YOUNG APPLE FRUITS DETECTION METHOD BASED ON IMPROVED YOLOv5
/
基于改进 YOLOv5 的苹果幼果检测方法

Yonghui DU [1], Ang GAO [1], Yuepeng SONG [1,2,3*], Jing GUO[1,2,3], Wei MA[4], Longlong REN [1,2,3,*]
[1] Shandong Agricultural University, College of Mechanical and Electrical Engineering/ China;
[2] Shandong Provincial Key Laboratory of Horticultural Machinery and Equipment/ China;
[3]Shandong Provincial Engineering Laboratory of Agricultural Equipment Intelligence/ China;
[4]Institute of Urban Agriculture, Chinese Academy of Agricultural Sciences/ China
E-mail: renlonglong@sdau.edu.cn
DOI: https://doi.org/10.35633/inmateh-73-07

***Keywords:** YOLOv5; Young apple fruit; Prediction layer; CA attention mechanism; GIOU loss function*

ABSTRACT

The intelligent detection of young apple fruits based on deep learning faced various challenges such as varying scale sizes and colors similar to the background, which increased the risk of misdetection or missed detection. To effectively address these issues, a method for young apple fruit detection based on improved YOLOv5 was proposed in this paper. Firstly, a young apple fruits dataset was established. Subsequently, a prediction layer was added to the detection head of the model, and four layers of CA attention mechanism were integrated into the detection neck (Neck). Additionally, the GIOU function was introduced as the model's loss function to enhance its overall detection performance. The accuracy on the validation dataset reached 94.6%, with an average precision of 82.2%. Compared with YOLOv3, YOLOv4, and the original YOLOv5 detection methods, the accuracy increased by 0.4%, 1.3%, and 4.6% respectively, while the average precision increased by 0.9%, 1.6%, and 1.2% respectively. The experiments demonstrated that the algorithm effectively recognized young apple fruits in complex scenes while meeting real-time detection requirements, providing support for intelligent apple orchard management.

摘要

基于深度学习的苹果幼果智能化检测面临诸多挑战如尺度大小不一、颜色与背景相近等，会导致误检或漏检的风险增加。为了有效解决这些问题，本文提出一种基于改进 YOLOv5 的苹果幼果检测方法，首先建立苹果幼果数据集，再者在检测模型的检测头中添加预测层，在检测脖颈（Neck）中添加四层 CA 注意力机制，并引入 GIOU 函数作为模型的损失函数，以提高模型的整体检测性能。在验证数据集上的准确率达到 94.6%，平均精度为 82.2%；与 YOLOv3、YOLOv4 和原始的 YOLOv5 检测方法相比，准确率分别提升 0.4%、1.3%、4.6%，平均精度分别提升 0.9%、1.6%、1.2%。试验证明，该算法能在满足实时检测要求的前提下，能够有效地识别复杂场景中的苹果幼果，为智能化苹果园管理提供支持。

INTRODUCTION

China, as the world's largest producer and consumer of apples, played a crucial role in rural revitalization and increasing farmers' income (*Jiang et al., 2023*). However, one of the constraints currently faced was the rising labor cost year by year, especially in the apple industry, which has become an important factor restricting its development. In apple production, the management of fruits and vegetables served as a crucial link. At present, it mainly depended on manual labor to complete, but there were problems of high labor intensity and low production efficiency. Therefore, there was an urgent need to improve the intelligent level of apple fruits and vegetables management to cope with the challenges of the current industrial development.

Achieving reliable detection and recognition of young apple fruits is crucial for the intelligence of fruit and vegetable management. In recent years, with the continuous advancement of deep learning technology, object detection techniques have been widely applied in the agricultural sector (*Kamilaris et al., 2018*). In the field of young apple fruit detection, various improved models have emerged successively. Tian et al. optimized the YOLOv3-based model with DenseNet, significantly improving detection speed to just 0.304 seconds per frame (*Tian et al., 2019*). To address the challenge of nighttime fruit detection, Sun et al. optimized the Retina-PVTv2 model by introducing a gradient-coordination mechanism, leading to a significant improvement in nighttime fruit detection accuracy (*Sun et al., 2022*).

Yonghui DU, M.Sc. Stud. Eng.; Ang Gao, M.Sc. Stud. Eng.; Yuepeng Song, Prof. Ph.D. Eng.; Jing GUO Prof. Ph.D. Eng.; Wei MA Prof. Ph.D. Eng.; Longlong Ren Lecturer Ph.D. Eng;

Li et al. focused on addressing the issue of small-sample datasets and proposed an integrated U-Net segmentation model, which significantly improved the accuracy and generalization ability of image segmentation, achieving accurate detection of nighttime young apple fruits (*Li et al., 2021*).

Sun et al. successfully applied the BFP-Net, an FPN layer-optimized model, to the task of detecting young apple fruits, effectively improving detection accuracy and efficiency (*Sun et al., 2022*). Wang et al. proposed an R-FCN young apple fruits detection network capable of detecting small targets with similar background colors (*Wang et al., 2019*). Song et al. achieved young apple fruits recognition by adding non-local modules and squeeze-and-excitation modules to the YOLOv4 model, achieving an average precision of 96.9% (*Song et al., 2021*). Jiang et al. added non-local attention mechanisms and convolutional block attention mechanisms to the YOLOv4 model, enabling efficient recognition under conditions such as high glare, shadows, blurriness, and severe occlusion (*Jiang et al., 2022*). Wang et al. improved the YOLOv5s model, achieving an accuracy of 95.8% in young fruits detection while reducing model load (*Wang et al., 2021*). Additionally, the detection of other fruits such as kiwi, green peach, and citrus bears some similarity to young fruits detection.

The above scholars had detected and achieved good detection results for young apple fruits. However, as far as the detection of young apple fruits was concerned, the problems of smaller target and similar colors to the background were still the focus of research. Therefore, in this paper, a young apple fruits detection method based on improved YOLOv5 was designed to better recognize and distinguish the young fruits from the background and improve the overall detection performance.

MATERIALS AND METHODS

YOLOv5 algorithm

The YOLOv5 algorithm was mainly composed of four parts: the Input, the Backbone, the Neck and the Head (*Park et al., 2023*). The Mosaic data enhancement technology (*Bochkovskiy et al., 2020*) was used on the input side, which was able to randomize the size, cropping and alignment of the data set to achieve more flexible and diverse data processing. The backbone network was the core of YOLOv5s network, which was mainly used to extract image features for target detection (*Arifando et al., 2023*). The neck network was located between the backbone network and the head network, and its function was to extract the features extracted from the backbone network in a deeper way to improve the robustness of the model. The head network was the output part of the target detection model and was responsible for target detection using the previously extracted features. It consists of multiple convolutional and detection layers, where the detection layer mainly implements the process of target detection.

Due to differences in network depth and width, YOLOv5 has undergone several evolutions, resulting in various versions, including YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x. Considering the design requirements of this study and to ensure accuracy, the research was based on the YOLOv5s as the foundational model, with improvements made using the YOLOv5-6.0 version. Figure 1 illustrates the network structure diagram of YOLOv5s.

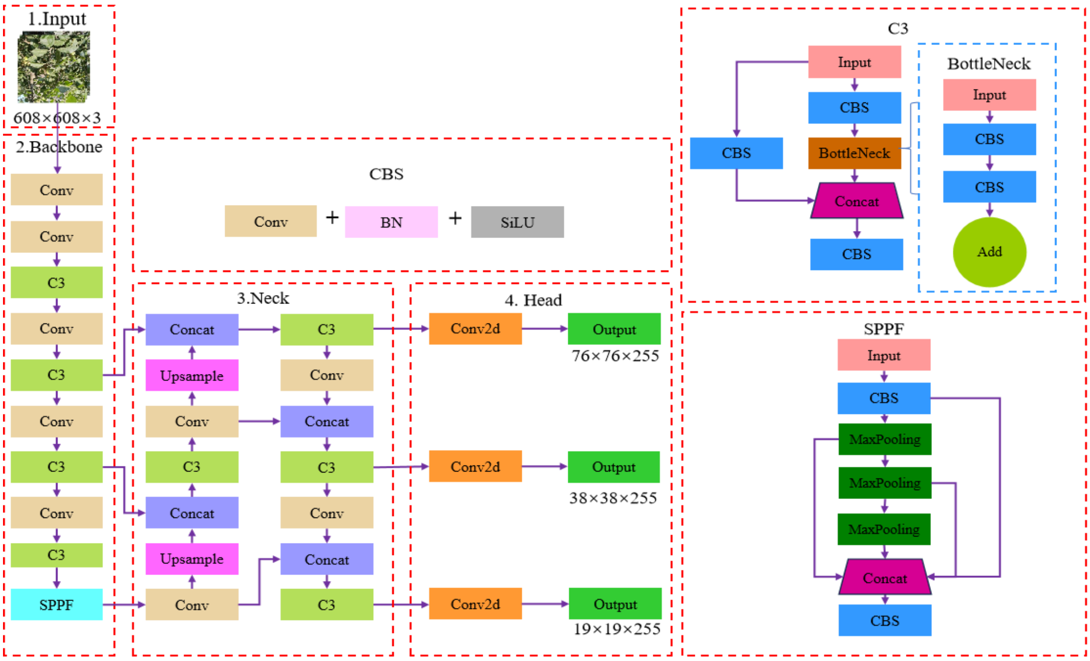


Fig. 1 - Network structure of YOLOv5s

Improved YOLOv5 algorithm

In this paper, the YOLOv5 model was improved on top of it. Firstly, a prediction layer was added to the detection head of the detection model. Secondly, four layers of CA attention mechanism were incorporated into the detection neck (Neck), and the GIOU function was introduced as the model's loss function to enhance the overall detection performance. The network structure of the improved model in this study is illustrated in Figure 2.

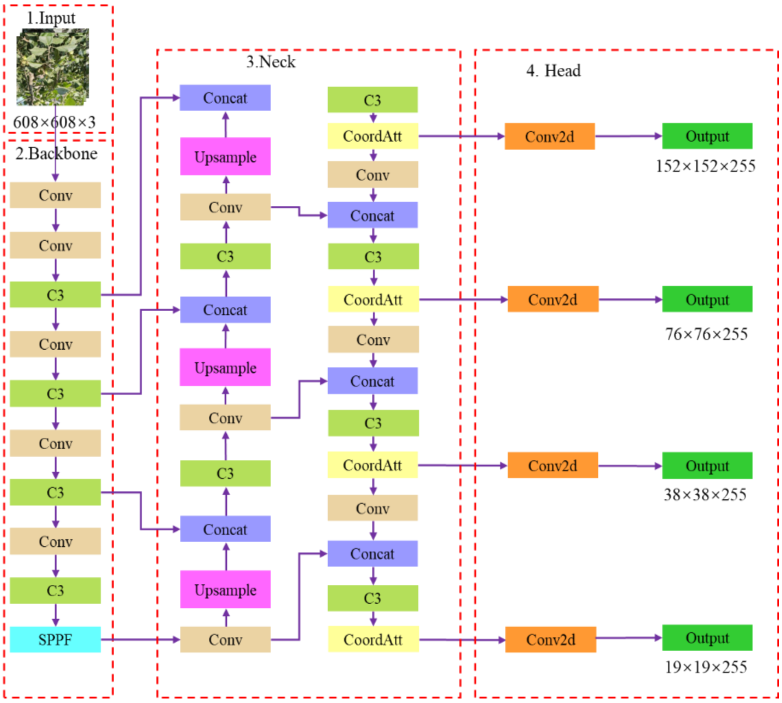


Fig. 2 - Improved network structure

Multi-scale detection head

As could be seen in Figure 1, in the original YOLOv5 model, YOLOv5 detected the target by three different sizes of prediction layers. If the input size was 608 × 608, the three prediction layer sizes of 19 × 19, 38 × 38, and 76 × 76 were obtained at the time of detection (*Carrasco et al., 2021*). As the depth of the input passed to the detection layer gradually increases, the possibility of information loss in the process of passing the features extracted from the shallow layer to the deeper layer gradually increases, placing higher demands on the feature extraction capability of the network structure (*Changgao et al., 2021*). This suggests that the features extracted in the deeper structure through multiple dimensionality reduction operations may not be able to adequately express the information of the image, and thus the design of appropriate multi-scale prediction layers becomes crucial.

The research object of this paper was young apple fruits. Since the size of the near apple fruits was larger relative to the size of the far ones during the shooting, it was often easy to miss the detection problem for the small-scale apple fruits. To solve this problem, this paper used multi-scale detection head to add a shallow prediction layer in the Head network. It was specifically used to detect small-scale young apple fruits. The improved Head network was shown in Figure 3.

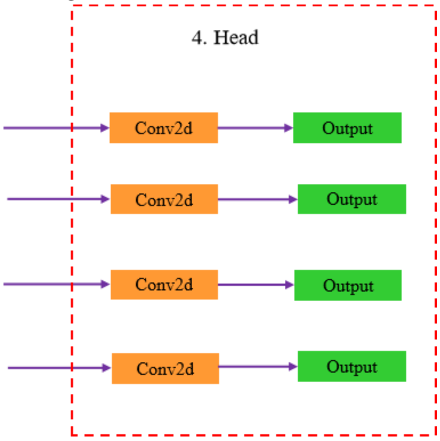


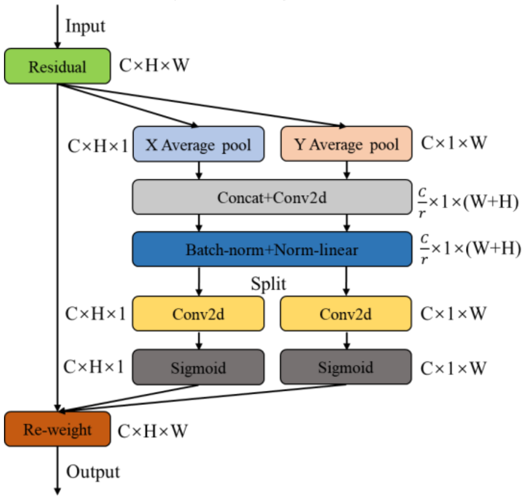
Fig. 3 - Improved Head network

Introduction of CA attention mechanisms

The attention mechanism was a widely used technique in current deep learning, endowing models with the ability to finely select information when processing images. Similar to the attention mechanism in human vision, this technique enables models to focus on key information within vast amounts of data, thereby disregarding unnecessary information (*Aijun et al., 2022*).

At the early stage of the growth of young apple fruits, they might have features that were similar to or not distinct enough from the background, which added difficulties to the detection. The introduction of CA (Coordinate attention) attention mechanism (*Hou et al., 2021*) could enhance the features related to young apple fruits and automatically select the most representative feature channels, thus improving the accuracy of recognition. CA attention mechanism improved the robustness of recognition by assigning different weights to each channel, which enabled the model to adaptively handle these changes (*Pham et al., 2023*). Compared with other complex attention mechanisms, CA attention mechanism had lower computational complexity, which made it more efficient in dealing with real-time or large-scale young apple fruits recognition tasks, and reduced the consumption of computational resources and time. The specific implementation process is shown in Fig.4.

In this study, CA attention mechanism modules were introduced after each C3 module in the Neck network to enhance the model's detection accuracy. The improved Neck structure is depicted in Figure 5.



Note: Input is the input; C is the number of image channels; H is the image height; W is the image width; r is the downsampling reduction ratio; X Average pool is the horizontal average pooling; Y Average pool is the vertical average pooling; Non-linear is the nonlinear activation function; Sigmoid is the activation function; Re-weight is the weight acquisition; Output is the output.

Fig. 4 - Schematic diagram of CA attention mechanism module

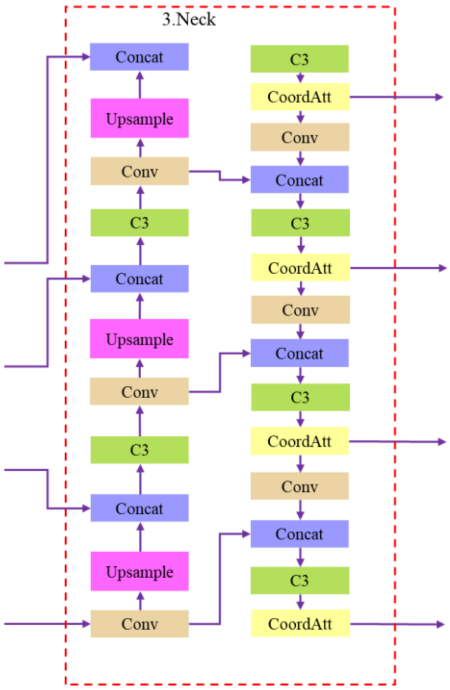


Fig. 5 - Diagram of the improved Neck structure

Optimizing the loss function

In the original YOLOv5 model, the CIOU loss function (*Zheng et al., 2023*) was adopted as the regression loss function of the model. It effectively improved the localization accuracy of the bounding box by integrating the intersection and concurrency ratio (IOU) between the predicted box and the real box, the center point distance, with the aspect ratio *(Aswal et al., 2023, Zhang et al. 2022*). However, in the specific task of young apple fruits recognition, there might be significant variations in the shape and size of young fruits, which makes the CIOU loss function show some limitations in terms of stability. In contrast, the GIOU loss function (*Rezatofighi et al., 2019*) had higher stability when dealing with targets with large variations in shape and size, which helped to improve the performance of the model in the task of young apple fruits recognition. Meanwhile, the GIOU loss function was able to consider the spatial relationship between the bounding boxes more comprehensively, which helped the model to locate the position of the young apple fruits more accurately. Therefore, in order to improve the stability of the model, the GIOU loss function was used instead of the CIOU loss function in this paper.

The GIOU loss function makes the predicted bounding box closer to the real bounding box by introducing the area of the smallest outer rectangle as a penalty term (*Rani, 2021*). This design makes the GIOU loss function perform stably in solving the problem of the degree of bounding box overlap, regardless of whether the two overlap or not. The GIOU loss function model is shown in Figure 6.

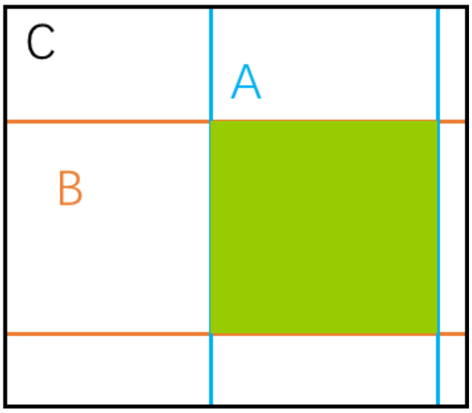


Fig. 6 - GIOU loss function model

$$GIOU = IOU - \frac{\left|C - |A \cup B|\right|}{|C|} \tag{1}$$

$$IOU = \frac{|A \cap B|}{|A \cup B|} \tag{2}$$

where: *A* is the area of the prediction frame; *B* is the area of the real frame; *C* is the area of the smallest rectangle containing the prediction frame *A* and the real frame *B*.

RESULTS

Data set preprocessing and evaluation indicators

The image data used in this study were sourced from the apple experimental field at Shandong Agricultural University, located at approximately 117.12296° E longitude and 36.200713° N latitude. The data were collected in early May 2023 to ensure their timeliness. All images in the datasets were uniformly formatted as .jpg files with dimensions of 608×608 pixels for ease of subsequent processing. The dataset was annotated using labelimg software to ensure data quality. In total, there were 7536 images in the dataset, which were divided into training, validation, and testing sets in an 8:1:1 ratio for model training and performance evaluation purposes.

To validate the effectiveness of the algorithm proposed in this paper, performance evaluation of the algorithm was conducted using metrics such as Precision (P), Recall (R), Average Precision (AP), and Mean Average Precision (mAP) (*Jubayer et al., 2021,* Ye *et al., 2023*). The following are the formulas used to calculate these parameters:

$$P = \frac{TP}{TP+FP} \tag{3}$$

$$R = \frac{TP}{TP+FN} \tag{4}$$

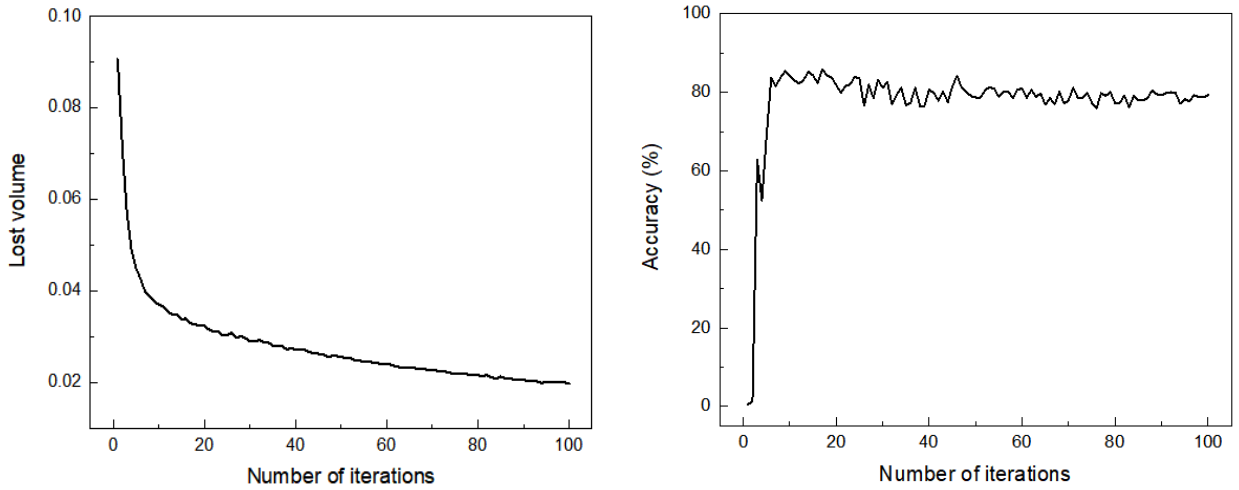
$$AP = \int_0^1 P(R)dR \tag{5}$$

$$mAP = \frac{1}{K}\sum_{i=1}^{K} AP(i) \quad (6)$$

where: P denotes accuracy, %; R denotes recall, %; AP denotes average precision, %; mAP denotes mean average precision, %; TP denotes the number of correctly detected young apple fruits, pcs; FP denotes the number of wrongly detected young apple fruits, pcs; FN denotes the number of missed young apple fruits, pcs; TN denotes the number of detected backgrounds, pcs; K is the number of categories of young apple fruits in the dataset, pcs.

Training of young apple fruits model

By adding a prediction layer, CA attention mechanism, and optimizing the GIOU loss function to the YOLOv5s model, the accuracy of the model in identifying young apple fruits was improved. To validate the proposed detection method, loss curve and accuracy curve of the improved model were plotted. Figure 7(a) depicted the loss curve during model training. After multiple iterations of training, the loss curve exhibited a noticeable decreasing trend, stabilizing around 0.025 with increasing iterations and gradually leveling off. Figure 7(b) illustrated the accuracy curve during model training. At the beginning of training, the recognition accuracy was relatively low. However, with an increase in the number of iterations, the accuracy gradually improved and stabilized around 80%. This clearly demonstrated the effectiveness of model training.



a) Model Training Loss Lost Plot; (b) Model training precision graph

Fig. 7 - Model training Loss, accuracy plot

Improvement program for young apple fruits

In the same experimental environment, eight tests of the improvement scheme were conducted to evaluate the impact of each improvement module in the YOLOv5 model improved in this paper on the performance of young apple fruit detection. The experimental results were shown in Table 1, where "√" indicates the use of the corresponding module.

Table 1

Test results

Test number	Predictive layer	CA	GIOU	Accuracy (%)	Mean average precision (%)
(1)				90.0	81.0
(2)	√			90.2	82.6
(3)		√		89.9	82.2
(4)			√	92.2	80.9
(5)	√	√		93.2	82.4
(6)	√		√	92.1	81.4
(7)		√	√	93.8	82.5
(8)	√	√	√	94.6	82.2

Trial (1) was the trial of the original model, which served as the baseline for the following seven sets of trial comparisons. Trial (2) was the trial with only the prediction layer added, which improved accuracy by 0.2% and mean average precision by 1.6%. Trial (3) was the trial with only the CA attention mechanism added, which showed a 0.1% decrease in accuracy but a 1.2% improvement in mean average precision.

Trial (4) was the trial with only the GIOU loss function introduced, and the accuracy was improved by 2.2%, but the mean average precision decreased by 0.1%. Trial (5) was a trial that added both the prediction layer and the CA attention mechanism, with the accuracy improved by 3.2% and the mean average precision improved by 1.4%. Trial (6) was a trial with the addition of both the prediction layer and the introduction of the GIOU loss function, with the accuracy improved by 2.1% and the mean average precision improved by 0.4%. Trial (7) was the trial of adding CA attention mechanism and introducing GIOU loss function at the same time, the accuracy was improved by 3.8%, and the mean average precision was improved by 1.5%. Trial (8) was the trial of the improved method in this paper, the accuracy was improved by 4.6%, and the mean average precision was improved by 1.2%. After the ablation test, it was found that all the improvement points and their synergies with each other positively affected the model and enhanced the detection of young apple fruits. This fully proved the effectiveness and correctness of the improvement method proposed in this paper, making it advantageous in the detection task.

Comparison tests of young apple fruits based on improved YOLOv5

Classical detection methods including YOLOv3, YOLOv4, and YOLOv5 were selected for young apple fruits detection and compared with the detection method proposed in this paper to validate its superiority in detecting young apple fruits. The detection results are presented in Table 2.

Table 2

Comparative test results

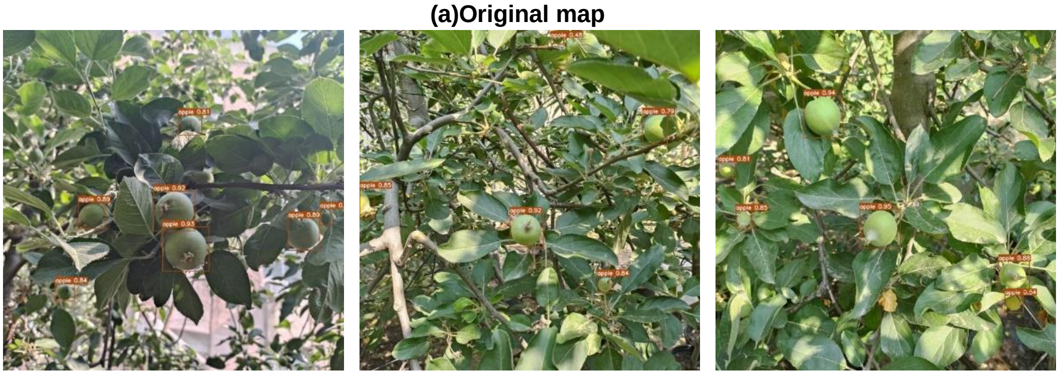
Detection Models	Accuracy (%)	Mean average precision (%)
YOLOv3	94.2	81.3
YOLOv4	92.9	80.6
YOLOv5	90.0	81.0
Methodology of this paper	94.6	82.2

Based on the comparison results in Table 2, it could be seen that the proposed detection method in this paper improved the accuracy by 0.4% and the mean averaged accuracy by 0.9% over the YOLOv3 model. Compared with the YOLOv4 model, the accuracy was improved by 1.3% and the mean average precision was improved by 1.6%. Compared with the original YOLOv5 model, the accuracy was improved by 4.6% and the mean average precision was improved by 1.2%. In each parameter index, the model in this paper outperforms the accuracy and mean average accuracy of the YOLOv3, YOLOv4 and YOLOv5 models.

Validation analysis of apple young fruits visualization based on improved YOLOv5

Using the same set of test images, YOLOv3, YOLOv4, YOLOv5, and the model proposed in this paper were tested and visually analyzed, as shown in Figure 8.

(a)Original map



(b) YOLOv5 Detection Method

(c) YOLOv3 Detection Method

(d) YOLOv4 Detection Methods

(e) The detection methods in this paper

Fig. 8 - Comparison of the detection effect of different methods on young apple fruits

In terms of the actual detection effect, the method in this paper showed better detection effect in the detection of young apple fruits at different distances and with leaves as the background. In contrast, YOLOv3, YOLOv4 and YOLOv5 had some degree of missed detection and false detection. The method in this paper could more accurately detect all the young apple fruits and showed higher detection accuracy. In conclusion, the method in this paper could significantly improve the effectiveness of feature extraction, effectively reduce the loss of detail information, and it improved the detection accuracy of young apple fruits.

CONCLUSIONS

To address current challenges in young apple fruits detection, such as significant variations in target sizes and diverse detection environments, in this paper, on the basis of YOLOv5 model, the prediction layer and CA attention mechanism were introduced on, and the GIOU loss function was used. It was proved through experiments that compared with the traditional YOLOv5 algorithm, the improved method proposed in this paper increased the prediction accuracy by 3.6% and the mean average accuracy by 1.2%. The method could better realize the detection of young apple fruits in natural scenes.

Although the current method has achieved certain results in young apple fruits detection, there are still some limitations, such as the single variety and lack of wide applicability. To overcome these limitations, future work will focus on further improvements, including collecting data from more varieties, optimizing network models, and addressing detection challenges in complex environments.

ACKNOWLEDGEMENT

We are very grateful to all the authors for their support and contribution with the manuscript. This work is partly supported by: Innovation Team Fund for the Fruit Industry of Shandong's Modern Agricultural Technology System (SDAIT-06-12), the State Key Laboratory of Mechanical System and Vibration (Grant No. MSV202002), and the Modern Agriculture Coarse Grain Industry Technology System Project of Shandong Province (SDAIT-15-05).

REFERENCES

[1] Aijun, G., Gao, A.N.G., Chunming, Y., Zhilong, Z., Zhang, J. I., & Jinglong, Z. (2022). Dropping ear detection method for corn harvester based on improved Mask-RCNN. *INMATEH-Agricultural Engineering*, 66(1), pp.31-40. Shandong/China. https://doi.org/10.35633/inmateh-66-03

[2] Arifando, R., Eto, S., & Wada, C. (2023). Improved YOLOv5-based lightweight object detection algorithm for people with visual impairment to detect buses. *Applied Sciences, 13*(9), 5802, Kyushu/Japan.

[3] Aswal, D., Shukla, P., & Nandi, G. C. (2021). Designing effective power law-based loss function for faster and better bounding box regression. *Machine Vision and Applications*, 32(4), 87, India.

[4] Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). Yolov4: Optimal speed and accuracy of object detection. *arxiv preprint arxiv:2004.10934, Russia.*

[5] Carrasco, D. P., Rashwan, H. A., García, M. Á., & Puig, D. (2021). T-YOLO: Tiny vehicle detection based on YOLO and multi-scale convolutional neural networks. *IEEE Access*, 11, 22430-22440, Spain.

[6] Changgao, Xia., Wanlei, Ni, Kun Jiang, & Xiaofan, Li (2022). A real-time detection model for identification of citrus during different growth stages in orchards. *INMATEH-Agricultural Engineering*, 68(3), pp.373-382. Jiangsu/China. https://doi.org/10.35633/inmateh-68-37

[7] Fan, X.Y., Fan, F. Kao, J.F. Yu, C.K. Hua, Jiang, Y., & Qian, C.H., (2023). Research on visual detection technology for liquid crystal panel electrode defect by improved YOLOv7 (基于改进 YOLOv7 的液晶面板电极缺陷视觉检测技术研究). *Journal of Electronic Measurement and Instrumentation* (09), 225-233. doi:10.13382/j.jemi.B2306576, Jiangsu/China.

[8] Hou, Q., Zhou, D., & Feng, J., (2021). Coordinate attention for efficient mobile network design. *In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. pp. 13713-13722, Singaporean.

[9] Jiang, G.Q., Yang, Z.Y., Huo, Z.Q., Luo, J.W., & Zhao, C.J., (2023). Apple detection method before thinning fruit based on improved YOLOv5 model (基于改进 YOLOv5 网络的疏果前苹果检测方法). *Jiangsu Agricultural Science* (14), 205-215. doi:10.15889/j.issn.1002-1302.2023, 14.028, Hebei/China.

[10] Jiang, M., Song, L., Wang, Y., Li, Z., & Song, H., (2022). Fusion of the YOLOv4 network model and visual attention mechanism to detect low-quality young apples in a complex environment. *Precision Agriculture*, 1-19, Shaanxi/China.

[11] Jubayer, F., Soeb, J. A., Mojumder, A. N., Paul, M. K., Barua, P., Kayshar, S., ... & Islam, A. (2021). Detection of mold on the food surface using YOLOv5. *Current Research in Food Science*, 4, 724-728, Bangladesh.

[12] Kamilaris, A., & Prenafeta-Boldú, F.X., (2018). Deep learning in agriculture: A survey. *Computers and electronics in agriculture*, 147, 70-90, Spain.

[13] Li, Q., Jia, W., Sun, M., Hou, S., & Zheng, Y., (2021). A novel green apple segmentation algorithm based on ensemble U-Net under complex orchard environment. *Computers and Electronics in Agriculture*, 180, 105900, Shandong/China.

[14] Pham, T. N., Nguyen, V. H., & Huh, J. H. (2023). Integration of improved YOLOv5 for face mask detector and auto-labeling to generate dataset for fighting against COVID-19. *The Journal of Supercomputing,* 79(8), 8966-8992, Korea.

[15] Park, J., Bae, J., Im, J., Kim, B., & Jeong, J. (2023). LED-Display Defect Detection based on YOLOv5 and Transformer. *IEEE Access*, Korea.

[16] Rani, E. (2021). LittleYOLO-SPP: A delicate real-time vehicle detection algorithm. *Optik*, 225, 165818, India.

[17] Rezatofighi, H., Tsoi, N., Gwak, J., Sadeghian, A., Reid, I., & Savarese, S. (2019). Generalized intersection over union: A metric and a loss for bounding box regression. *In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 658-666), Australia.

[18] Song, H. B., Jiang, M. Mei, Wang, Y. F., & Song, L., (2021). Efficient detection method for young apples based on the fusion of convolutional neural network and visual attention mechanism (融合卷积神经网络与视觉注意机制的苹果幼果高效检测方法). *Journal of Agricultural Engineering* (09), 297-303, Shaanxi/China.
[19] Sun, M., Xu, L., Chen, X., Ji, Z., Zheng, Y., & Jia, W., (2022). Bfp net: balanced feature pyramid network for small apple detection in complex orchard environment. *Plant Phenomics*, Shandong/China.
[20] Sun, M., Xu, L., Luo, R., Lu, Y., & Jia, W., (2022). GH Former-Net: Towards more accurate small green apple/begonia fruit detection in the nighttime. *Journal of King Saud University-Computer and Information Sciences*, 34(7), 4421-4432, Shandong/China.
[21] Tian, Y., Yang, G., Wang, Z., Wang, H., Li, E., & Liang, Z., (2019). Apple detection during different growth stages in orchards using the improved YOLO-V3 model. *Computers and electronics in agriculture*, 157, 417-426, Beijing/China.
[22] Wang, D., & He, D., (2021). Channel pruned YOLO V5s-based deep learning approach for rapid and accurate apple fruitlet detection before fruit thinning. *Biosystems Engineering*, 210, 271-281, Shaanxi/China.
[23] Wang, D., & He, D., (2019). Recognition of apple targets before fruits thinning by robot based on R-FCN deep convolution neural network (基于 R-FCN 深度卷积神经网络的机器人疏果前苹果目标的识别). *Journal of Agricultural Engineering* (03), 156-163, Shaanxi/China.
[24] Wanzhi Zhang, Yuelin Han, Chen Huang, & Zhiwei Chen. (2022). Recognition method for seed potato buds based on improved YOLOv3-tiny. *INMATEH-Agricultural Engineering,* 67(2), Shandong/China.
[25] Ye Xia, Xiaohui Lei, Herbst, A., & Xiaolan LYU. (2023). Research on pear inflorescence recognition based on fusion attention mechanism with Yolov5. *INMATEH-Agricultural Engineering,* 69(1), pp.11-20. Nanjing/China. https://doi.org/10.35633/inmateh-69-01
[26] Zhang X. H., Yan J. X., Ma B., Ju J. S., Shen Q. F., & Wu Y. J., (2022). Research on abnormal detection method of side guard based on improved YOLOv5s (基于改进 YOLOv5s 的护帮板异常检测方法研究). *Journal of Engineering Design* (06), 665-675, Shaanxi/China.
[27] Zheng, Z., Wang, P., Liu, W., Li, J., Ye, R., & Ren, D. (2020, April). Distance-IoU loss: Faster and better learning for bounding box regression. *In Proceedings of the AAAI conference on artificial intelligence* Vol. 34, No. 07, pp. 12993-13000, Tianjin/China.