2025 5th International Conference on Artificial Intelligence and Industrial Technology Applications (AIITA)

Tea Pest and Disease Detection Algorithm Based on Deep Learning

1st Chengtao Liu*
School of Electronic and Information
Xi'an Polytechnic University
Xi'an, Shaanxi ,China
20150706@xpu.edu.cn

2nd Jiahao Cai
School of Electronic and Information
Xi'an Polytechnic University
Xi'an, Shaanxi ,China
16492560@qq.com

3rd Zhuoxi Wu
School of Electronic and Information
Xi'an Polytechnic University
Xi'an, Shaanxi,China
wuzhuoxi@xpu.edu.cn

4th Jingbing Yang
School of Information Engineering
Yangling Vocational &Technical College
YangLing, China
13619259015@163.com

***Abstract*—Tea cultivation faces significant challenges from various pests and diseases that directly impact both yield and quality. Traditional manual detection methods are limited by their inefficiency and subjective factors such as inspector experience and expertise. To address these limitations, this paper proposes two innovative improvements to the traditional YOLOv5 target detection algorithm: the integration of Multidimensional Collaborative Attention (MCA) and Structured Intersection over Union (SIoU). The MCA mechanism enhances feature extraction capabilities across multiple scales and directions, while the SIoU loss function optimizes boundary regression by considering directional relationships between predicted and actual bounding boxes. Experimental results demonstrate that our improved algorithm achieves superior performance compared to the classical YOLOv5, with increased precision (96.6%), recall (95.8%), and mAP (69.8%). This enhanced detection system provides a more efficient and accurate solution for tea pest and disease identification, offering significant practical value for tea farmers in implementing timely and effective pest control measures.**

***Keywords*—tea, diseases and pests, Yolo v5, MCA, SIoU**

I. INTRODUCTION

Tea industry encompasses planting, processing, sales and other links, forming a huge industrial cluster. During tea cultivation, plants often face threats from various pests and diseases, which not only affect the growth and development of tea plants but also lead to quality deterioration and economic losses. Therefore, addressing pest and disease problems in tea cultivation is of great importance.

Currently many researchers and scholars have achieved some results in tea testing using modern technology. Wang et al. developed a CNN-based tea pest identification model integrated with a MySQL pest control database [1]. Similarly, Huang et al. optimized impurity frame clustering using the K-Means algorithm to determine more suitable anchor frame sizes for tea impurity characteristics [2]. In another study, Wang et al. employed the YOLOv5 model to estimate tea production based on seasonal changes and bud detection across different tea varieties [3]. Furthermore, Ren et al. established a comprehensive tea grade quality evaluation method [4], while Jin et al. combined improved YOLOX-small and PSP-net models for tea shoot and picking area detection [5]. Ding et al. applied transfer learning strategies with pre-trained models to identify tea tree variety morphology [6]. Additionally, Huang et al. enhanced detection performance for fuzzy, occluded, and diseased leaflets by integrating Retinex algorithm with Faster R-CNN [7].

Although the above studies have achieved some results, there are fewer studies on tea pest and disease detection. Therefore, this paper conducts some research on tea pest and disease detection, and the experiments show that the improved algorithm can further optimize the target frame regression process of the model and improve the accuracy and recall of detection. In addition, the effectiveness and feasibility of the deep learning-based tea pest and disease detection algorithm designed in this paper are further verified by designing the system interface.

II. THEORETICAL BASIS

Yolo v5 model has the lowest number of parameters among similar models and is widely used in practical scenarios. Based on these advantages, Yolo v5 is selected as the basic model for tea pest detection. Various attention mechanisms can be implemented in the Yolo v5 model structure to enhance its performance. The integration of attention mechanisms enables the model to process information more selectively and efficiently, thereby improving detection accuracy. The implementation process requires careful consideration of the model architecture and strategic selection of appropriate attention mechanisms. It is crucial to maintain a balance between model enhancement and stability, avoiding potential performance degradation from excessive use of attention mechanisms.

A. MCA Attention Mechanism

The structure of the MCA attention mechanism is shown in Fig.1, by adjusting the E2, E3, and E4 feature maps to the same resolution, followed by their concatenation to form a multi-scale information feature map. A 1x1 convolution is applied to reduce channel numbers, thereby reducing computational demands. Next, cross-axis transformations are applied with substitutions in both horizontal and vertical directions to enhance directional sensitivity. The processed feature maps are summarized and merged with E1, and the model's output is obtained through convolution. Multi-scale horizontal and vertical convolutions focus on different directional features, using cross-processing to ensure comprehensive information capture. The MCA structure enables the network to efficiently extract multi-scale and multi-directional features with low computational cost, improving the model's overall performance.

Special Research Project of Shaanxi Department of Education (23JK0467), Science and technology plan project of Shaanxi Provincial Science and Technology Department(2024JC-YBQN-0678). Shaanxi Provincial Science and Technology Department (2023-JC-QN-0462).

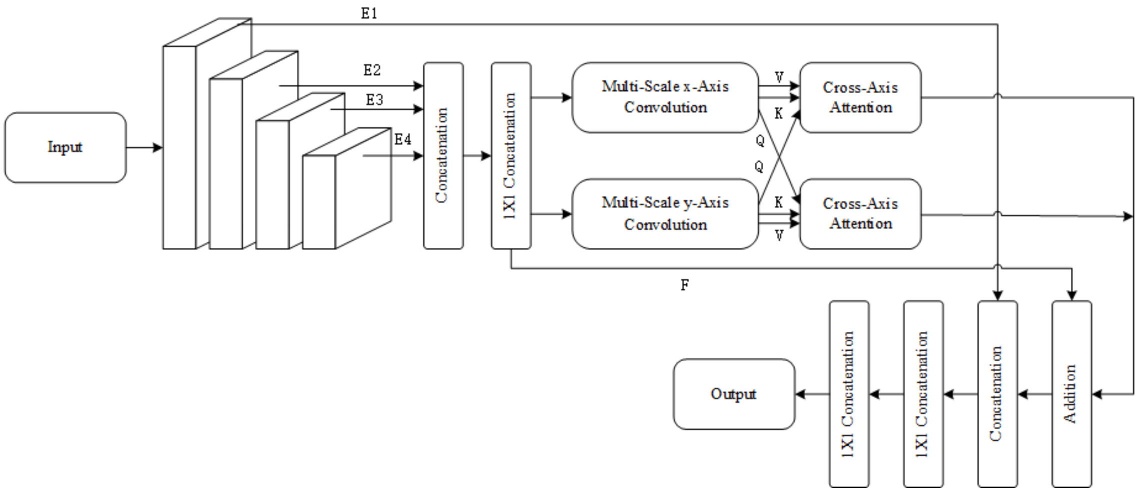


Fig. 1 MCA attention mechanism structure Diagram

In this study, an MCA Gate module is defined under the models\common.py file, and the specific block diagram of the algorithm is shown in Fig.2.

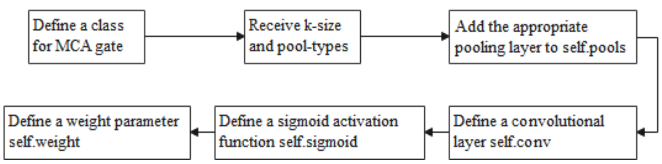


Fig. 2 Block diagram of MCA gate module definition

Fig. 3 presents a detailed diagram of the MCA Attention mechanism. The input feature maps first enter the paths in both the x and y directions. These maps are fused after convolutions of different sizes are applied. Then, attention is computed across the axes through Q-swapping of the x and y axes. Finally, the output feature maps are obtained.

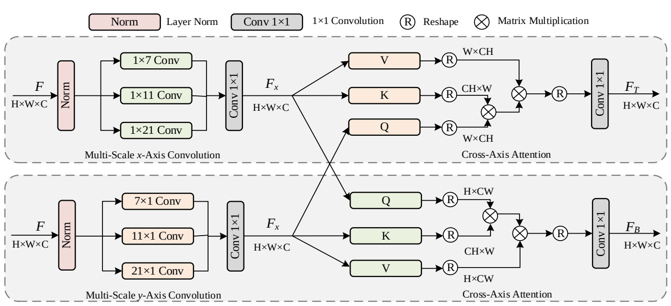
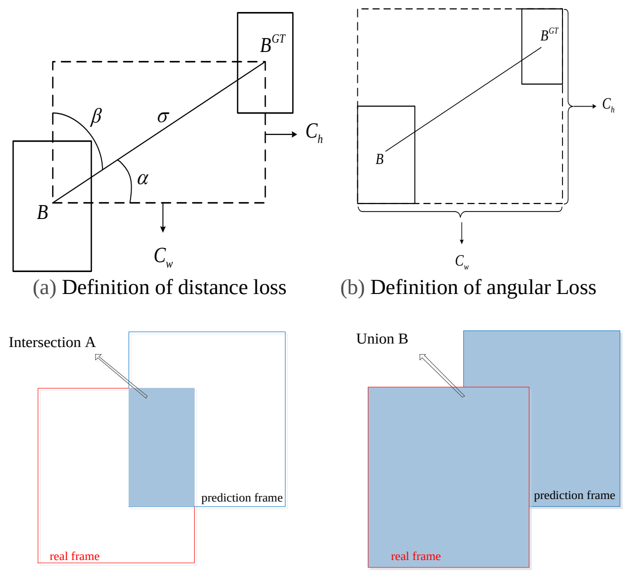


Fig. 3 Detailed diagram of the MCA attention mechanism

B. *Border regression loss function SIoU*

The common loss function does not consider the directional relationship between the real frame and the predicted frame, which leads to slow convergence. To address this issue, the SIoU function introduces the vector angle between the real frame and the predicted frame and redefines the associated loss function.



(a) Definition of distance loss (b) Definition of angular Loss

(c) Definition of Iou loss

Fig.4 Some definitions of the marginal regression loss function

(1) Angular loss is defined as shown in Fig. 4(a).

$$\Lambda = 1-2*\sin^2(\arcsin(\frac{c_h}{\sigma})-\frac{\pi}{4}) = \cos(2*(\arcsin(\frac{c_h}{\sigma})-\frac{\pi}{4})) \quad (1)$$

where c_h is the height difference between the center points of the real frame and the predicted frame, and σ is the distance between the center points of the real frame and the predicted frame, which is $\arcsin(\frac{c_h}{\sigma})$ in fact equal to the angle α.

$$\frac{c_h}{\sigma} = \sin(\alpha) \quad (2)$$

$$\sigma = \sqrt{(b_{c_x}^{gt}-b_{c_x})^2+(b_{c_y}^{gt}-b_{c_y})^2} \quad (3)$$

$$c_h = \max(b_{c^y}^{gt}, b_{c^y}) - \min(b_{c^y}^{gt}, b_{c^y}) \quad (4)$$

$(b_{c^x}^{gt}, b_{c^y}^{gt})$ are the coordinates of the center of the real frame and (b_{c^x}, b_{c^y}) are the coordinates of the center of the predicted frame.It may be noted that when α is $\frac{\pi}{2}$ or 0, the angular loss is 0, During training minimiz α if $\alpha < \frac{\pi}{4}$, otherwise minimize β.

(1) The distance loss is defined as shown in Fig. 4(b).

$$\Delta = \sum_{t=x,y}(1-e^{-\gamma\rho_t}) = 2 - e^{-\gamma\rho_x} - e^{-\gamma\rho_y} \tag{5}$$

$$\rho_x = \left(\frac{b_{c^x}^{gt} - b_{c^x}}{c_w}\right)^2, \rho_y = \left(\frac{b_{c^y}^{gt} - b_{c^y}}{c_h}\right)^2 \quad \gamma = 2 - \Lambda \tag{6}$$

Here (c_w, c_h) is the width and height of the smallest outer rectangle of the real and predicted boxes.

(2) Shape loss, defined as follows:

$$\Omega = \sum_{t=W,h}(1-e^{-W_t})^{\theta} = (1-e^{-W_w})^{\theta} + (1-e^{-W_h})^{\theta} \tag{7}$$

$$W_w = \frac{|W - W^{gt}|}{\max(W, W^{gt})}, W_h = \frac{|h - h^{gt}|}{\max(h, h^{gt})} \tag{8}$$

(W, h) and (W^{gt}, h^{gt}) are the width and height of the predicted and real frames, respectively, and θ is the degree of attention paid to shape loss by the control.

(3) The IoU loss is defined as shown in Fig. 4(c).

$$IoU = \frac{\text{intersection A}}{\text{union B}} \tag{9}$$

In summary, the SIoU Loss loss function is defined as follows:

$$Loss_{SIoU} = 1 - IoU + \frac{\Delta + \Omega}{2} \tag{10}$$

III. Experimentation and Analysis

This part first uses the classical Yolo v5 algorithm to train tea containing pests and diseases, and explains the relevant parameters of the operation results.Then the algorithm is improved by adding MCA attention mechanism and SIoU loss function, and the experimental results are compared.As shown in Fig.5, set the epoch of the training round to 100 and run train.py to get the training result graph of the classic Yolo v5 algorithm.

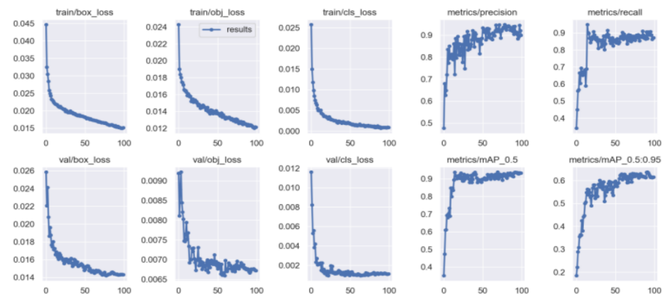


Fig. 5 Classic Yolo v5 training results

The meanings of each data indicator in Fig. 5 are as follows:

(1) train/box_loss refers to the index used to calculate the boundary frame loss of objects in the object detection model.Bounding box is usually used in object detection tasks to represent the position and size of objects in the predicted image.

(2) train/obj-loss is a loss function used to measure the accuracy of the model's prediction of the existence of object in object detection model.

(3) train/cls_loss is a loss function used to measure the performance of the model in classification tasks in the object detection model.

(4) Precision looks at how many instances predicted to be positive samples are actually positive samples.

(5) metrics/recall measures the proportion of positive examples correctly identified by the model among all samples that are actually positive examples.

(6) val/box_loss in Yolo v5 represents the position loss of each prediction box on the validation set.

(7) val/obj-loss represents the object existential loss on the validation set in the object detection model.

(8) val/cls_loss in object detection models typically represents a classification loss on a validation set.

(9) mAP_0.5 refers to the average accuracy when the IOU threshold is 0.5.

(10) mAP_0.5:0.95 indicates that when the IOU threshold changes from 0.5 to 0.95, the average value of maps corresponding to each threshold is taken.

Based on the classic Yolo v5, add MCA attention mechanism for training. Specifically, after adding the new module to the parsing module in the models\ Yol. py file, set the epoch of the training round to 100 and run train.py. The training results are shown in Fig. 6.

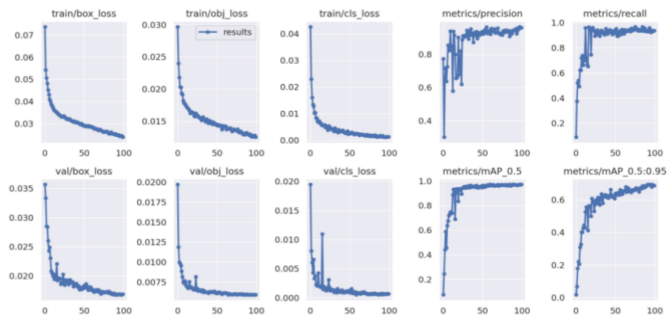


Fig. 6 Added MCA attention mechanism training results

On the basis of the above algorithm, SIoU loss function is added for training. Specifically, set the epoch of the training round to 100, run train.py, and get the training result as shown in Fig. 7.

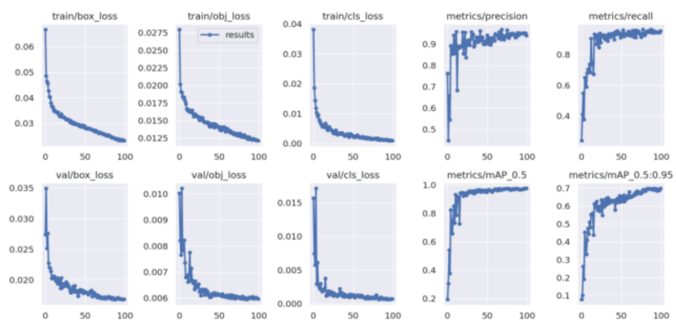
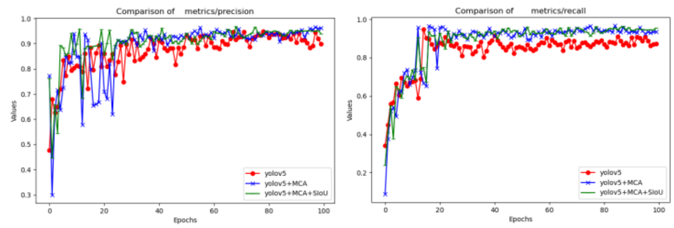


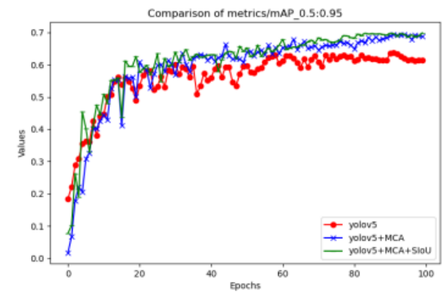
Fig.7 Training results of adding SIoU loss function

The comparison graphs of the three algorithms are shown in Fig. 8. In the subgraphs of Fig. 8, the red dotted line represents the results calculated by the classical Yolo v5 algorithm, the blue dashed line represents the results with the addition of the MCA attention mechanism, and the green solid line represents the results after introducing the SIoU loss function in conjunction with the MCA attention mechanism.

Fig.8(a) compares the training accuracy of the three algorithms, where the horizontal axis indicates the number of training iterations and the vertical axis represents the training accuracy values. Observing Fig.8(a), it is evident that the addition of the MCA attention mechanism enhances training precision, while the introduction of the SIoU loss function reduces the number of iterations needed to achieve better precision, indicating improved algorithm efficiency.Fig.8(b) presents the training recall comparison, with the same axes as Fig.8(a). It shows that the MCA attention mechanism increases training recall, and the SIoU loss function contributes to algorithm stability as the number of training iterations increases, further indicating enhanced efficiency. Fig.8(c) illustrates the comparison of mAP_0.5:0.95 values. Again, the horizontal axis represents the number of training iterations, and the vertical axis shows the mAP_0.5:0.95 values. The addition of the MCA attention mechanism improves the mAP values, and the SIoU loss function allows the algorithm to achieve better results with fewer iterations, thus improving detection efficiency.



(a) Comparison plot of training accuracy (b) Recall comparison plot



(c) mAP_0.5:0.95 comparison plot

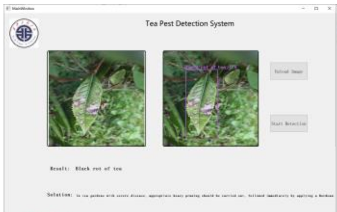
Fig. 8 Comparison of the three algorithms

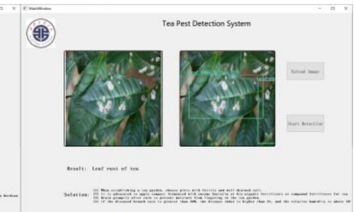
Table 1 displays the data comparison for algorithm improvements.By examining Table 1, it is evident that parameters such as precision, recall, and mAP_0.5:0.95 show improvement when using the classical Yolo v5 algorithm enhanced with the MCA attention mechanism and the SIoU loss function.

TABLE1 COMPARISON OF ALGORITHMIC IMPROVEMENT DATA

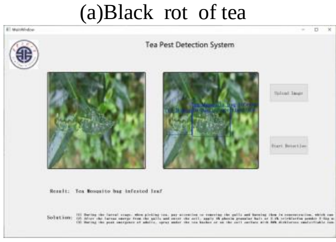
arithmetic	***accurate***	***Recal rate***	***mAP_0.5:0.95***
Classic Yolo v5	0.89969	0.86995	0.61455
Classic Yolo v5 + MCA	0.96147	0.93562	0.68756
Classic Yolo v5+ MCA +SIoU	0.96603	0.95869	0.69855

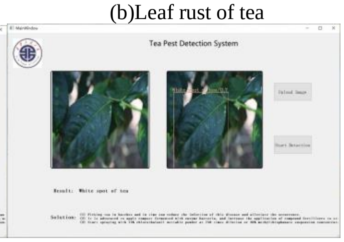
In this study, we selected a publicly available image dataset of tea pests from the Kaggle platform after an extensive search of relevant academic resources. The dataset includes six categories: Black Rot of Tea, Leaf Rust of Tea, Red Spider Infested Tea Leaf, Tea Mosquito Bug Infested Leaf, Tea Leaf, and White Spot of Tea. Each category is named according to the specific disease it represents. The dataset consists of over 5,400 images, with more than 4,700 used for training, over 270 for validation, and the remainder for testing to evaluate the model's performance. Testing showed that the improved algorithm can quickly recognize and detect six different types of tea pests and diseases, providing corresponding solutions based on the identified pests and diseases. The detection results are illustrated in Fig. 9.



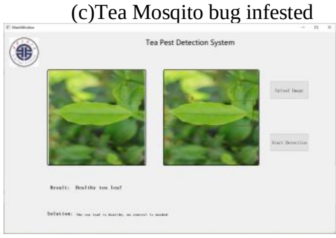


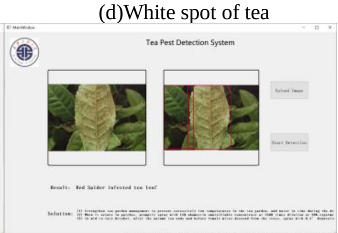
(a)Black rot of tea (b)Leaf rust of tea





(c)Tea Mosqito bug infested (d)White spot of tea





(e)Healthy tea (f)Red spider infested tea

Fig. 9 Results of tea pest and disease

IV. CONCLUSION

In this paper, we address the shortcomings of current tea pest detection methods by improving the Yolo v5 algorithm. The introduction of the MCA attention mechanism enhances the precision, recall rate, and mAP value of the deep learning-based tea pest detection algorithm on the dataset, thus making the algorithm more accurate. Additionally, the integration of the SIoU loss function allows the algorithm to achieve better training results with fewer iterations, improving the training speed. Testing demonstrates that the improved algorithm can accurately identify and classify tea pests and diseases. This research holds practical significance as it enhances detection efficiency and reduces labor costs compared to traditional manual detection methods, thereby promoting the development of the tea economy.

REFERENCES

[1] XiaoPan W, ZanLi Z, DaCang W. Design of The Intelligent Identification APP for Tea Pests Basedon Deep Learning[J]. Science & Technology Vision, 2020(06): 226-228.

[2] Shaohua H , Xifeng L. Detecting the impurities in tea using an improved YOLOv5 model[J].Transactions of the Chinese Society of Agricultural Engineering, 2022, 38(17): 329-336.

[3] Shumao W , Cuiping Yu , JunHui Ma, et al.Tea yield estimation using UAV images and deep learning[J]. Industrial Crops Products, 2024, (12): 212118358-.

[4] Guangxin R , Rui W, Lingling Y, et al. Description of tea quality using deep learning and multi-sensor feature fusion[J]. Journal of Food Composition and Analysis, 2024, (126): 105924-.

[5] Meng J, Wang Y, Zhang J, et al.Tea Bud and Picking Point Detection Based on Deep Learning[J]. Forests, 2023, 4(6).

[6] Ding Y, Huang H, Cui H, et al. A Non-Destructive Method for Identification of Tea Plant Cultivars Based on Deep Learning[J]. Forests, 2023, 14(4).

[7] Gensheng H, Huaiyu W, Yan Z, et al. Detection and severity analysis of tea leaf blight based on deep learning[J]. Computers and Electrical Engineering, 2021, (90):107023-.