawberry harvesting robot in non-structural environment based on Mask-RCNN. Comput Electron Agric, 2019; 163: 104846.

[9] Zhao Y S, Liang G, Huang Y X, Liu C L. A review of key techniques of vision-based control for harvesting robot. Comput Electron Agric, 2016; 127: 311–323.

[10] An Q L, Wang K, Li Z Y, Song C Y, Tang X Y, Song J. Real-time monitoring method of strawberry fruit growth state based on YOLO improved model. IEEE Access, 2022; 10: 124363–124372.

[11] Du X Q, Cheng H C, Ma Z H, Lu W W, Wang M X, Meng Z C, et al. DSW-YOLO: A detection method for ground-planted strawberry fruits under different occlusion levels. Comput Electron Agric, 2023; 214: 108304.

[12] Wang Y, Yan G, Meng Q L, Yao T, Han J F, Zhang B. DSE-YOLO: Detail semantics enhancement YOLO for multi-stage strawberry detection. Comput Electron Agric, 2022; 198: 107057.

[13] Shen S, Duan F M, Tian Z W, Han C X. A novel deep learning method for detecting strawberry fruit. Applied Sciences, 2024; 14: 4213.

[14] He Y L, Peng Y F, Wei C Y, Zheng Y D, Yang C C, Zou T Y. Automatic disease detection from strawberry leaf based on improved YOLOv8. Plants, 2024; 13: 2556.

[15] Li C L, Wu H N, Zhang T, Lu J H, Li J H. Lightweight network of multi-stage strawberry detection based on improved YOLOv7-Tiny. Agriculture, 2024; 14: 1132.

[16] Li Y H, Zhang M Y, Zhang C Y, Liang H, Li P, Zhang W W. YOLO-CCS: Vehicle detection algorithm based on coordinate attention mechanism. Digital Signal Processing: A Review Journal, 2024; 153: 104632.

[17] Shan W, Yue Y R. Apple defect detection in complex environments. Electronics, 2024; 13: 4844.

[18] Xiao L S, Li W Z, Yao S, Liu H T, Ren D H. High-precision and lightweight small-target detection algorithm for low-cost edge intelligence. Sci Rep, 2024; 14: 23542.

[19] Chen J J, Zhu J N, Li Z, Yang X B. YOLOv7-WFD a novel convolutional neural network model for helmet detection in high-risk workplaces. IEEE Access, 2023; 11: 113580–113592.

[20] Huang X H, Zhu J H, Huo Y. SSA-YOLO: An improved YOLO for hot-rolled strip steel surface defect detection. IEEE Trans Instrum Meas, 2024; 73: 5040017.
[21] Gao W W, Zong C L, Wang M Y, Zhang H F, Fang Y. Intelligent identification of rice leaf disease based on YOLO V5-EFFICIENT. Crop Protection, 2024; 183: 106758.
[22] Hu X Y, Kong D F, Liu X Y, Zhang J W, Zhang D D. FM-STDNet: High-speed detector for fast-moving small targets based on deep first-order network architecture. Electronics, 2023; 12: 1829.
[23] Wang Q H, Liu Y K, Zheng Q, Tao R, Liu Y. SMC-YOLO: A high-precision maize insect pest-detection method. Agronomy, 2025; 15: 195.
[24] Cao X M, Zhong P, Huang Y H, Huang M T, Huang Z Y, Zou T L, et al. Research on lightweight algorithm model for precise recognition and detection of outdoor strawberries based on improved YOLOv5n. Agriculture, 2025; 15(1): 90.
[25] Chen S L, Liao Y H, Lin F, Huang B. An improved lightweight YOLOv5 algorithm for detecting strawberry diseases. IEEE Access, 2023; 11: 54080–54092.
[26] Weng S Z, Yu S, Guo B Q, Tang P P, Liang D. Non-destructive detection of strawberry quality using multi-features of hyperspectral imaging and multivariate methods. Sensors, 2020; 20: 3074.
[27] Guo R Q, Ji P Y, Zhang Y P, Hu J Q, Liu W L, Li X J, et al. Steel surface defect detection based on improved GCHS-YOLO algorithm. IEEE Access, 2024; 12: 190865–190875.
[28] Su Y, Tan W X, Dong Y F, Xu W, Huang P P, Zhang J X, Zhang D K. Enhancing concealed object detection in active millimeter wave images using wavelet transform. Signal Processing, 2024; 216: 109303.
[29] Xu G P, Liao W T, Zhang X, Li C, He X W, Wu X L. Haar wavelet downsampling: A simple but effective downsampling module for semantic segmentation. Pattern Recognition, 2023; 143: 109819.
[30] Wu H T, Mo X T, Wen S J, Wu K L, Ye Y, Wang Y M, et al. DNE-YOLO: A method for apple fruit detection in diverse natural environments. Journal of King Saud University- Computer and Information Sciences, 2024; 36: 102220.
[31] Xiang S, Wang S Y, Xu M, Wang W Y, Liu W G. YOLO POD: A fast and accurate multi-task model for dense soybean pod counting. Plant Methods, 2023; 19: 8.
[32] Zhao S Y, Liu J Z, Wu S. Multiple disease detection method for greenhouse-cultivated strawberry based on multiscale feature fusion Faster R_CNN. Comput Electron Agric, 2022; 199: 107176.
[33] Lu M, Sheng W Q, Zou Y, Chen Y T, Chen Z G. WSS-YOLO: An improved industrial defect detection network for steel surface defects. Measurement, 2024; 236: 115060.
[34] Malik M H, Zhang T, Li H, Zhang M, Shabbir S, Saeed A. Mature tomato fruit detection algorithm based on improved HSV and watershed algorithm. IFAC-PapersOnLine, 2018; 51(17): 431–436.