Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 9, No. 3, 1777-1794
2025
Publisher: Learning Gate
DOI: 10.55214/25768484.v9i3.5690

A diseases and pests identification system in tea leaves using improved YOLOv5 deep learning model

Xianghong Deng[1], Zhiwei Zhou[1], Mingze Lei[2], Xuwen Zheng[2*]
[1]Hunan Mechanical & Electrical Polytechnic, Changsha410151, China.
[2]Faculty of Engineering, Mahasarakham University, MahaSarakham44150, Thailand; 2240053955@qq.com (X.Z.).

Abstract: People all throughout the world enjoy tea. The traditional detection of tea diseases relies heavily on farm experts, which often takes a lot of time. Computer vision technology and artificial intelligence can automatically detect diseases and control their spread in tea leaves in real time. This study suggested a deep learning-based version of the YOLOv5 algorithm. It then improved the model structure of YOLOv5 and integrated Omni Dimensional Dynamic Convolution (ODConv) and Convolutional Block Attention Module (CBAM) attention mechanisms into YOLOv5. The findings demonstrate that the enhanced YOLOv5's accuracy is better than that of previous approaches and 5.06% greater than it was prior to the enhancement. Using the improved YOLOv5 detection algorithm to achieve a high-precision model for identifying tea pests and diseases and combining it with the Pyside6 library to design an interface recognition system, the development of the target detection recognition page is completed. In conclusion, this work offers a useful deep learning-based technique for the quick and precise diagnosis of tea disorders in the area of autonomous tea disease detection.
Keywords: *Aattention mechanisms, Deep learning, Smart agriculture, Tea pests and diseases.*

1. Introduction

People everywhere all like to drink tea, which is extremely health beneficial. In China, tea is one of the traditional beverages [1]. Therefore, tea is an important cash crop, and tea trees are also planted in a large area on the earth.However, tea infections can also lower tea quality and result in significant financial losses for tea gardens during the planting of tea trees. According to statistics, there are more than 900 kinds of tea tree pests and diseases recorded. When tea leaves are healthy, they have a vivid color, but when they are affected by diseases, the color of the tea leaves will change significantly. Plants are susceptible to many illnesses as a result of unfavorable seasonal and environmental factors [2]. Farmers in the fields judge and identify tea diseases with their naked eyes and past experience [3].The precise identification and detection of tea-related illnesses, coupled with the prompt adoption of prevention and control measures, holds immense significance in minimizing production losses, enhancing tea quality, and augmenting the income of tea farmers [4].Computer image processing technology has been widely used in recent years to address a variety of agricultural science issues due to the quick development of smart and precision agriculture [5].Many machine learning algorithms have been used for tea leaf disease classification [6].Deep learning is a branch of machine learning that involves writing algorithms that mimic human brain functions. The convolutional neural network (CNN) is now the most widely used deep learning technique [7].Deep learning (DL) has become more significant because of its powerful computing capabilities [8]. Among conventional techniques, deep learning is distinct because it eliminates the need for a laborious manual feature extraction procedure and enables the network to learn discriminative features straight from pixel intensities [9].To get high-level properties, deep learning models employ many layers that process data through nonlinear

History: Received: 10 January 2025; Revised: 28 February 2025; Accepted: 3 March 2025; Published: 22 March 2025
* Correspondence: 2240053955@qq.com

transformations [10]. Deep learning algorithms are highly accurate and have layers with phases for feature extraction [11].

In the study, an automatic system was developed using Support Vector Machine (SVM) to detect three different types of tea diseases with a rather limited set of characteristics[3]. Researchers proposed the use of Convolutional Neural Networks (CNN) for the detection of tea diseases. Konstantinos P. Ferentinos[12]used deep learning models, VGG (highest success rate) and AlexNet OWTBn architectures (lowest final average testing error),for plant disease detection and diagnosis. The system uses three models, two of which are pre-trained. They are Faster RCNN (Inception-v2) and VGG16. The other one is a manually trained model, specifically a sequential model or CNN[13].In this study, the model based on a faster region-based convolutional neural network (Faster R-CNN) is trained using over 1000 photos of tea leaves [14].This study presented AX-RetinaNet, an improved variant of RetinaNet, as an enhanced target detection and recognition network for automatically detecting and recognizing tea illnesses in natural scene photos [4].

In order to identify tomato diseases and pinpoint the areas of damaged tomato leaves, this study enhanced the Faster RCNN algorithm [15].With the objective of enhancing the detection performance for blurry, occluded, and small diseased leaves, a deep learning framework known as "Faster Region-based Convolutional Neural Network" is employed for detecting TLB (Tea Leaf Blight) leaves[16].To recognize tree leaves, two deep learning models—Faster R-CNN and YOLOv5—representing two-stage and one-stage techniques are used [17].The BiFPN feature fusion network and Adaptive Spatial Feature Fusion (ASFF) technologies have been used to improve the multi-scale feature fusion of tea illnesses, increasing the model's resistance to interference from complicated backgrounds Lin, et al. [18]. Z. Xue, R. Xu, D. Bai and H. Lin [19] replaced the original YOLOv5's spatial pyramid pooling fast (SPPF) module with a receptive field block (RFB) module and incorporated self-attention and convolution (ACmix) and convolutional block attention module (CBAM) into YOLOv5.This article uses the enhanced YOLOv5 to identify peach tree leaf diseases [20].The original Bottleneck CSP module in the main trunk of the YOLOv5s network architecture is replaced by the BottleneckCSP-2 module, which has been enhanced and redesigned [21].The convolutional neural network's calculations are reduced in the backbone network by substituting the Involution Bottleneck module for the bottleneck module in the C3 module [22].The article makes a selection of the YOLOv5 weakly supervised model and subsequently discovers, through rigorous experiments, that incorporating the GAM attention mechanism into the YOLOv5 network framework significantly enhances the ability to identify Apolygus lucorum [23].This research proposes an enhanced SE-YOLOv5 network model for tomato viral illness detection [24]. A shallow feature layer is the basis for the enhanced YOLO v5 algorithm[25].By employing ShuffleNetv2 as the enhanced backbone network, the model structure is subsequently fine-tuned for simplification. During the feature fusion phase, the CBAM module is introduced to further augment the effective feature information pertaining to lychees[26].A multi-attention mechanism module, consisting of the Convolutional Block Attention Module (CBAM) and Efficient Channel Attention (ECA), was placed into the network's neck [27].The proposed method incorporates an Efficient Channel Attention module (ECA) within the C3 module of the YOLOv5s network model's backbone structure. Additionally, a Global Attention Mechanism module (GAM) is interposed between the neck structure and the head structure[28].In order to address the high miss rate of small targets in wheat population photos, an integrated Transformer small-target detection head (TSDH) was added concurrently[29].With compact, dense, and overlapping crop disease targets, an enhanced YOLOv5s model was put forth to detect a number of prevalent crop illnesses[30].A. Wang, T. Peng, H. Cao, Y. Xu, X. Wei and B. Cui[31]proposed TIA YOLOv5, which adds a transformer encoder block to the backbone network to improve the model's sensitivity to weeds and proposes a Channel Feature Fusion with Inner Convolution (CFFI) strategy for channel feature fusion. The YOLOv5 model has been improved by utilizing a Coordinate Attention (CA) module and a regression loss function for bounding boxes to detect and accurately count pod targets on live plants [32].To facilitate information interchange between channels and refine features, S. Zhu, W. Ma, J. Wang, M. Yang, Y. Wang, and C.

Wang[33]introduced hierarchical mixed-scale units (HMUs) within the neck network. By incorporating CAM into the original YOLOv5s network, the model achieves greater accuracy in locating and identifying fruits within dense images [34].The YOLOv5 model is enhanced and used for the detection of thick and small tea shoots [35].

The categorization and identification of illnesses and pests in tea leaves has improved somewhat thanks to deep learning models. On the other hand, the model's performance and parameters can still be improved. Based on the above research, this paper proposes according to the characteristics of several tea diseases, the identification model of diseases and insect pests is constructed to accurately identify the types of tea diseases and insect pests, and then provide reasonable suggestions for farmers to actively prevent diseases and insect pests, and solve the problem of intelligent tea garden disease and insect pests' identification.

Therefore, in order to balance detection speed and accuracy, we have built a lightweight network model based on the design of YOLOv5.The following are this paper's primary contributions:

(1) The backbone structure module has been improved, and CSP has been upgraded to CSPK, which enhances feature propagation, aids in better information transmission across different layers of the network, thereby improving performance.

(2) In the YOLOv5 model, we have incorporated the Convolutional Block Attention Module (CBAM) to enhance the recognizability of feature images. Secondly, owing to the extensive area of tea lesion and strong background contrast, the YOLOv5 backbone network has introduced the CBAM attention mechanism, which improves the directionality of tea lesion recognition, resulting in faster convergence speed and enhanced reasoning and training capabilities of the detection algorithm.

(3) ODC, dynamic convolution enhances the adaptability of features, while attention mechanisms enhance the network's ability to focus on the specific information required for the current task.

(4)A visual interface for detecting and recognizing tea disease images has been designed, and the functions of tea disease and pest diseases include importing and initializing training models; Adjusting confidence and IOU (Intersection over Union) thresholds, image upload, detection, visualization of results, export of results, and ending the detection; video upload, detection, visualization of results, export of results, and ending the detection.

The following is the structure of the paper: The tea plant dataset, the enhanced YOLOv5s model, the experimental setup, and the evaluation measures are all covered in detail in Section 2. The findings are examined and discussed in Section 3. Our tea plant disease and pest detection method is shown in Section 4. The conclusions are summed up in the last section.

2. Materials and Methods

2.1. Dataset

All the photos of tea plant diseases and pests utilized in this study were captured outdoors in natural environments. The images were captured at the Xiangcha Tea Farm in Changsha, Hunan Province, China. Image collection was performed under clear weather conditions, with adequate light and a favourable collection environment. Since our photos were taken in natural settings, we captured both images with dense targets and images with fewer targets. Phenomena such as occlusion by leaves, overlapping leaves, and large areas of vegetation are present in the photos. There are also shadows and reflections on the leaves in the images. A few typical samples from the dataset are displayed in Figure 1.

Figure 1.
Some illustrative examples from our dataset.

The photos of tea plant diseases and pests are stored in JPG format. To label the data, we employed LabelImg [36]. The label files are ultimately saved in VOC format. Labeling boxes are added and corresponding labeling files are generated for the regions of the image suffering from tea diseases. The image labeling is shown in

DOI: 10.55214/25768484.v9i3.5690

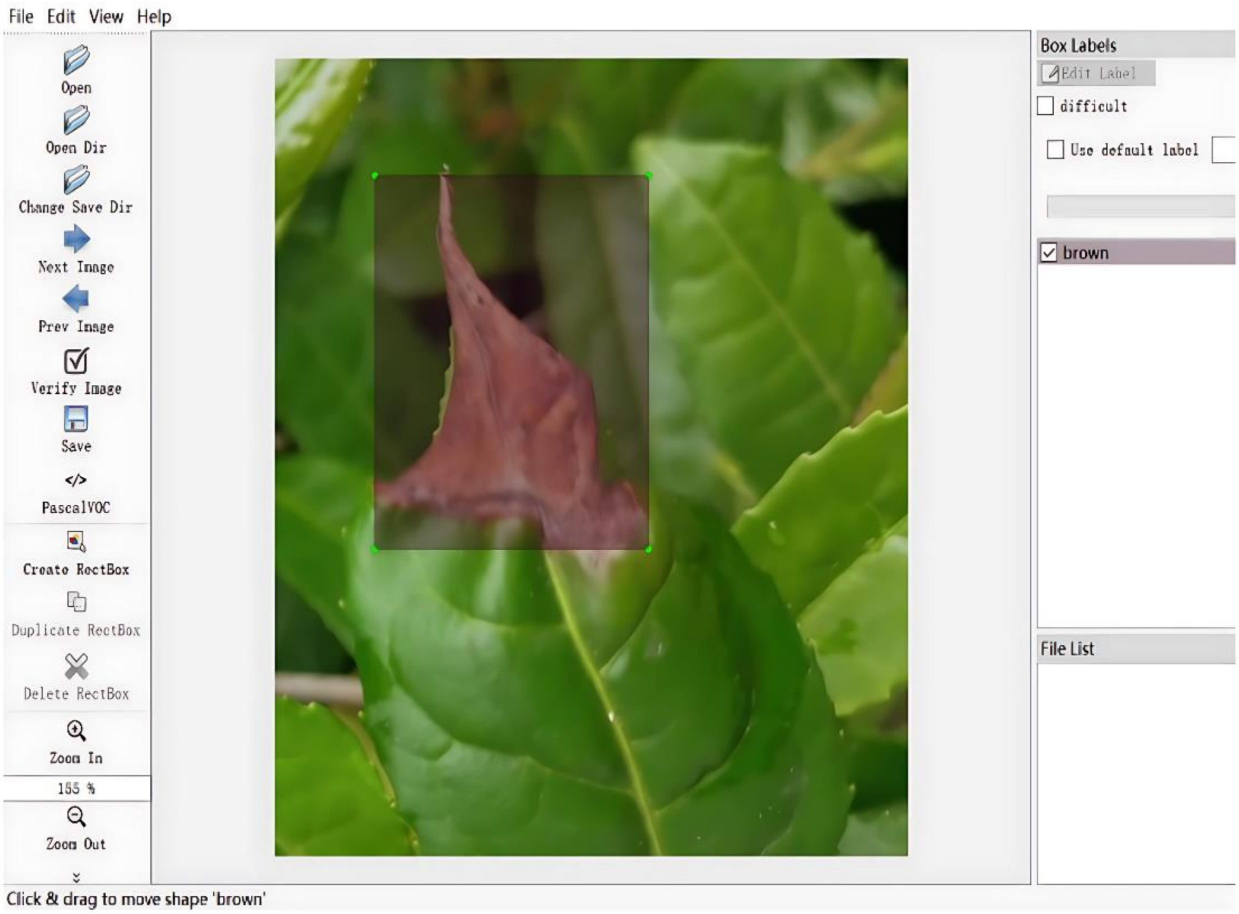


Figure 2. To increase the size of the final dataset, we employed data augmentation. Lastly, an 8:1:1 randomization process was used to divide the annotated photos into training, validation, and testing sets. By employing random scaling, cropping, and layout techniques for stitching, mosaic data augmentation creates a single image out of four.

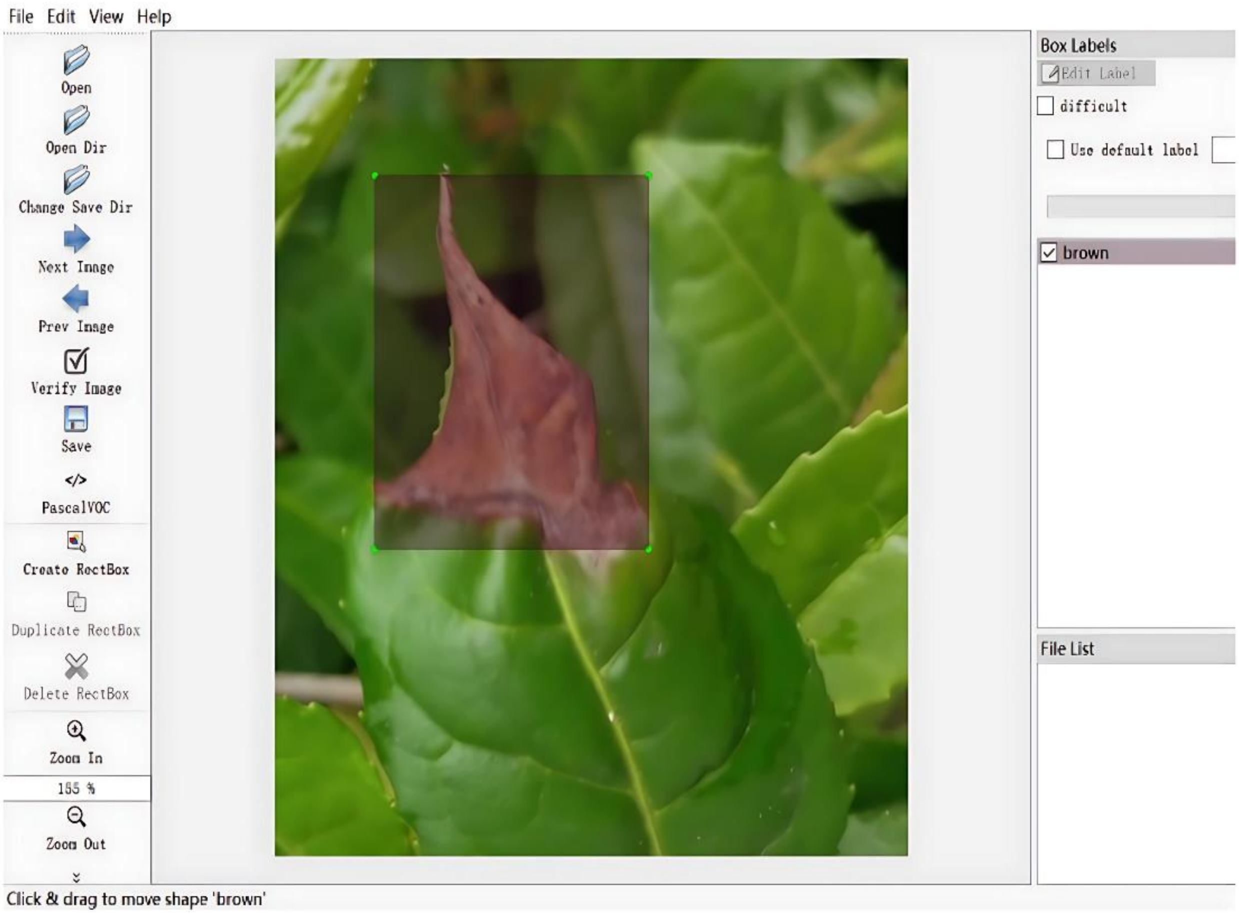


Figure 2.
Image Labeling fabrication.

2.2. Original YOLOv5s

The Yolo algorithm stands for "You Only Look Once," and it utilizes a single convolutional neural network (CNN) model to achieve end-to-end object detection. YOLOv5 can be divided into YOLOv5n, YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x based on the various network widths and depths.Based on YOLOv5s-5.3, we present an upgraded version of the YOLOv5s disease and pest detection model in this study. On some researchers' devices, the image inference speed of the Yolov5s model has reached 455FPS, which is a significant advantage that has been widely used by many scholars[37].The Yolov5 algorithm utilizes deep learning techniques to swiftly and accurately identify objects in images, while annotating their positions and class information, as shown in

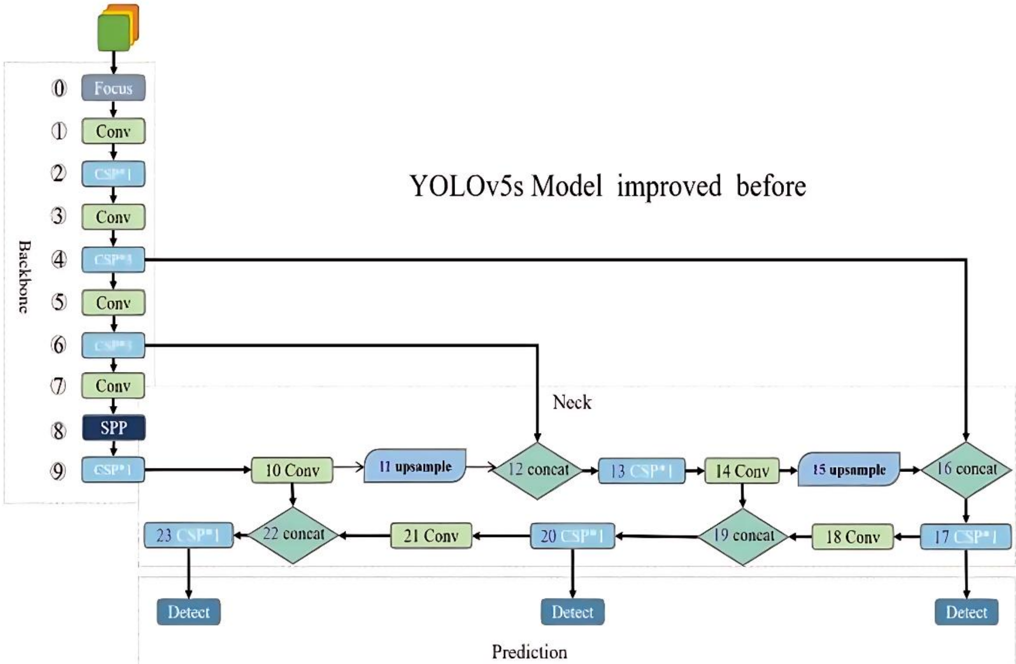


Figure 3.

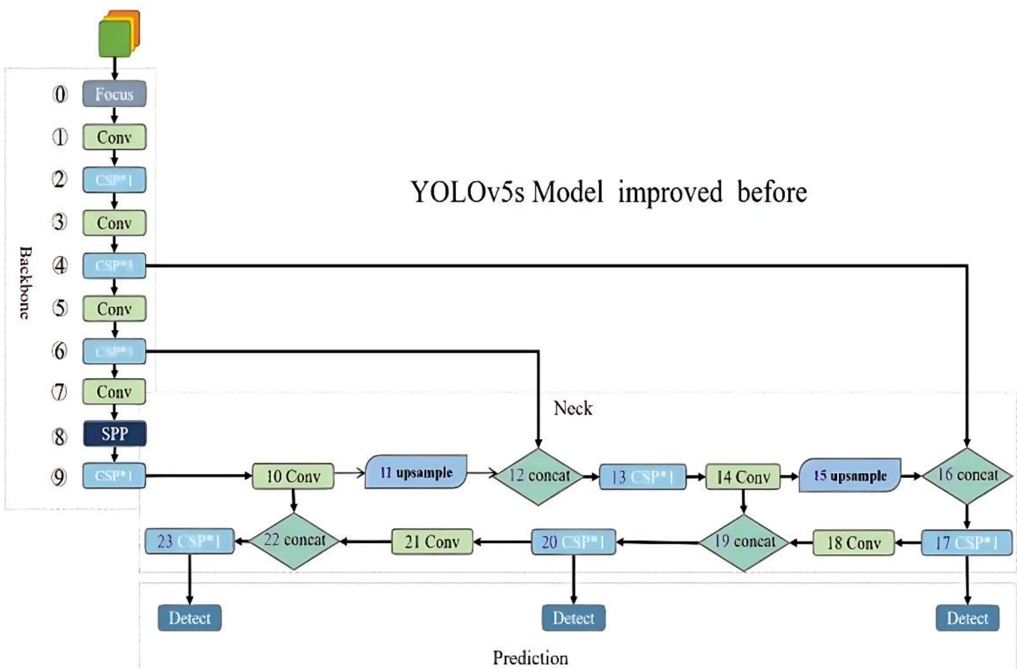


Figure 3.
The network structure of Yolov5s-5.3 model.

As shown in Fig. 3, The Yolov5 Version 5.3 network structure model includes three parts: Backbone function: feature extraction. Neck function: mix and combine features in one wave, and transfer these features to the prediction layer. Head function: make final prediction output. The input, backbone, neck, and prediction are the four primary parts of the YOLOv5 network topology. YOLOv5's input part serves as the network's entrance point. Here, the input data is randomly cropped and concatenated using the Mosaic data augmentation approach. The backbone network is the feature extraction part of YOLOv5, and its feature extraction capability directly impacts the overall performance of the network. The input feature map is split into two halves by YOLOv5's use of the CSPNet (Cross Stage Partial Network) structure during the feature extraction stage. One portion is directly down sampled, while the other is processed through a sequence of convolutional layers. Ultimately, feature maps from these two sections are combined. This design enables the network to have stronger non-linear expression ability, which can better handle complex backgrounds and diverse objects in object detection tasks, as shown in

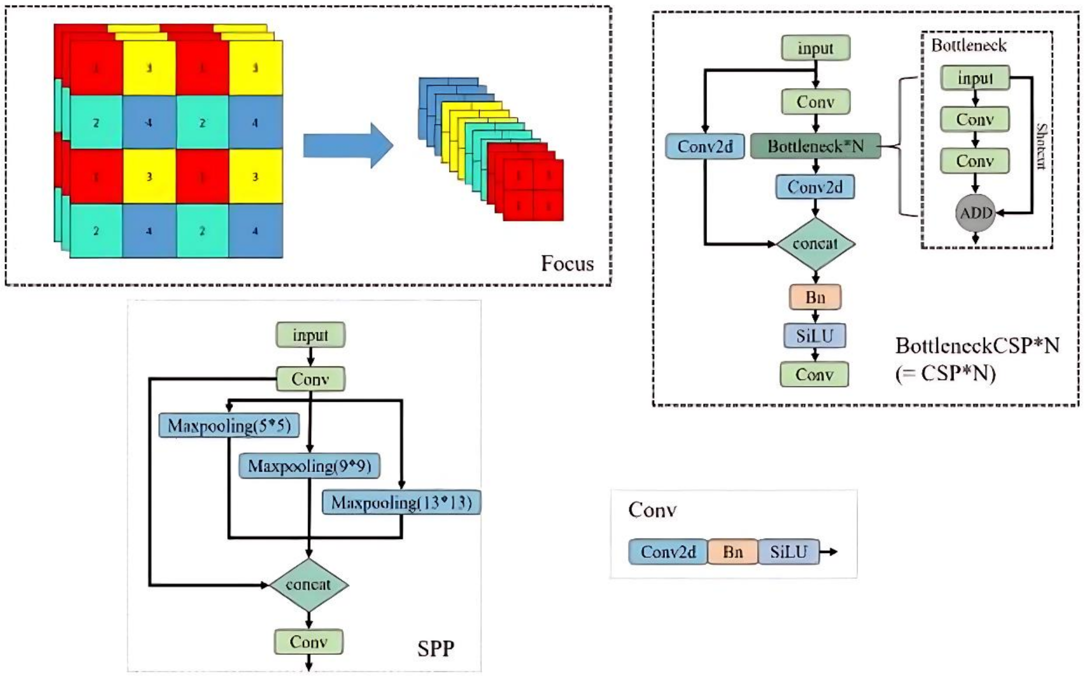


Figure 4.

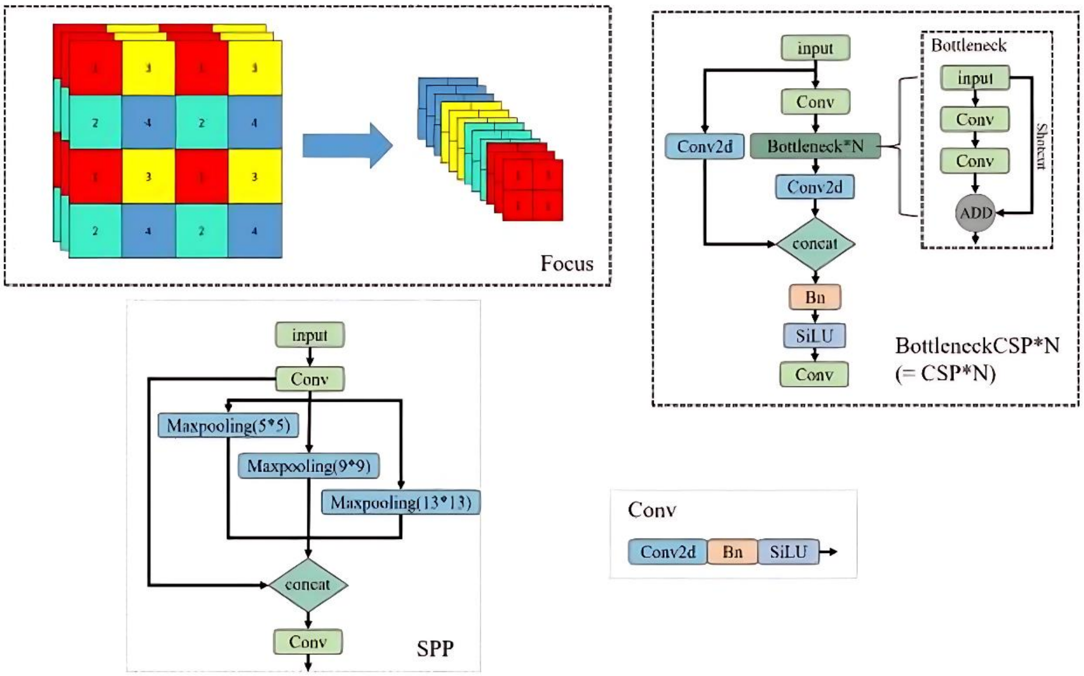


Figure 4.
The component networks in the original YOLOv5s.

The convolutional module is a commonly used fundamental block in convolutional neural networks, consisting primarily of convolutional layers, Batch Normalization (BN) layers, and activation functions. One of the fundamental layers in a convolutional neural network is the convolutional layer, which serves to extract local spatial information from input features. Following the convolutional layer, a BN layer is added, functioning as a normalization layer that standardizes the distribution of feature values within the neural network. The activation function is a nonlinear function that introduces nonlinear transformation capabilities into the neural network. Common activation functions include sigmoid, ReLU, LeakyReLU, ELU, and others. The three detect detectors that make up the head's main body use grid-based anchoring to identify objects on feature maps with varying scales.

2.3. Improved YOLOv5s

The algorithm must be enhanced in order to increase the accuracy of the YOLOv5 series models.

- The network's feature extraction capability is enhanced by improving the feature extraction structure;
- A tiny target detection layer is added to enhance the detection performance of small targets.
- Enhance the model's accuracy and the backbone network.

As shown in Fig. 5, the improvements to this model include changing CSP to CSPK, enhancing the convergence of feature maps, preventing feature loss, and adding BN layers and the LeakyReLU activation function to improve the stability and generalization performance of the model. This will allow the model to pick up additional features. And by adding ODConv2d and CDAM to the model's head to further enhance feature adaptability and improve the performance of focusing on essential features, as shown in

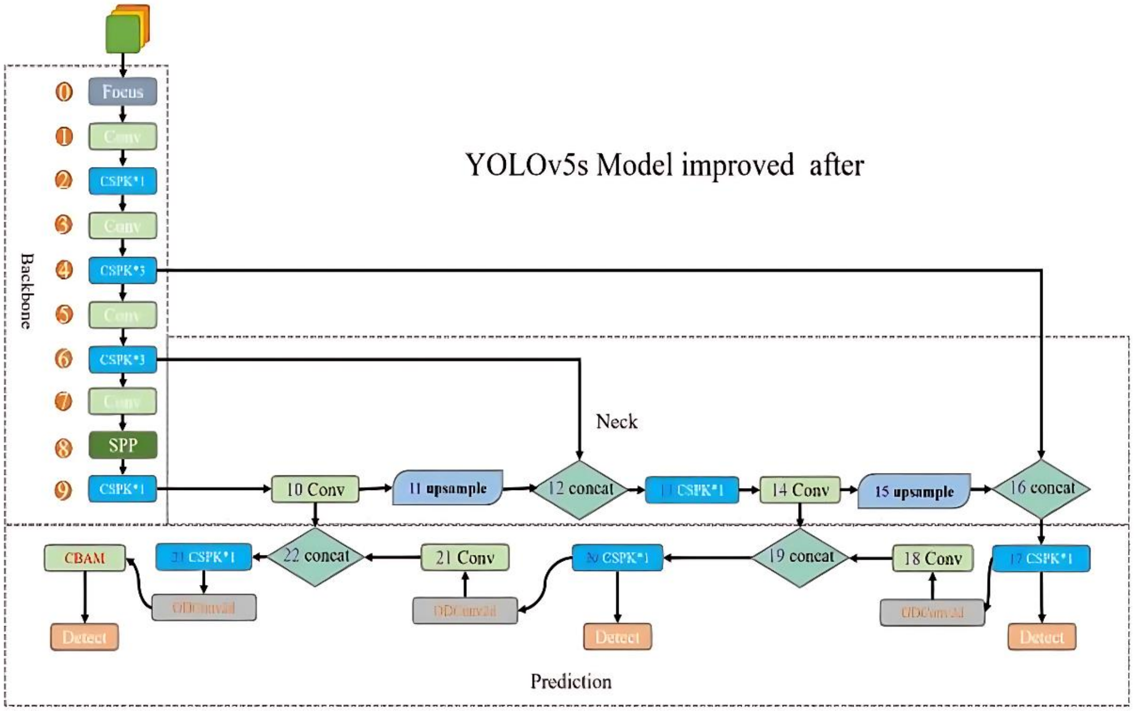


Figure 5.

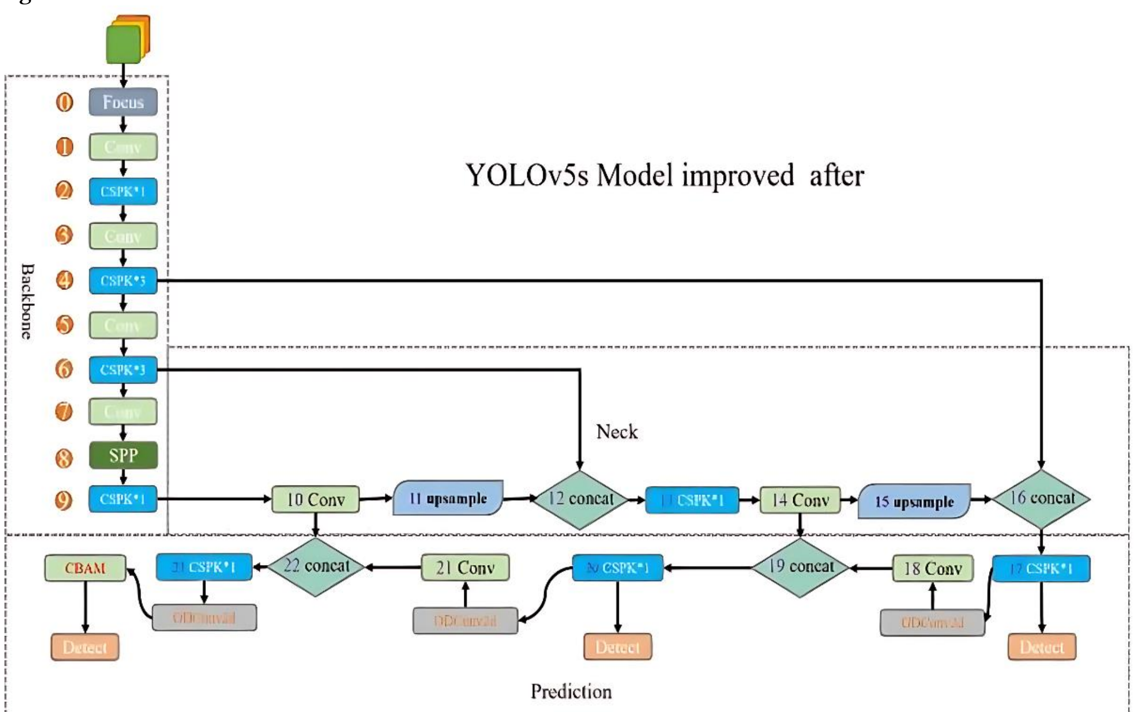


Figure 5.
The network structure of the improved YOLOv5s.

In neural network architectures, CSP (Cross Stage Partial) is a connection approach that is mostly used to improve the information transmission and interaction of features across several stages or scales.

The term "conv2d" describes a convolution procedure that is frequently used to process picture data when it is applied to two-dimensional data. It takes in as input a three-dimensional tensor, of which the first and second dimensions stand for the image's height and breadth, and the third dimension for the image's channel count. CSP plays a crucial role in boosting the network's representational capacity through improved feature extraction and representation, which leads to improved performance in a variety of computer vision applications.

The CSPK structure can enhance feature propagation more, aiding in better information transmission across different layers of the network, thereby improving performance, as shown in

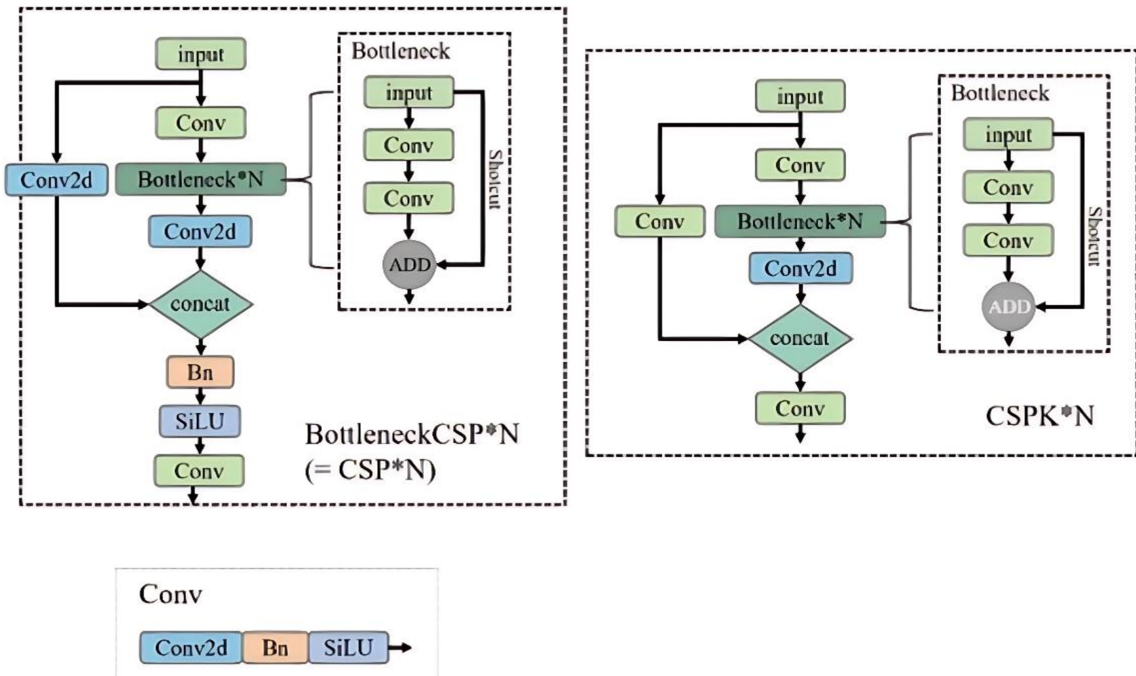


Figure 6.

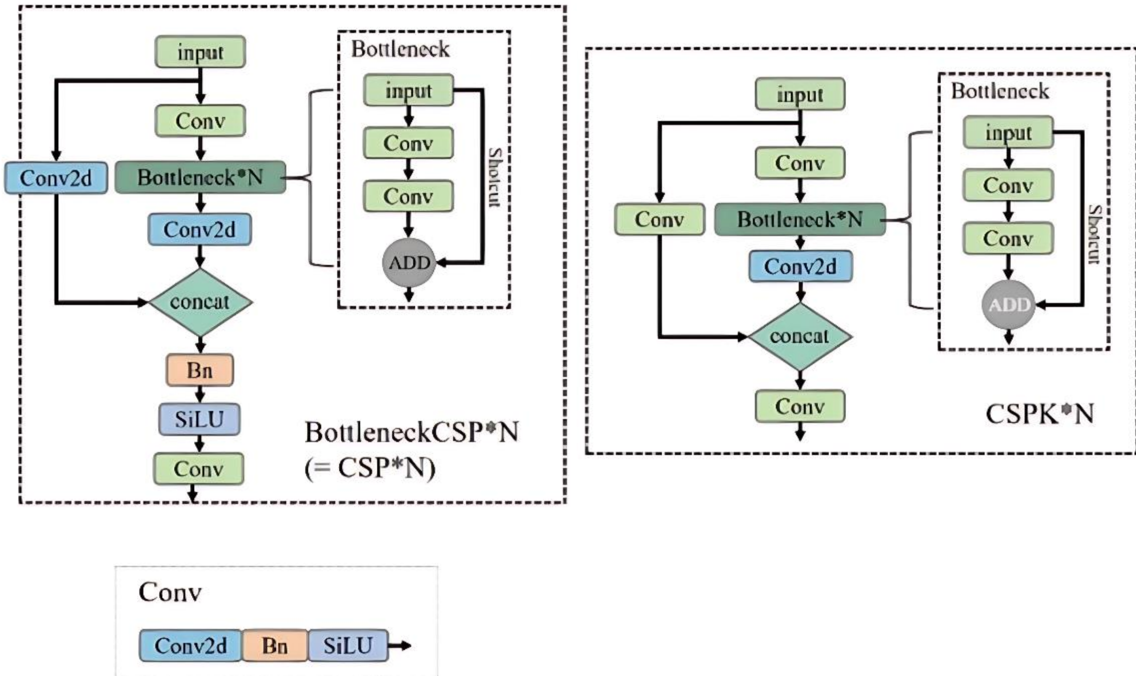


Figure 6.
Comparison of CSP and CSPK Structure.

Adding ODConv2d and CBAM (see

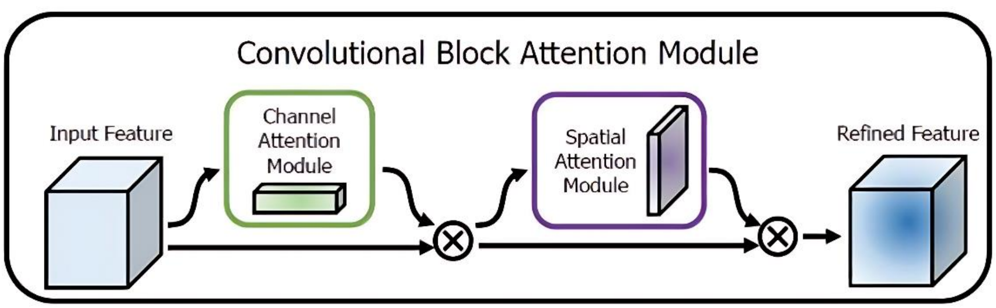


Figure 7) to the model's head to further enhance feature adaptability and improve the performance of focusing on essential features. Dynamic convolution refers to the concept where the weights of the convolutional kernel are not fixed during the convolution process. Instead, they can be dynamically adjusted based on the input data. This adaptability of the convolutional kernel allows it to better capture the features present in the input data, thereby enhancing the performance of convolutional networks. Attention mechanisms are crucial neural network modules that enable networks to focus more effectively on relevant information when processing sequential data. Attention mechanisms guarantee that the network gives more attention to crucial information during data processing and transmission by allocating varying weights to various input data components according to their importance.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 9, No. 3: 1777-1794, 2024
DOI: 10.55214/25768484.v9i3.5690

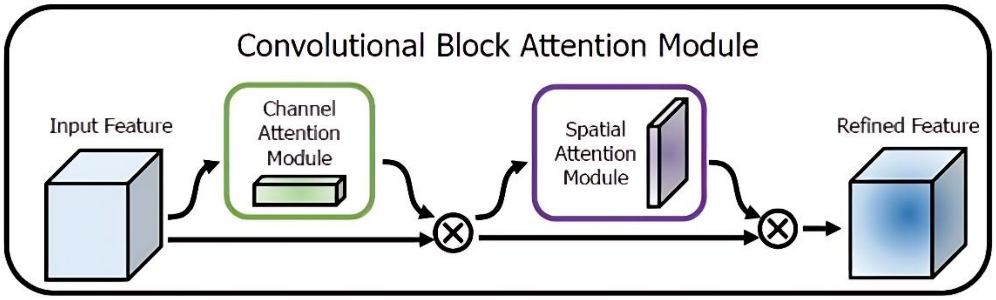


Figure 7.
Diagram of CBAM.

The channel submodule uses the shared network's max pooling and average pooling outputs, as seen in Figure 7, while the spatial submodule uses two similar outputs that are pooled along the channel axis and then sends them to the convolutional layer.To learn the attention of convolutional kernels along the four dimensions of kernel space in any convolutional layer, ODConv employs a revolutionary multidimensional attention mechanism and parallel technique. The number of convolution kernel input channels, the convolution kernel's receptive field, the number of convolution kernel output channels, and the number of convolution kernels are the four distinct forms of attention.As an alternative to conventional convolution, ODConv can be inserted into many CNN architectures. That is to say, this module is plug-and-play.As a result, while both dynamic convolution and attention mechanisms contribute to improving the performance of neural networks, they serve different purposes. Dynamic convolution enhances the adaptability of features, while attention mechanisms enhance the network's ability to focus on the specific information required for the current task.

2.4. Experiment Setup

This experiment is based on a windows operating system, using PyTorch learning framework to load the model, apply Python language to write the program, and save the model parameters with the best effect during the training process. Table 1 displays the experimental platform's configuration table.For the improved YOLOv5s (including other comparative models) used in this paper, the input images were of 640 × 640 pixels, with a batch size of 16, momentum of 0.937, weight decay of 0.0005, IoU of 0.2, hsv_h of 0.015, hsv_s of 0.7, hsv_v of 0.4, mosaic of 1.0, scale of 0.5, translate of 0.1, mixup of 0.0, and a training time of 150 epochs. Random initialization techniques were used to initialize the weights, allowing all models to be trained from scratch.

Table 1.
The experiment platform's configuration table.

Setting for Experiments	Specifics
Mirroring	PyTorch1.7.0 Python3.8(ubuntu18.04) Cuda 11.0
GPU	Tesla T4(16GB) * 1
CPU	8 vCPU Intel Xeon Processor (Skylake, IBRS)
Memory	56GB

2.5. Evaluation Metrics

We will employ a number of standard evaluation measures to evaluate the models' performance in our investigation[38].

2.5.1. Confusion Matrix

The four categories of categorization task prediction outputs in machine learning and deep learning are known as confusion matrices and include the following: One tool for visualizing and summarizing classification problem performance is the confusion matrix[39].

DOI: 10.55214/25768484.v9i3.5690

True Positive (TP) [2]: Both the label value and the anticipated value are positive cases, suggesting that the prediction is accurate. False Negative (FN): An inaccurate prediction is made when the label value is positive and the projected value is negative. False Positive (FP): An inaccurate prediction is indicated when the label value is a negative example despite the prediction being a positive example. True Negative (TN): This indicates that the prediction is accurate as both the label value and the predicted value are negative cases.

2.5.2. Precision

Precision(P) refers to precision, and its definition is as follows:

$$\text{Precision} = \frac{\text{TP}}{\text{FP} + \text{TP}} \times 100\% \qquad (1)$$

Accuracy's denominator comprises TP (true positives) and FP (false positives). TP counts cases predicted positive that are actually positive, while FP counts wrongly predicted positives (actual negatives). Accuracy reflects the correlation between predicted positives and actual positives/negatives. Higher accuracy means lower FP, fewer misclassified cases, and higher purity in predicted positives, translating to fewer false positives.

2.5.3. Recall

Recall(R) refers to the recall rate, and the definition of Recall is as follows:

$$\text{Recall} = \frac{\text{TP}}{\text{FN} + \text{TP}} \times 100\% \qquad (2)$$

According to this definition, the denominator of Recall is the sum of TP and FN, where TP is the number of positive cases predicted, and the actual value is also positive; FN is the number of negative cases predicted, but the actual value is positive.

Analyzing the formula reveals that Recall encompasses both predicted positive and negative cases alongside actual positive cases. As Recall increases, False Negatives (FN) decrease, indicating a reduction in positive cases misclassified as negative. This translates to identifying a greater number of positive cases, signifying fewer missed detections.

2.5.4. P-R curve

Precise and Recall coordinates, respectively, surround the P-R curve. The lines of different colors represent the PR curves of different categories, while the thick blue lines represent the average PR curves of all categories.

When assessing a model's prediction outcomes, the region bounded by the axes and the P-R curve can be used as a reference. One can conclude that the prediction results of one model are unquestionably superior to those of the other if the P-R curve of one model fully encloses the P-R curve of the other model.

2.5.5. AP

AP (Average Precision): In spite of its name, AP is not determined by averaging precision numbers. Rather, it uses integration to find the area under the PR curve (Precision-Recall curve) for every class. A model's overall Precision and Recall are comparatively greater if its AP is higher, which means the area bounded by the PR curve and the axes is larger.

$$\text{AP} = \int_0^1 \text{P(R)dR} \qquad (3)$$

2.5.6. mAP

Determine the average AP value for each category using the mAP (mean of average precision) formula. The accuracy of predictions for each category can be represented by AP, and the accuracy of the overall model is represented by mAP, which is the average of the APs for all categories. The region that the PR curve and coordinate axis enclose increases with mAP. We frequently discuss the accuracy of a particular object detection technique, commonly known as mAP.

$$mAP = \frac{\sum_{i=1}^{N} APi}{N} \tag{4}$$

2.5.7. mAP@0.5

The YOLO model displays mAP@0.5. When the IOU (Intersection over Union) criterion is 0.5, this expression shows the value of mAP. The object prediction is deemed accurate when the IOU between the predicted box and the annotated box is larger than 0.5; mAP is then computed using this assumption. Map@0.5, generally speaking. It is among the metrics used to assess the YOLO model.

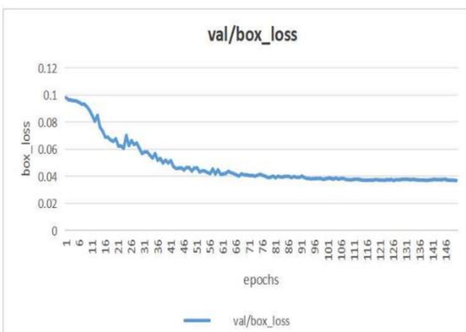
2.5.8. mAP@ [0.5:0.95]

Another way to express the mean Average Precision at different Intersection over Union (IoU) thresholds in the YOLO model is as mAP @ [0.5:0.95]. Ten IoU thresholds are set with a step size of 0.05 within the range of [0.5, 0.95], and the Average Precision at each of these thresholds is computed. The mAP @ [0.5:0.95] is then calculated by taking the mean of these values.

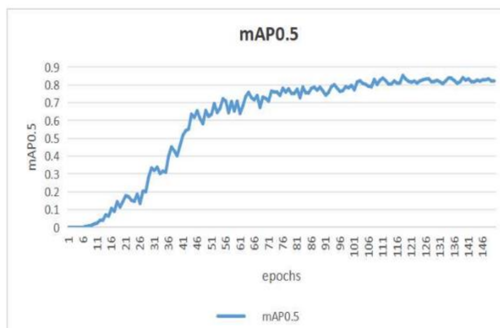
3. Results and Discussion

3.1. The Diseases and Pests in Tea Leaves Detection

From the

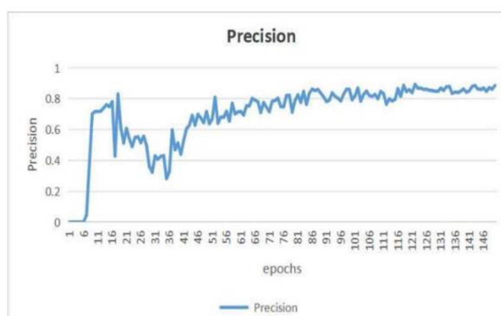


(A) Box loss values



(B) Precision values

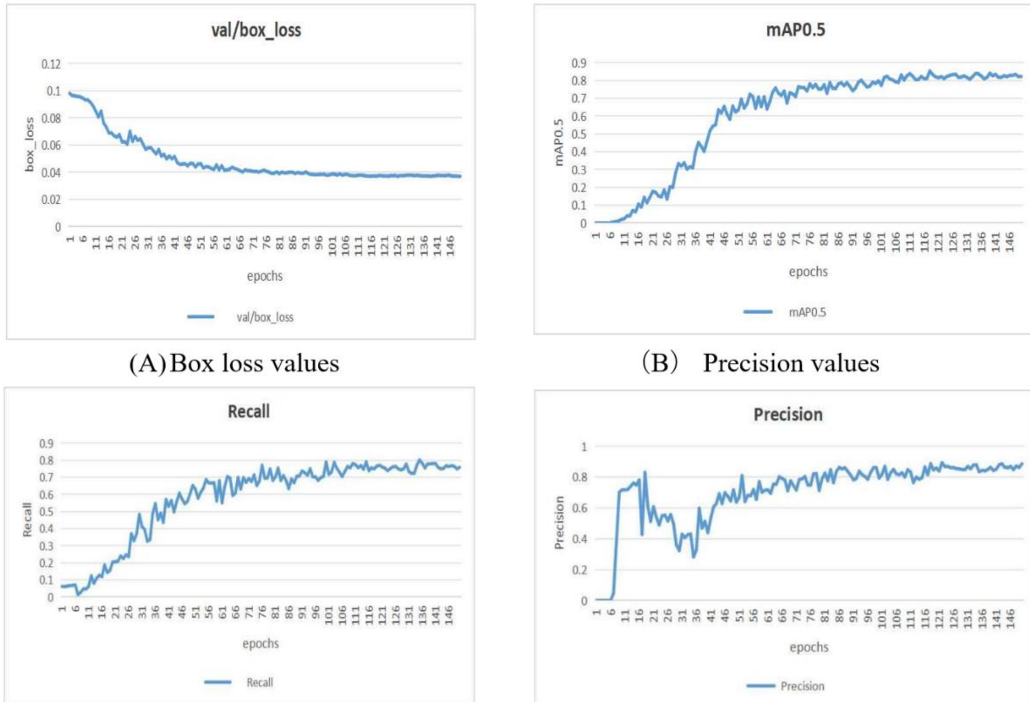
(C) Recall



(D)mAP0.5 values

Figure 8 below, it can be seen that as the number of training iterations increases, the training validation loss of the model gradually decreases, indicating that the model continuously learns more accurate features. Following training, we assessed the model using the dataset's validation set and got the following outcomes.

ISSN: 2576-8484
Vol. 9, No. 3: 1777-1794, 2024
DOI: 10.55214/25768484.v9i3.5690



(A)Box loss values (B) Precision values
(C) Recall (D)mAP0.5 values

Figure 8.
The training Results.

Consequently, we decided to choose the weight that yielded the best accuracy as our ideal weight. The PR curve of our enhanced YOLOv5 model on the validation data is displayed in Figure 9. Figure 9 shows that the model's accuracy and recall were both high.

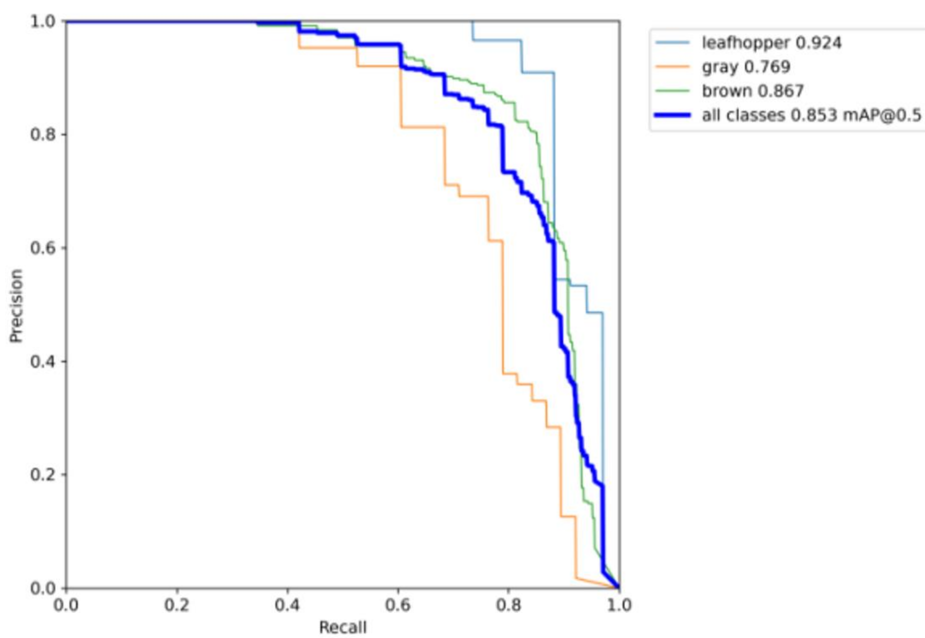
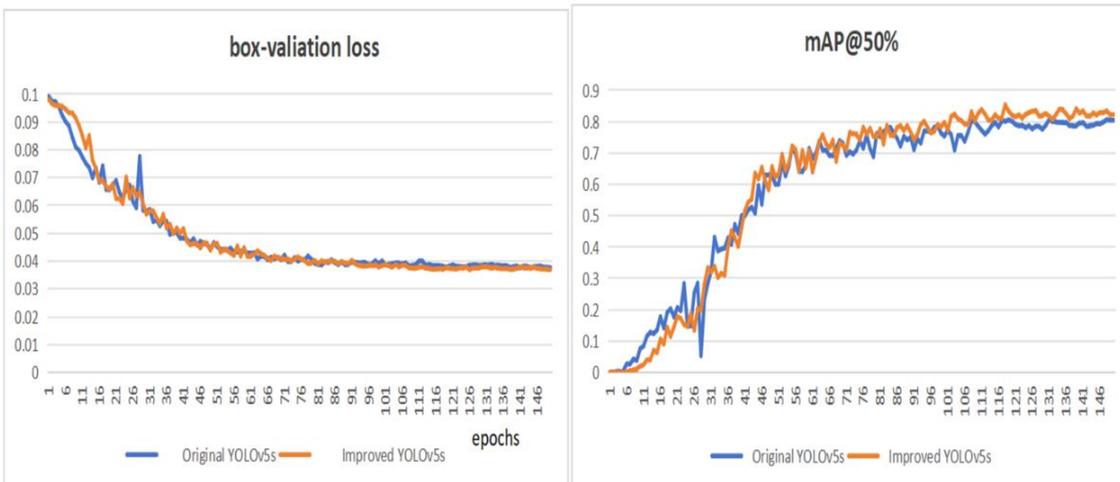


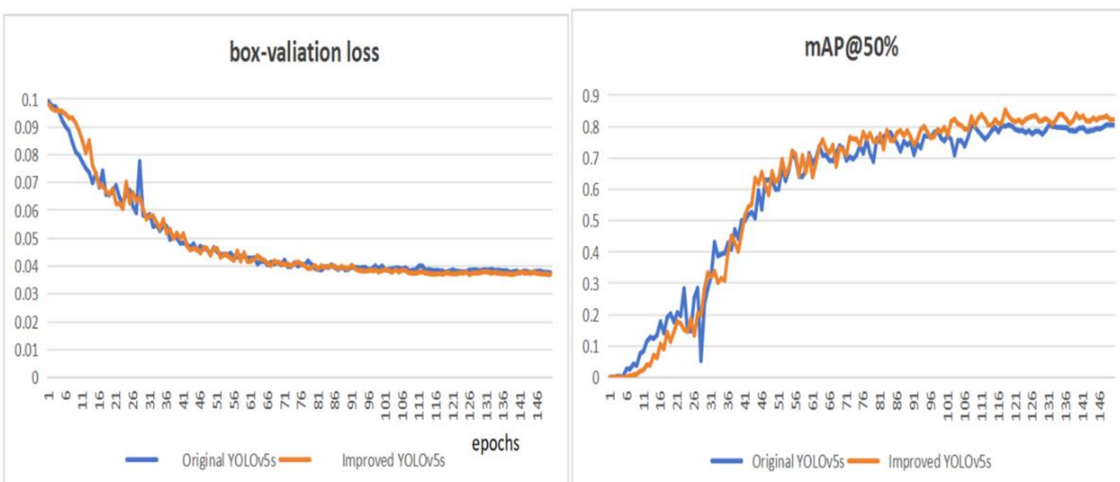
Figure 9.
The Precision-Recall Curve.

As shown in



(A)box-validation loss (B) mAP@50%

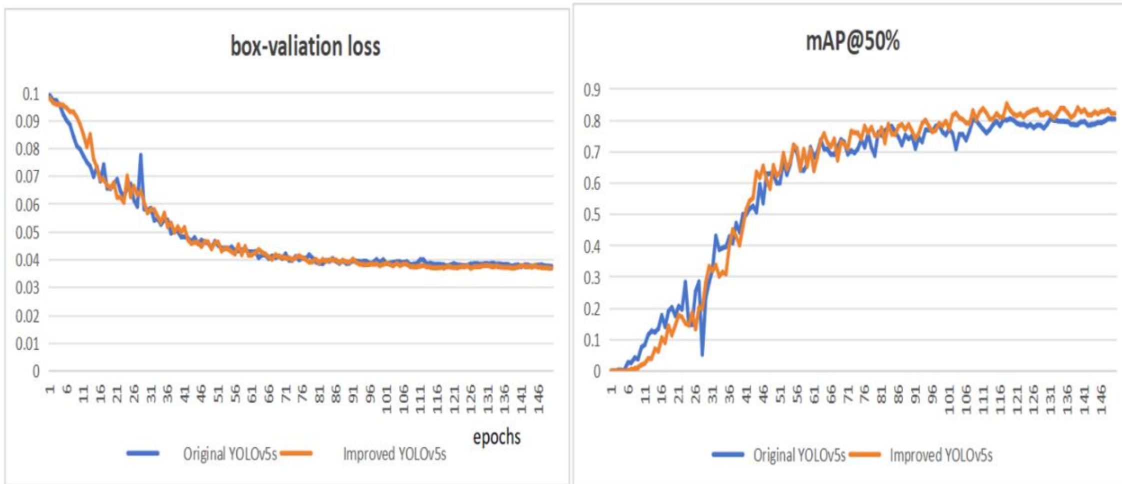
Figure 10(A) and



(A)box-validation loss (B) mAP@50%

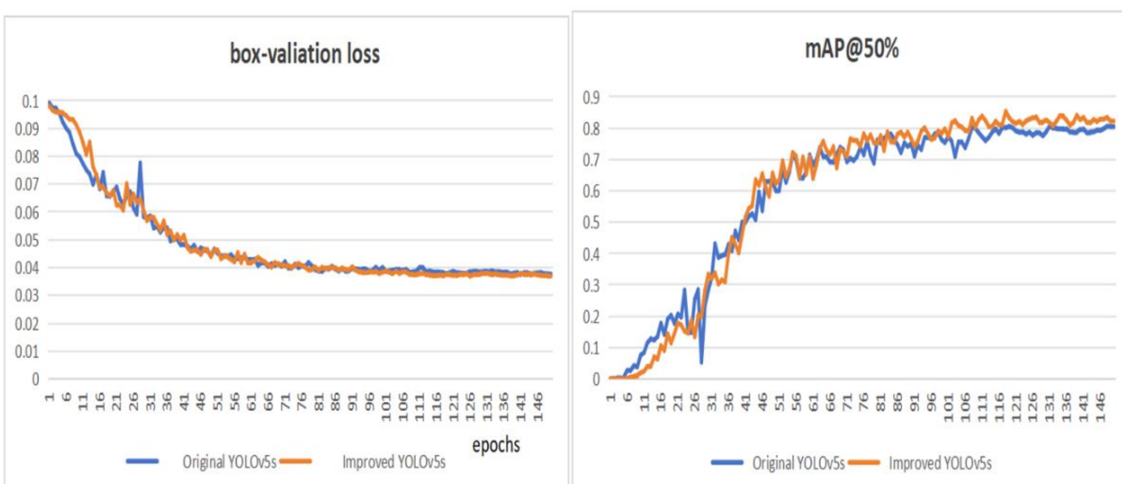
Figure 10(B), the enhanced YOLOv5s model outperforms its original counterpart.

DOI: 10.55214/25768484.v9i3.5690



(A)box-validation loss (B) mAP@50%

Figure 10(A) depicts the box validation loss, which quantifies the accuracy of locating pests and diseases in tea images. This metric exhibits a steady downward trend, indicative of the model's training performance. Notably, the improved YOLOv5s achieves a lower box validation loss compared to the original, highlighting the effectiveness of the deeper neural network architecture. As the model undergoes training, its performance gains momentum. Figure 10(B) further demonstrates that the reduced box validation loss correlates with an increase in mean Average Precision (mAP). Specifically, the enhanced YOLOv5s achieves a mAP of 85.32%, surpassing the 81.3% of the original model.



(A)box-validation loss (B) mAP@50%

Figure 10.
The training outcome of models.

3.2. Comparison of the Model's Output

We used the test set to compare the performance of the enhanced YOLO v5 model suggested in this study with other common convolutional neural networks, including Faster-RCNN, SSD, YOLOv3, YOLOv4, and YOLO v5, for the detection of tea illness. The test set results are displayed in Table 2, which shows that our model has significantly outperformed other models. The mAP has increased by over 4 percentage points when compared to the same series of YOLO v5, suggesting that this model is useful for detecting tea sickness.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 9, No. 3: 1777-1794, 2024
DOI: 10.55214/25768484.v9i3.5690

Table 2.
The results of tea disease detection based on different methods.'

Model	Precision	Recall	mAP（%）
Faster-RCNN	78.46%	50.45%	73.95%
SSD	80.70%	51.46%	62.36%
YOLO v3	86.02%	64.98%	72.74%
YOLO v4	87.56%	56.23%	66.86%
YOLO v5	83.52%	71.71%	81.23%
Our method	88.79%	74.97%	85.36%

The effect of our improved model performance can be seen more intuitively in

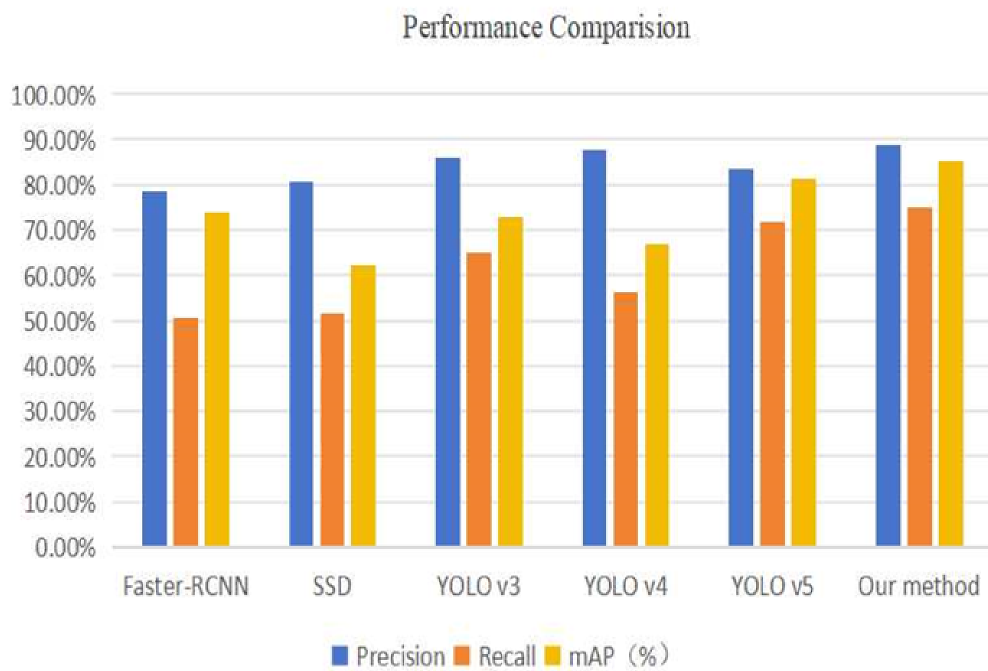


Figure 11.

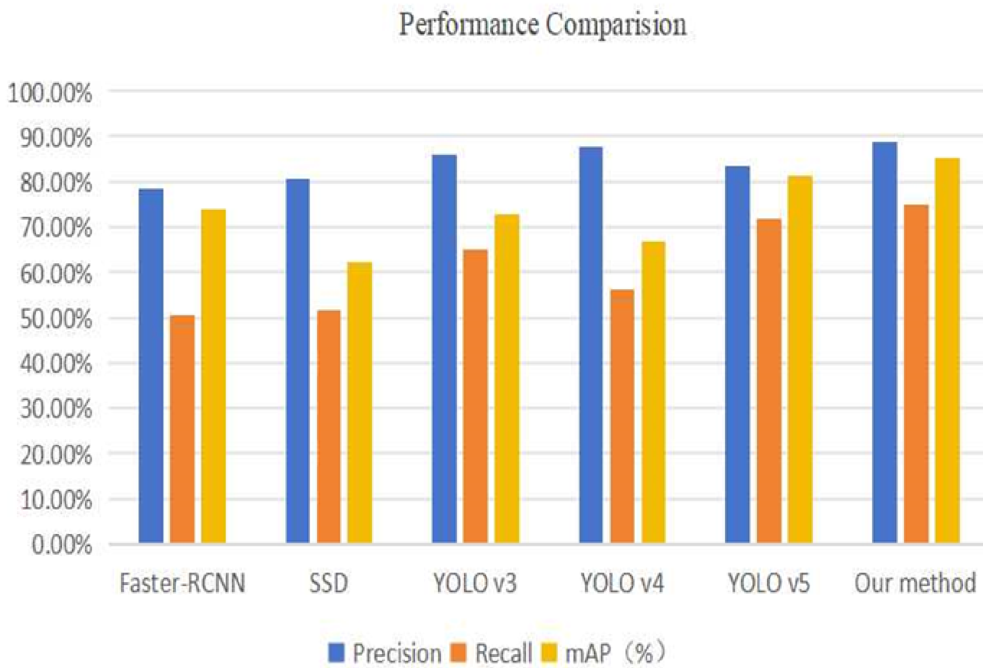


Figure 11.
The performance of models.

4. Tea Disease and Pest Detection System

This system uses an improved YOLOv5 object detection model to train a tea pest and disease dataset and uses the Pyside6 library to build a page display system. The main libraries used in the code include PyTorch, NumPy, OpenCV, PyQt, etc., to complete the development of a tea pest and disease target detection and recognition page. Pyside6 is one of the GUI (Graphical User Interface) programming solutions for Python language, which can quickly create GUI applications for Python programs. Design a graphical interface using Qt Designer, and then convert the designed UI file into Python code using Pyside6.The graphical interface contains multiple UI controls, such as labels, buttons, text boxes, multiple selection boxes, etc. Through the signal slot mechanism in Pyside6, UI controls can be connected to program logic code. The supported functions of the system include model import and initialization; confidence score and IOU threshold adjustment, image upload, detection, visualization of results, result export, and ending detection; video upload, detection, visualization of results, result export, and ending detection; camera feed upload, detection, visualization of results, and ending detection; as well as the list of detected objects, location information, and more. The software interface designed below is simple and elegant overall. Users can click the Select Image button to upload a single image for detection and recognition. Click the detection result display button again to display the input image detection results in the lower left corner of the system, as shown in Figure 12.

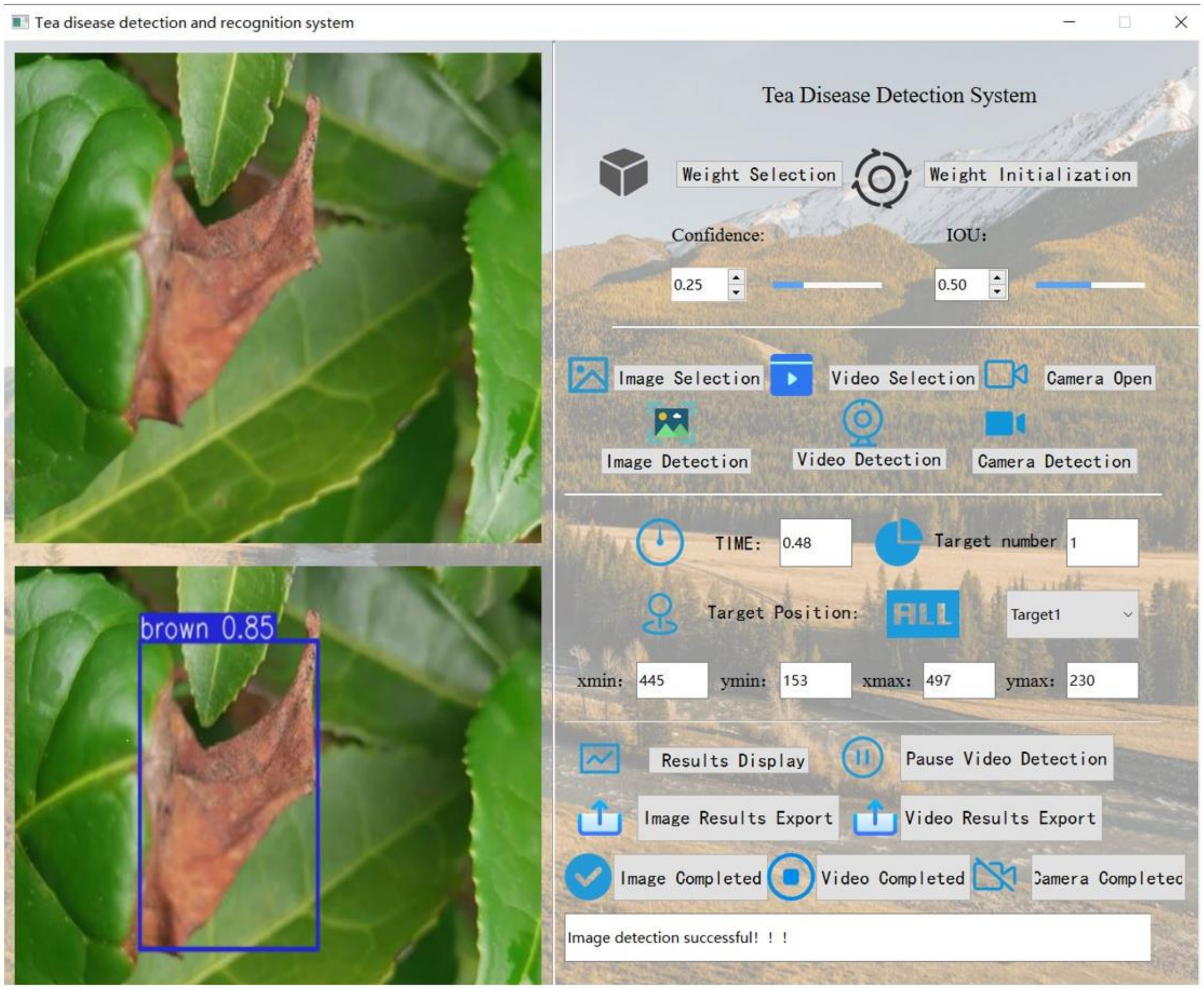


Figure 12.
Tea disease and pest detection display system.

5. Conclusions

There are many different types of tea diseases and pests, and at the moment, identifying these issues primarily depends on the knowledge of experts. As a result, this paper suggests an ensemble learning-based approach for identifying tea plant diseases and pests. To achieve effective disease and pest recognition, we conducted the following work. First, in the object detection domain, the popular YOLOv5 model is selected. Second, we enhanced the YOLOv5 model for disease and pest identification in three ways.We upgraded CSP to CSPK, which better enhances feature propagation, aiding in better information transmission between different layers of the network, thus improving performance. To help the model better focus on objectives like leaf blight, we also included the ODConv and CBAM attention processes. The effectiveness of our model was finally empirically verified in comparison to the original YOLOv5 model, Faster R-CNN, SSD, YOLOv4, and YOLOv3 models. Furthermore, we thoroughly tested the entire system, culminating in the development of a seamless, high-precision object identification system interface—which serves as the presentation interface for this paper—complete with a user interface, test photos, and videos.

In future endeavors, our focus will remain on refining the model and exploring more streamlined

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 9, No. 3: 1777-1794, 2024
DOI: 10.55214/25768484.v9i3.5690

approaches with diminished parameter density. We also intend to delve into methods for deploying models designed for tea plant pest and disease detection. To enhance the comprehensiveness of our research, it is imperative to increase the image capture at varying heights, encompassing the entire tea disease statistics process, and augment the training samples to boost recognition accuracy. Furthermore, we envision integrating detection methods with drones, minimizing model parameter counts, and achieving real-time detection. However, it's worth noting that swift detection necessitates specific hardware configurations. We will strive to continually enhance the model while extending our research on diverse tea diseases and broadening its applicability. We firmly believe that this approach can make meaningful contributions to sustainable, green, and smart agriculture.

Transparency:
The authors confirm that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.

Acknowledgments:
For the facility support, the authors would like to thank the College of Electrical Engineering, Hunan Mechanical & Electrical Polytechnic, and Faculty of Engineering, Mahasarakham University.

Copyright:

References

[1] S. Xiaoxiao, M. Shaomin, X. Yongyu, C. Zhihao, and S. Tingting, "Image recognition of tea leaf diseases based on convolutional neural network," in *2018 International Conference on Security, Pattern Analysis, and Cybernetics (SPAC)*, 2018: IEEE, pp. 304-309, doi: https://doi.org/10.1109/spac46244.2018.8965555.

[2] H. Ulutas and V. Aslantas, "A fast and accurate method for classifying tomato plant health status using machine learning and image processing," *Elektronika ir Elektrotechnika*, vol. 29, no. 2, pp. 54-68, 2023. https://doi.org/10.5755/j02.eie.33866

[3] M. S. Hossain, R. M. Mou, M. M. Hasan, S. Chakraborty, and M. A. Razzak, "Recognition and detection of tea leaf's diseases using support vector machine," in *IEEE International Colloquium on Signal Processing & Its Applications*, 2018, doi: https://doi.org/10.1109/cspa.2018.8368703.

[4] W. Bao, T. Fan, G. Hu, D. Liang, and H. Li, "Detection and identification of tea leaf diseases based on AX-RetinaNet," *Scientific Reports*, vol. 12, no. 1, p. 2183, 2022. https://doi.org/10.1038/s41598-022-06181-z

[5] J. Chen, Q. Liu, and L. Gao, "Visual tea leaf disease recognition using a convolutional neural network model," *Symmetry*, vol. 11, no. 3, p. 343, 2019. https://doi.org/10.3390/sym11030343

[6] R. Latha *et al.*, "Automatic detection of tea leaf diseases using deep convolution neural network," in *2021 International Conference on Computer Communication and Informatics (ICCCI)*, 2021: IEEE, pp. 1-6, doi: https://doi.org/10.1109/iccci50826.2021.9402225.

[7] L. Trihardianingsih, A. Sunyoto, and T. Hidayat, "Classification of Tea Leaf Diseases Based on ResNet-50 and Inception V3," *Sinkron: jurnal dan penelitian teknik informatika*, vol. 8, no. 3, pp. 1564-1573, 2023. https://doi.org/10.33395/sinkron.v8i3.12604

[8] S. Bhowmik, A. K. Talukdar, and K. K. Sarma, "Detection of disease in tea leaves using convolution neural network," in *2020 Advanced Communication Technologies and Signal Processing (ACTS)*, 2020: IEEE, pp. 1-6, doi: https://doi.org/10.1109/acts49415.2020.9350413.

[9] Y. Raju and M. Narayana, "Satellite image segmentation using UNet++ with Vgg19 deep learning model," *Edelweiss Applied Science and Technology*, vol. 8, no. 6, pp. 5707-5722, 2024. https://doi.org/10.55214/25768484.v8i6.3242

[10] L. C. Jaffrin and J. Visumathi, "Impact of feature selection techniques on machine learning and deep learning techniques for cardiovascular disease prediction-an analysis," *Edelweiss Applied Science and Technology*, vol. 8, no. 5, pp. 1454-1471, 2024. https://doi.org/10.55214/25768484.v8i5.1848

[11] A. M. Ridha, M. J. Mohammed, H. A. Saber, M. H. Chyad, and M. H. Abdulameer, "Advanced deep learning approaches for early detection and localization of ocular diseases," *Edelweiss Applied Science and Technology*, vol. 8, no. 6, pp. 3708-3721, 2024. https://doi.org/10.55214/25768484.v8i6.2813

[12] K. P. Ferentinos, "Deep learning models for plant disease detection and diagnosis," *Computers and electronics in agriculture*, vol. 145, pp. 311-318, 2018. https://doi.org/10.1016/j.compag.2018.01.009
[13] M. H. Kamrul, M. Rahman, M. R. I. Robin, M. S. Hossain, M. H. Hasan, and P. Paul, "A deep learning based approach on categorization of tea leaf," in *Proceedings of the International Conference on Computing Advancements*, 2020, pp. 1-8, doi: https://doi.org/10.1145/3377049.3377122.
[14] S.-H. Lee, C.-C. Wu, and S.-F. Chen, "Development of image recognition and classification algorithm for tea leaf diseases using convolutional neural network," in *2018 ASABE Annual International Meeting*, 2018: American Society of Agricultural and Biological Engineers, p. 1, doi: https://doi.org/10.13031/aim.201801254.
[15] Y. Zhang, C. Song, and D. Zhang, "Deep learning-based object detection improvement for tomato disease," *IEEE access*, vol. 8, pp. 56607-56614, 2020. https://doi.org/10.1109/access.2020.2982456
[16] G. Hu, H. Wang, Y. Zhang, and M. Wan, "Detection and severity analysis of tea leaf blight based on deep learning," *Computers & Electrical Engineering*, vol. 90, p. 107023, 2021. https://doi.org/10.1016/j.compeleceng.2021.107023
[17] L. Wang and W. Q. Yan, "Tree leaves detection based on deep learning," in *Geometry and Vision: First International Symposium, ISGV 2021, Auckland, New Zealand, January 28-29, 2021, Revised Selected Papers 1*, 2021: Springer, pp. 26-38, doi: https://doi.org/10.1007/978-3-030-72073-5_3.
[18] J. Lin, D. Bai, R. Xu, and H. Lin, "TSBA-YOLO: An improved tea diseases detection model based on attention mechanisms and feature fusion," *Forests*, vol. 14, no. 3, p. 619, 2023. https://doi.org/10.3390/f14030619
[19] Z. Xue, R. Xu, D. Bai, and H. Lin, "YOLO-tea: A tea disease detection model improved by YOLOv5," *Forests*, vol. 14, no. 2, p. 415, 2023. https://doi.org/10.3390/f14020415
[20] Y. Li, A. Li, X. Li, and D. Liang, "Detection and identification of peach leaf diseases based on YOLO v5 improved model," in *Proceedings of the 5th International Conference on Control and Computer Vision*, 2022, pp. 79-84, doi: https://doi.org/10.1145/3561613.3561626.
[21] B. Yan, P. Fan, X. Lei, Z. Liu, and F. Yang, "A real-time apple targets detection method for picking robot based on improved YOLOv5," *Remote Sensing*, vol. 13, no. 9, p. 1619, 2021. https://doi.org/10.3390/rs13091619
[22] Z. Chen *et al.*, "Plant disease recognition model based on improved YOLOv5," *Agronomy*, vol. 12, no. 2, p. 365, 2022. https://doi.org/10.3390/agronomy12020365
[23] Y. Wang, R. Xu, D. Bai, and H. Lin, "Integrated Learning-Based Pest and Disease Detection Method for Tea Leaves," *Forests*, vol. 14, no. 5, p. 1012, 2023. https://doi.org/10.3390/f14051012
[24] J. Qi *et al.*, "An improved YOLOv5 model based on visual attention mechanism: Application to recognition of tomato virus disease," *Computers and electronics in agriculture*, vol. 194, p. 106780, 2022. https://doi.org/10.1016/j.compag.2022.106780
[25] R. Li and Y. Wu, "Improved YOLO v5 wheat ear detection algorithm based on attention mechanism," *electronics*, vol. 11, no. 11, p. 1673, 2022. https://doi.org/10.3390/electronics11111673
[26] L. Wang, Y. Zhao, Z. Xiong, S. Wang, Y. Li, and Y. Lan, "Fast and precise detection of litchi fruits for yield estimation based on the improved YOLOv5 model," *Frontiers in Plant Science*, vol. 13, p. 965425, 2022. https://doi.org/10.3389/fpls.2022.965425
[27] Z. Ye *et al.*, "Recognition of terminal buds of densely-planted Chinese fir seedlings using improved YOLOv5 by integrating attention mechanism," *Frontiers in Plant Science*, vol. 13, p. 991929, 2022. https://doi.org/10.3389/fpls.2022.991929
[28] H. Zang *et al.*, "Detection method of wheat spike improved YOLOv5s based on the attention mechanism," *Frontiers in Plant Science*, vol. 13, p. 993244, 2022. https://doi.org/10.3389/fpls.2022.993244
[29] X. Song *et al.*, "Real-time determination of flowering period for field wheat based on improved YOLOv5s model," *Frontiers in Plant Science*, vol. 13, p. 1025663, 2023. https://doi.org/10.3389/fpls.2022.1025663
[30] Y. Zhao, Y. Yang, X. Xu, and C. Sun, "Precision detection of crop diseases based on improved YOLOv5 model," *Frontiers in Plant Science*, vol. 13, p. 1066835, 2023. https://doi.org/10.3389/fpls.2022.1066835
[31] A. Wang, T. Peng, H. Cao, Y. Xu, X. Wei, and B. Cui, "TIA-YOLOv5: An improved YOLOv5 network for real-time detection of crop and weed in the field," *Frontiers in Plant Science*, vol. 13, p. 1091655, 2022. https://doi.org/10.3389/fpls.2022.1091655
[32] H. He, X. Ma, H. Guan, F. Wang, and P. Shen, "Recognition of soybean pods and yield prediction based on improved deep learning model," *Frontiers in Plant Science*, vol. 13, p. 1096619, 2023. https://doi.org/10.3389/fpls.2022.1096619
[33] S. Zhu, W. Ma, J. Wang, M. Yang, Y. Wang, and C. Wang, "EADD-YOLO: An efficient and accurate disease detector for apple leaf using improved lightweight YOLOv5," *Frontiers in Plant Science*, vol. 14, p. 1120724, 2023. https://doi.org/10.3389/fpls.2023.1120724
[34] J. Lin *et al.*, "Improved YOLOX-Tiny network for detection of tobacco brown spot disease," *Frontiers in Plant Science*, vol. 14, p. 1135105, 2023. https://doi.org/10.3389/fpls.2023.1135105
[35] O. M. Lawal, S. Zhu, and K. Cheng, "An improved YOLOv5s model using feature concatenation with attention mechanism for real-time fruit detection and counting," *Frontiers in Plant Science*, vol. 14, 2023. https://doi.org/10.3389/fpls.2023.1153505
[36] D. Tzutalin, "Tzutalin/labelImg," *Github*, 2015.

[37] J. Qian and H. Lin, "A forest fire identification system based on weighted fusion algorithm," *Forests*, vol. 13, no. 8, p. 1301, 2022. https://doi.org/10.3390/f13081301

[38] K. S. Kouassi, M. Diarra, E. Kouassi Hilaire, and K. B. Jean-Claude, "Detection of cocoa pod diseases using a hybrid feature extractor combining CNN and vision transformer with dual classifier," *Edelweiss Applied Science and Technology*, vol. 9, no. 1, pp. 668-681, 2025. https://doi.org/10.55214/25768484.v9i1.4209

[39] S. Ozbay and O. Tunc, "Deep learning in analysing paranasal sinuses," *Elektronika ir Elektrotechnika*, vol. 28, no. 3, pp. 65-70, 2022. https://doi.org/10.5755/j02.eie.31133