


Article

Plant Disease Recognition Model Based on Improved YOLOv5

Zhaoyi Chen ¹, Ruhui Wu ², Yiyan Lin ¹, Chuyu Li ¹, Siyu Chen ¹, Zhineng Yuan ², Shiwei Chen ²
and Xiangjun Zou ^{1,3,*} 

¹ College of Engineering, South China Agricultural University, Guangzhou 510642, China; chenzhaoyi@stu.scau.edu.cn (Z.C.); linyiyan@stu.scau.edu.cn (Y.L.); lichuyu@stu.scau.edu.cn (C.L.); darkdrgonflycsy@gmail.com (S.C.)

² Guangdong Agribusiness Tropical Agriculture Institute Co., Ltd., Guangzhou 511365, China; 15595625042@163.com (R.W.); yuanzhin7@gmail.com (Z.Y.); hzchenmeng@163.com (S.C.)

³ Foshan-Zhongke Innovation Research Institute of Intelligent Agriculture, Foshan 528010, China

* Correspondence: xjzou@scau.edu.cn

Abstract: To accurately recognize plant diseases under complex natural conditions, an improved plant disease-recognition model based on the original YOLOv5 network model was established. First, a new InvolutionBottleneck module was used to reduce the numbers of parameters and calculations, and to capture long-distance information in the space. Second, an SE module was added to improve the sensitivity of the model to channel features. Finally, the loss function ‘Generalized Intersection over Union’ was changed to ‘Efficient Intersection over Union’ to address the former’s degeneration into ‘Intersection over Union’. These proposed methods were used to improve the target recognition effect of the network model. In the experimental phase, to verify the effectiveness of the model, sample images were randomly selected from the constructed rubber tree disease database to form training and test sets. The test results showed that the mean average precision of the improved YOLOv5 network reached 70%, which is 5.4% higher than that of the original YOLOv5 network. The precision values of this model for powdery mildew and anthracnose detection were 86.5% and 86.8%, respectively. The overall detection performance of the improved YOLOv5 network was significantly better compared with those of the original YOLOv5 and the YOLOX_nano network models. The improved model accurately identified plant diseases under natural conditions, and it provides a technical reference for the prevention and control of plant diseases.

Keywords: plant diseases recognition; YOLOv5; InvolutionBottleneck; SE module; EIOU



Citation: Chen, Z.; Wu, R.; Lin, Y.; Li, C.; Chen, S.; Yuan, Z.; Chen, S.; Zou, X. Plant Disease Recognition Model Based on Improved YOLOv5.

Agronomy **2022**, *12*, 365.

<https://doi.org/10.3390/agronomy12020365>

agronomy12020365

Academic Editor: Imre J. Holb

Received: 24 December 2021

Accepted: 29 January 2022

Published: 31 January 2022

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Agricultural production is an indispensable part of a nation’s economic development. Crops are affected by climate, which may make them susceptible to pathogen infection during the growth period, resulting in reduced production. In severe cases, the leaves fall off early and the plants die. To reduce the economic losses caused by diseases, it is necessary to properly diagnose plant diseases. Currently, two methods are used: expert diagnoses and pathogen analyses. The former refers to plant protection experts, with years of field production and real-time investigatory experience, diagnosing the extent of plant lesions. This method relies highly on expert experience and is prone to subjective differences and a low accuracy [1]. The latter involves the cultivation and microscopic observation of pathogens. This method has a high diagnostic accuracy rate, but it is time consuming, and the operational process is cumbersome, making it not suitable for field detection [2,3].

In recent years, the rapid development of machine vision and artificial intelligence has accelerated the process of engineering intelligence in various fields, and machine vision technology has also been rapidly improved in industrial, agricultural and other complex scene applications [4–9]. In response to the plant disease detection problem, disease

detection methods based on visible light and near-infrared spectroscopic digital images have been widely used. Near-infrared spectroscopic and hyperspectral images contain continuous spectral information and provide information on the spatial distributions of plant diseases. Consequently, they have become the preferred technologies of many researchers [10–13]. However, the equipment for acquiring spectral images is expensive and difficult to carry; therefore, this technology cannot be widely applied. The acquisition of visible light images is relatively simple and can be achieved using various ordinary electronic devices, such as digital cameras and smart phones, which greatly reduces the challenges of visible light image-recognition research [14,15].

Because of the need for real-time monitoring and sharing of crop growth information, visible light image recognition has been successfully applied to the field of plant disease detection in recent years [16–20]. A variety of traditional image-processing methods have been applied. First, the images are segmented, then the characteristics of plant diseases are extracted and, finally, the diseases are classified. Shrivastava et al. [21] proposed an image-based rice plant disease classification approach using color features only, and it successfully classifies rice plant diseases using a support vector machine classifier. Alajas et al. [22] used a hybrid linear discriminant analysis and a decision tree to predict the percentage of damaged leaf surface on diseased grapevines, with an accuracy of 97.79%. Kianat et al. [23] proposed a hybrid framework based on feature fusion and selection techniques to classify cucumber diseases. They first used the probability distribution-based entropy approach to reduce the extracted features, and then, they used the Manhattan distance-controlled entropy technique to select strong features. Mary et al. [24] used the merits of both the Gabor filter and the 2D log Gabor filter to construct an enhanced Gabor filter to extract features from the images of the diseased plant, and then, they used the k-nearest neighbor classifier to classify banana leaf diseases. Sugiarti et al. [25] combined the grey-level co-occurrence matrix extraction function with the naive Bayes classification to greatly improved the classification accuracy of apple diseases. Mukhopadhyay et al. [26] proposed a novel method based on image-processing technology, and they used the non-dominated sorting genetic algorithm to detect the disease area on tea leaves, with an average accuracy of 83%. However, visible light image-recognition based on traditional image processing technologies requires the artificial preprocessing of images and the extraction of disease features. The feature information is limited to shallow learning, and the generalization ability of new data sets needs to be improved.

However, deep learning methods are gradually being applied to agricultural research because they can automatically learn the deep feature information of images, and their speed and accuracy levels are greater than those of traditional algorithms [27–30]. Deep learning has also been applied to the detection of plant diseases from visible light images. Abbas et al. [31] proposed a deep learning-based method for tomato plant disease detection that utilizes the conditional generative adversarial network to generate synthetic images of tomato plant leaves. Xiang et al. [32] established a lightweight convolutional neural network-based network model with channel shuffle operation and multiple-size modules that achieved accuracy levels of 90.6% and 97.9% on a plant disease severity and PlantVillage datasets, respectively. Tan et al. [33] compared the recognition effects of deep learning networks and machine learning algorithms on tomato leaf diseases and found that the metrics of the tested deep learning networks are all better than those of the measured machine learning algorithms, with the ResNet34 network obtaining the best results. Alita et al. [34] used the EfficientNet deep learning model to detect plant leaf disease, and it was superior to other state-of-the-art deep learning models in terms of accuracy. Mishra et al. [35] developed a sine-cosine algorithm-based rider neural network and found that the detection performance of the classifier improved, achieving an accuracy of 95.6%. In summary, applying deep learning to plant disease detection has achieved good results.

As a result of climatic factors, rubber trees may suffer from a variety of pests and diseases, most typically powdery mildew and anthracnose, during the tender leaf stage.

Rubber tree anthracnose is caused by *Colletotrichum gloeosporioides* and *Colletotrichum acutatum* infections, whereas rubber tree powdery mildew is caused by *Oidiumheveae* [36,37]. The lesion features of the two diseases are highly similar, making them difficult to distinguish, which has a certain impact on the classification results of the network model. Compared with traditional image processing technology, deep convolutional neural networks have greater abilities to express abstract features and can obtain semantic information from complex images. Target detection algorithms based on deep learning can be divided into two categories, one-stage detection algorithms (such as the YOLO series) and two-stage detection algorithms (such as FasterR-CNN). The processing speeds of the former are faster than those of the latter, which makes them more suitable for the real-time detection of plant diseases in complex field environments.

In this paper, we report our attempts to address the above issues, as follows: First, we used convolutional neural networks to automatically detect rubber tree powdery mildew and anthracnose in visible light images, which has some practical benefits for the prevention and control of rubber tree diseases. Second, we focused on solving the existing difficulties in detecting rubber tree diseases using YOLOv5, and we further improved the detection accuracy of the model. Consequently, a rubber tree disease recognition model based on the improved YOLOv5 was established, with the aim of achieving the accurate classification and recognition of rubber tree powdery mildew and anthracnose under natural light conditions. The main contributions of our work are summarized below:

- (1) In the backbone network, the Bottleneck module in the C3 module was replaced with the InvolutionBottleneck module that reduced the number of calculations in the convolutional neural network;
- (2) The SE module was added to the last layer of the backbone network to fuse disease characteristics in a weighted manner;
- (3) The existing loss function Generalized Intersection over Union (GIOU) in YOLOv5 was replaced by the loss function Efficient Intersection over Union (EIOU), which takes into account differences in target frame width, height and confidence;
- (4) The proposed model can realize the accurate and automatic identification of rubber tree diseases in visible light images, which has some significance for the prevention and control of rubber tree diseases.

The remainder of this article is organized as follows: In Section 2, we give a brief review of the original YOLOv5 model, and the improved YOLOv5 model is proposed. In Section 3, we list the experimental materials and methods. Experiments and analyses of the results are covered in Section 4. Finally, the conclusions are summarized in Section 5.

2. Principle of the Detection Algorithm

2.1. YOLOv5 Network Module

YOLOv5 [38] is a one-stage target recognition algorithm proposed by Glenn Jocher in 2020. On the basis of differences in network depth and width, YOLOv5 can be divided into four network model versions: YOLOv5s, YOLOv5m, YOLOv5l and YOLOv5x. Among them, the YOLOv5s network has the fastest calculation speed, but the average precision is the lowest, whereas the YOLOv5x network has the opposite characteristics. The model size of the YOLOv5 network is approximately one-tenth that of the YOLOv4 network. It has faster recognition and positioning speeds, and the accuracy is no less than that of YOLOv4. The YOLOv5 network is composed of three main components: Backbone, Neck and Head. After the image is inputted, Backbone aggregates and forms image features on different image granularities. Then, Neck stitches the image features and transmits them to the prediction layer, and Head predicts the image features to generate bounding boxes and predicted categories. The YOLOv5 network uses *GIOU* as the network loss function, as shown in Equation (1).

$$GIOU = IOU - \frac{|C - (A \cup B)|}{|C|} \quad (1)$$

where $A, B \subseteq S \subseteq R^n$ represent two arbitrary boxes, C represents the smallest convex box, $C \subseteq S \subseteq R^n$, enclosing both A and B and $IOU = |A \cap B| / |A \cup B|$.

When the input network predicts image features, the optimal target frame is filtered by combining the loss function $GIOU$ and the non-maximum suppression algorithm.

2.2. Improved YOLOv5 Network Construction

2.2.1. InvolutionBottleneck Module Design

In the Backbone, the Bottleneck module in the C3 module was replaced with the InvolutionBottleneck module. The two inherent principles of standard convolution kernels are spatial-agnostic and channel-specific, whereas those of involution [39] are the opposite. Convolutional neural networks usually increase the receptive field by stacking convolution kernels of different sizes. Using different kernel calculations for each channel causes a substantial increase in the number of calculations. In the Backbone, the InvolutionBottleneck module was used to replace the Bottleneck module, which alleviated the kernel redundancy by sharing the involution kernel along the channel dimension, and this is beneficial for capturing the long-distance information of the spatial range and reducing the number of network parameters. The output feature map Y of the Involution-convolution operation is defined as shown in Equations (2) and (3).

$$H_{i,j} = \phi(X_{\psi_{i,j}}) \quad (2)$$

$$Y_{i,j,k} = \sum_{(u,v) \in \Delta k} H_{i,j,u+[K/2],v+[K/2],[kG/C]} X_{i+u,j+v,k} \quad (3)$$

where C represents the calculation channel, ϕ represents the generation function of the convolution kernel and $\psi_{i,j}$ represents the index to the pixel set $H_{i,j}$. The convolution kernel $H_{i,j,\dots,g} \in R^{K \times K}$ ($g = 1, 2, \dots, G$) was specifically customized for the pixel $X_{i,j} \in R^C$ located at the corresponding coordinates (i, j) , but it is shared on the channel. G represents the number of groups sharing the same convolution kernel. The size of the convolution kernel depends on the size of the input feature map.

2.2.2. SE Module Design

The squeeze-and-excitation network [40] is a network model proposed by Hu et al. (2017) that focuses on the relationship between channels. It aims to learn each image feature according to the loss function, increase the weight of effective image features and reduce the weight of invalid or ineffective image features, thereby training the network model to produce the best results. The SE modules with different structures are shown in Figure 1.

The SE module is a calculation block that can be built on the transformation between the input feature vector X and the output feature map u , and the transformation relationship is shown in Equation (4):

$$u_c = v_c * X = \sum_{s=1}^{C'} v_c^s * x^s \quad (4)$$

where $*$ represents convolution, $v_c = [v_c^1, v_c^2, \dots, v_c^{C'}]$, $X = [X^1, X^2, \dots, X^{C'}]$ and $u_c \in R^{H \times W}$. v_c^s represents a 2D spatial kernel, which denotes a single channel of v_c that acts on the corresponding channel of X .

In this paper, the SE module was added to the last layer of the Backbone, allowing it to merge the image features of powdery mildew and anthracnose in a weighted manner, thereby improving the network performance at a small cost.

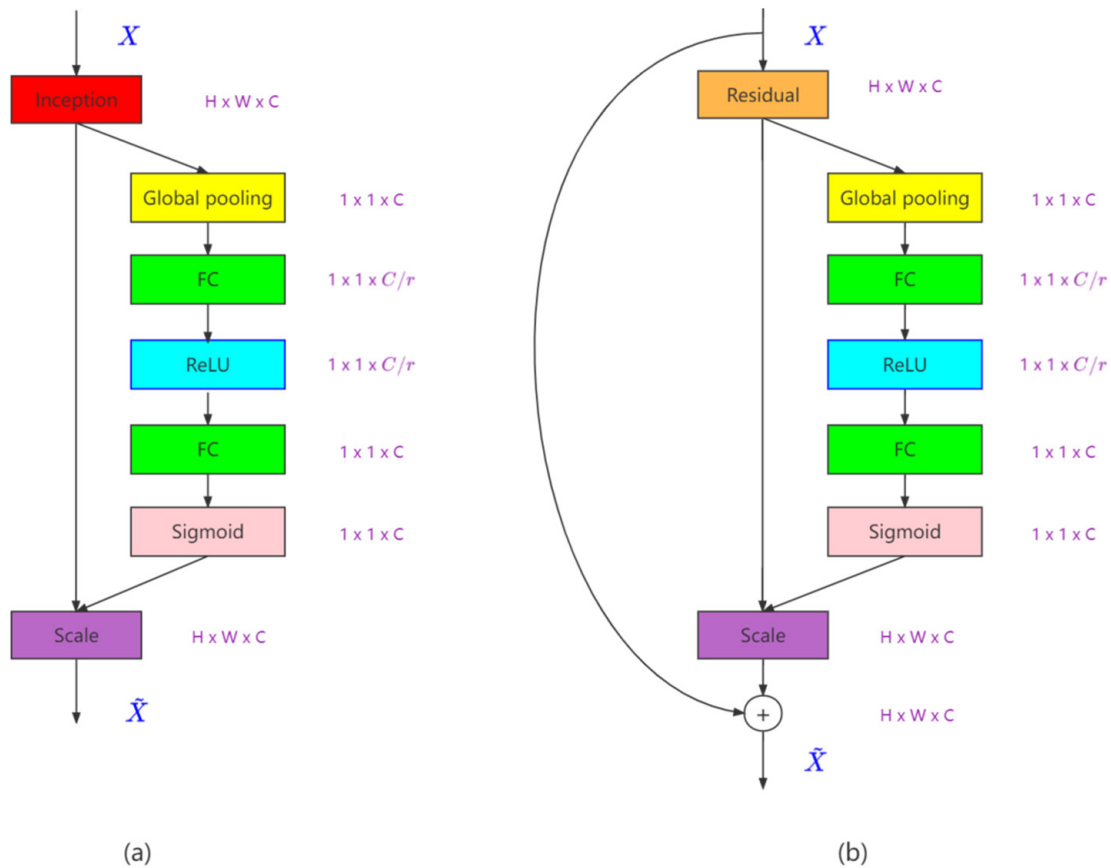


Figure 1. SE modules with different structures. (a) SE module with Inception structure; (b) SE module with Residual structure.

2.2.3. Loss Function Design

The loss function was changed from *GIOU* to *EIOU* [41]. The *GIOU* function was proposed on the basis of the *IOU* function. It solves the problem of the *IOU* not being able to reflect how the two boxes intersect. However, if the anchor and target boxes are part of a containment relationship, then *GIOU* will still degenerate into *IOU*. Therefore, we changed the loss function *GIOU* to *EIOU*. *EIOU* was obtained on the basis of complete-*IOU* loss (*CIOU*), and it not only takes into account the central point distance and the aspect ratio, but also the true discrepancies in the target and anchor boxes' widths and heights. The *EIOU* loss function directly minimizes these discrepancies and accelerates model convergence. The *EIOU* loss function is shown in Equation (5).

$$L_{EIOU} = L_{IOU} + L_{dis} + L_{asp} = 1 - IOU + \frac{\rho^2(b, b^{st})}{C^2} + \frac{\rho^2(w, w^{st})}{C_w^2} + \frac{\rho^2(h, h^{st})}{C_h^2} \quad (5)$$

where C_w and C_h represent the width and height, respectively, of the smallest enclosing box covering the two boxes; b and b^{st} represent the central points of the predicted and target boxes, respectively; ρ represents the Euclidean distance; C represents the diagonal length of the smallest enclosing box covering the two boxes. The loss function *EIOU* is divided into three parts: the *IOU* loss L_{IOU} , the distance loss L_{dis} and the aspect loss L_{asp} .

Combined with the InvolutionBottleneck and the SE modules, the whole improved YOLOv5 network model framework is constructed, as shown in Figure 2.

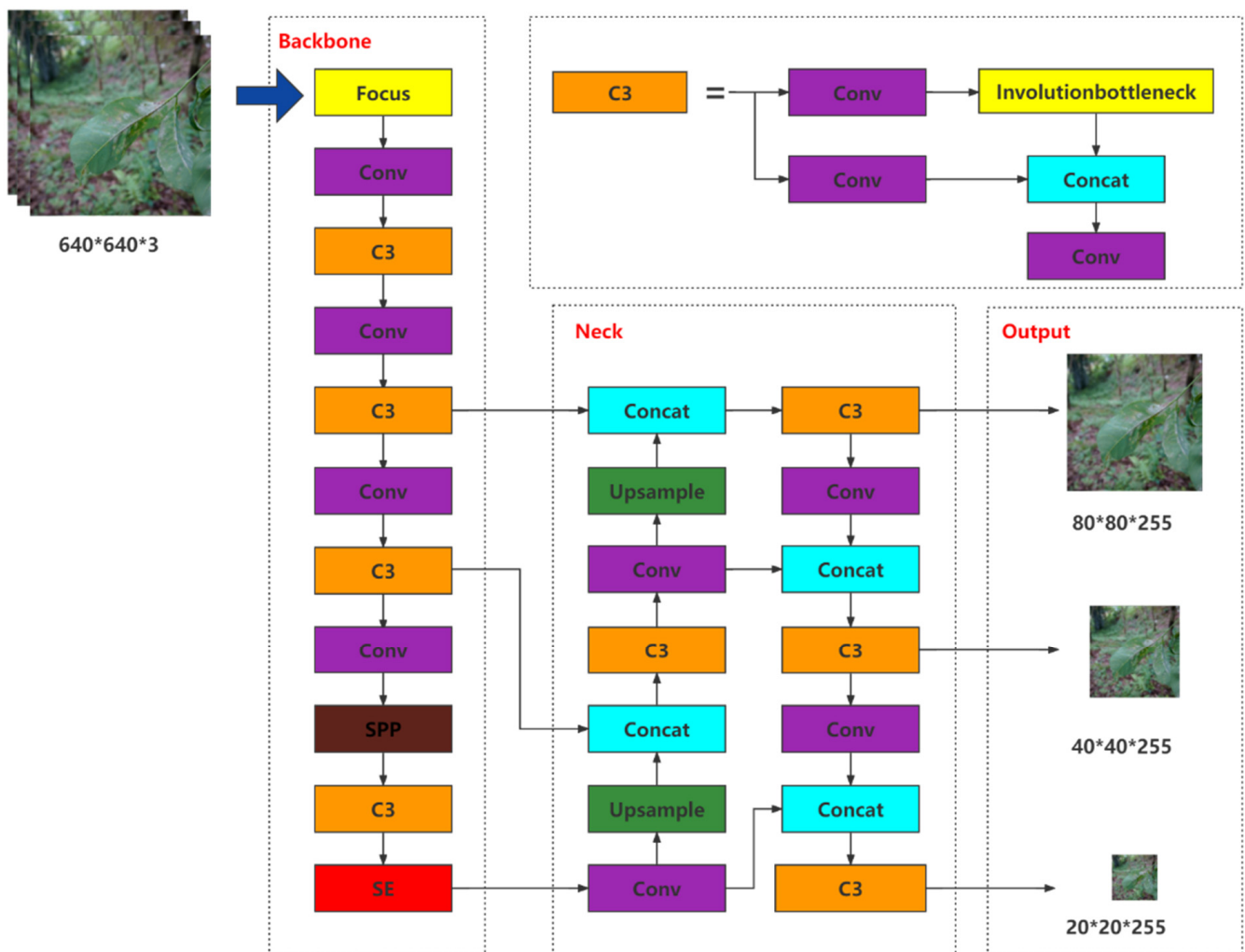


Figure 2. The improved YOLOv5 network model structure.

3. Materials and Methods

3.1. Experimental Materials

The images of rubber tree diseases were collected from a rubber plantation in Shengli State Farm, Maoming City, China. It is located at $22^{\circ}6' N$, $110^{\circ}80' E$, with an altitude of 34–69 m, an average annual precipitation of 1698.1 mm and an annual average temperature of 19.9 – 26.5 °C. The high humidity and warm climate are conducive to widespread epidemics of powdery mildew and anthracnose. To ensure the representativeness of the image set, they were collected under natural light conditions. A Sony digital ILCE-7m3 camera was used to photograph powdery mildew and anthracnose of rubber leaves at different angles, with an image resolution of 6000×4000 pixels. There were 2375 images in the rubber tree disease database, including 1203 powdery mildew images and 1172 anthracnose images, which were used for the training and testing of disease recognition models. We identified these two diseases with the guidance of plant protection experts. Images of these rubber tree diseases are shown in Figure 3.

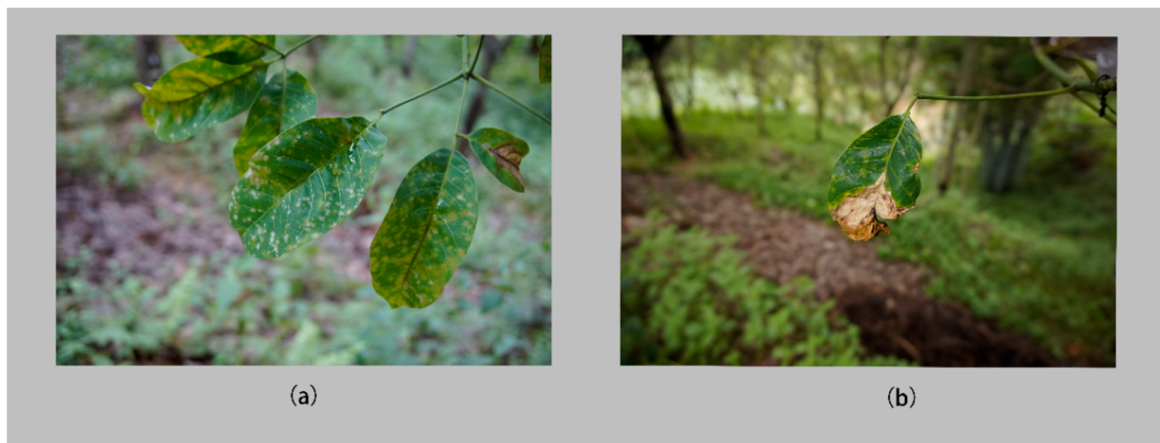


Figure 3. Rubber tree diseases images. (a) Powdery mildew image; (b) Anthracnose image.

3.2. Data Preprocessing

Before the images were inputted into the improved YOLOv5 network model, the mosaic data enhancement method was used to expand the image set. The images were spliced using several methods, such as random scaling, random cropping and random arrangement, which not only expanded the image set, but also improved the detection of small targets. In addition, before training the model, adaptive scaling and filling operations were performed on the images of rubber tree diseases, and the input image size was normalized to 640×640 pixels. The preprocessing results are shown in Figure 4.

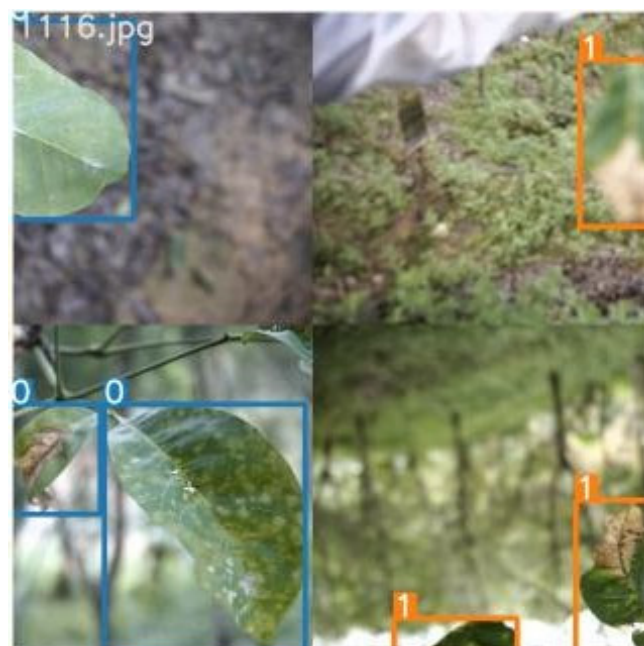


Figure 4. Image preprocessing result.

3.3. Experimental Equipment

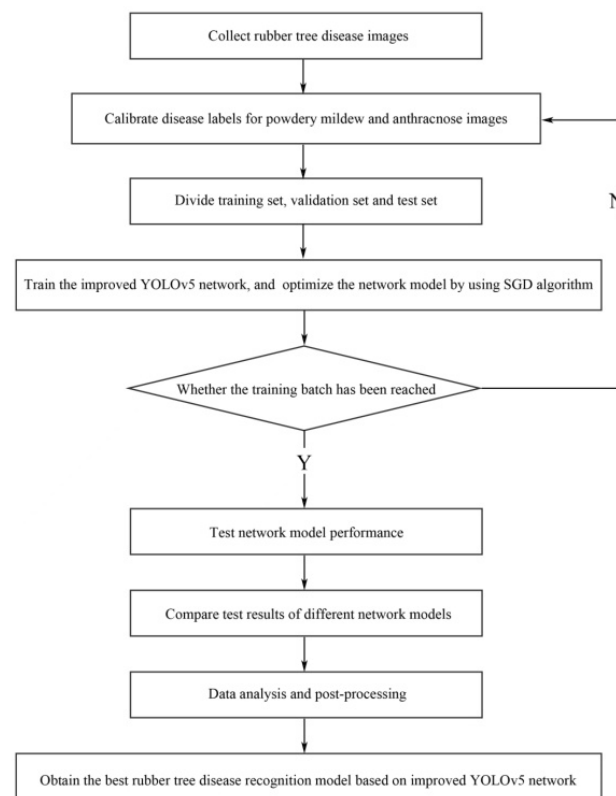
A desktop computer was used as the processing platform, the operating system was Ubuntu 18.04 and the Pytorch framework and the YOLOv5 environment were built in the Anaconda3 environment. The program was written in Python 3.8, and the CUDA version was 10.1. For hardware, the processor was an Intel Core i3-4150, the main frequency was 3.5 GHz, the memory was 3G and the graphics card was a GeForce GTX 1060 6G. The specific configurations are provided in Table 1.

Table 1. Test environment setting.

Parameter	Configuration
Operating system	Ubuntu 18.04
Deep learning framework	Pytorch 1.8
Programming language	Python 3.8
GPU accelerated environment	CUDA 10.1
GPU	GeForce GTX 1060 6G
CPU	Intel(R) Core(TM) i3-4150 CPU @ 3.50GHz

3.4. Experimental Process

First, the manual labeling method was used to mark each rubber disease image for powdery mildew or anthracnose to obtain training label images, and then the disease image set was divided at a 4:1:1 ratio into training, validation and test sets. The training set was inputted into the improved YOLOv5 network of different structures for training. The training process was divided into 80 batches, with each batch containing 96 images. The Stochastic Gradient Descent algorithm was used to optimize the network model during the training process, and the optimal network weight was obtained after the training was completed. Subsequently, the performance of the network model was determined using the test set and compared with the test results of the original YOLOv5 and the YOLOX_nano networks. The network model with the best result was selected as the rubber tree disease recognition model. The test process is shown in Figure 5.

**Figure 5.** Test flow chart.

4. Results and Analysis

4.1. Convergence Results of the Network Model

The training and verification sets were inputted into the network for training. After 80 batches of training, the loss function value curves of the training and verification sets were determined (Figure 6), and they included the detection frame loss, the detection object loss and the classification loss.

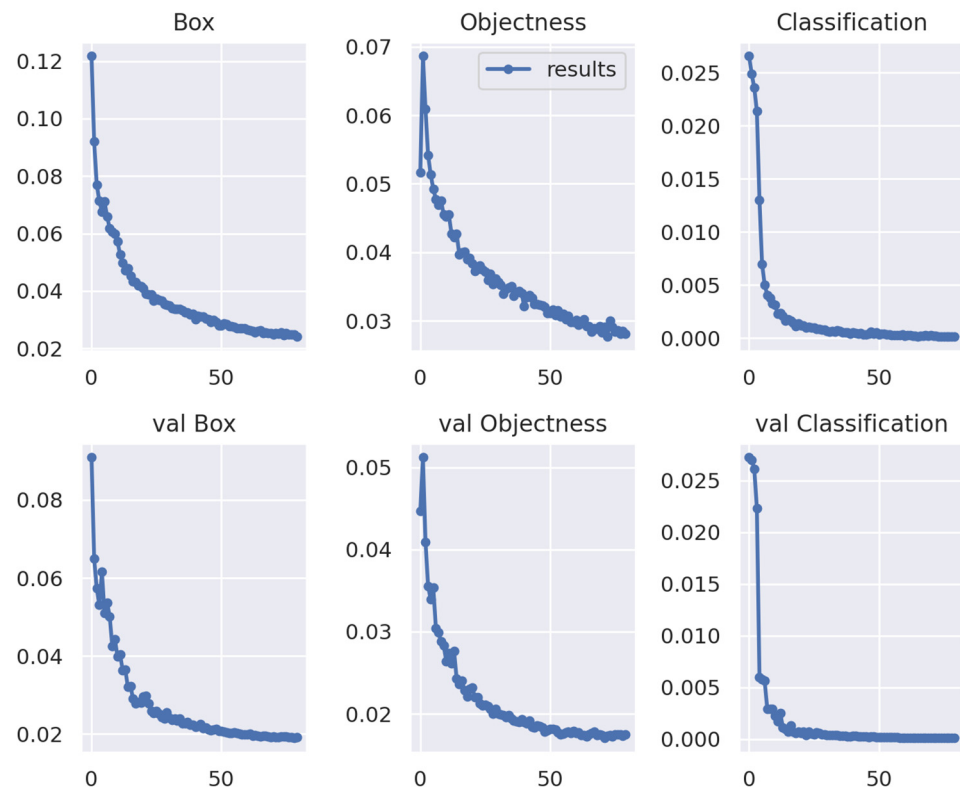


Figure 6. Convergence of the loss functions of training and validation sets.

The loss of the detection frame indicates whether an algorithm can locate the center point of an object well and whether the detection target is covered by the predicted bounding box. The smaller the loss function value, the more accurate the prediction frame. The object loss function is essentially a measure of the probability that the detection target exists in the region of interest. The smaller the value of the loss function, the higher the accuracy. The classification loss represents the ability of the algorithm to correctly predict a given object category. The smaller the loss value, the more accurate the classification.

As shown in Figure 6, the loss function value had a downward trend during the training process, the Stochastic Gradient Descent algorithm optimized the network and the network weight and other parameters were constantly updated. Before the training batch reached 20, the loss function value dropped rapidly, and the accuracy, recall rate and average accuracy rapidly improved. The network continued to iterate. When the training batch reached approximately 20, the decrease in the loss function value gradually slowed. Similarly, the increases in parameters such as average accuracy also slowed. When the training batch reached 80, the loss curves of the training and validation sets showed almost no downward trends, and other index values also tended to have stabilized. The network model basically reached the convergence state, and the optimal network weight was obtained at the end of training.

4.2. Verification of the Network Model

To evaluate the detection performance of the improved YOLOv5 network, it was crucial to use appropriate evaluation metrics for each problem. The precision, recall, average precision and mean average precision were used as the evaluation metrics, and they were respectively defined as follows:

$$P = \frac{TP}{TP + FP} \quad (6)$$

$$R = \frac{TP}{TP + FN} \quad (7)$$

$$AP_i = \int_0^1 P(R)d(R) \quad (8)$$

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (9)$$

where TP represents the number of positive samples that are correctly detected, FP represents the number of negative samples that are falsely detected and FN represents the number of positive samples that are not detected.

In total, 200 powdery mildew images and 200 anthracnose images were randomly selected as the test set and inputted into the improved YOLOv5 network for testing. The test results were compared with those of the original YOLOv5 and the YOLOX_nano networks. The comparison results are shown in Figure 7.

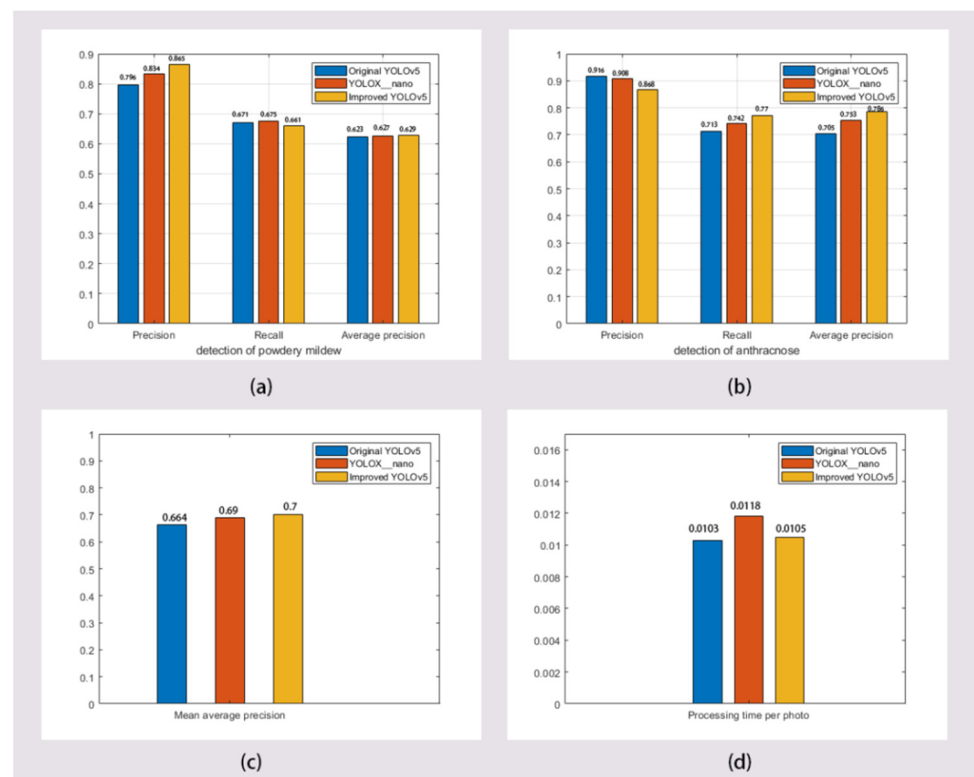


Figure 7. Performance comparison of all the network models. (a) Powdery mildew recognition results; (b) Anthracnose recognition results; (c) Mean average precision; (d) Processing times per photo.

As shown in Figure 7, the detection performance of the improved YOLOv5 network was better than that of the original YOLOv5 network for each of the tested two diseases of rubber trees. For the detection of powdery mildew, precision increased by 8.7% and average precision increased by 1%; however, recall decreased by 1.5%. For the detection of anthracnose, average precision increased by 9.2% and recall increased by 9.3%; however, precision decreased by 5.2%. Overall, the mean average precision increased by 5.4%. The improved YOLOv5 network achieved 86.5% and 86.8% precision levels for the detection of powdery mildew and anthracnose, respectively. In summary, the improved YOLOv5 network's performance was greatly enhanced compared with that of the original YOLOv5 network; consequently, it achieved more accurate rubber tree disease identification and location functions. As shown in Figure 7, the performance of the improved YOLOv5 network was better than those of the original YOLOv5 and YOLOX_nano networks for the detection of the two diseases of rubber trees. Compared with the original YOLOv5

network, the precision of powdery mildew detection increased by 8.7% and the average precision increased by 1%; however, recall decreased by 1.5%. The average precision of anthracnose detection increased by 9.2% and recall increased by 9.3%; however, precision decreased by 5.2%. Overall, the mean average precision increased by 5.4%. Compared with the YOLOX_nano network, the precision of powdery mildew detection increased by 3.7% and the average precision increased by 0.3%; however, recall decreased by 2%. The precision of anthracnose detection increased by 4.4% and recall increased by 3.8%; however, precision decreased by 4.4%. Overall, the mean average precision increased by 1.4%. The improved YOLOv5 network achieved 86.5% and 86.8% levels for the detection of powdery mildew and anthracnose, respectively. In summary, the improved YOLOv5 network's performance was greatly enhanced compared with those of the original YOLOv5 and YOLOX_nano networks; consequently, it more accurately locates and identifies rubber tree diseases.

4.3. Comparison of Recognition Results

The original YOLOv5, the YOLOX_nano and the improved YOLOv5 networks were used to detect two kinds of diseases of rubber trees to verify the actual classification and recognition effects of the improved network. A comparison of test results is shown in Figure 8.

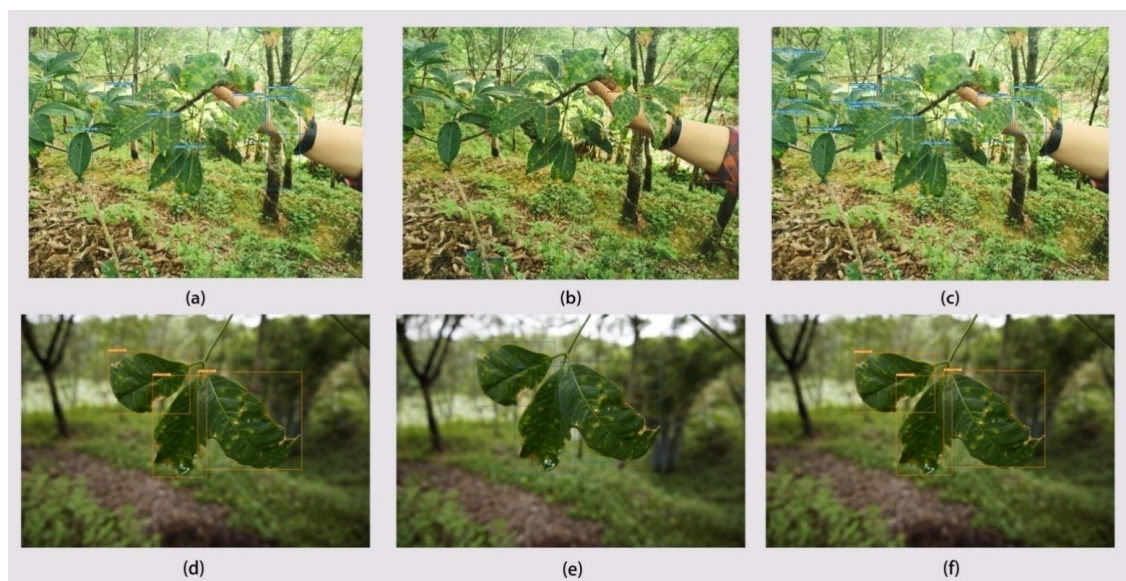


Figure 8. Comparison of the recognition effects of all the network models. (a–c) Powdery mildew recognition effects of the (a) original YOLOv5; (b) YOLOX_nano; and (c) improved YOLOv5 network models; (d–f) Anthracnose recognition effects of the (d) original YOLOv5; (e) YOLOX_nano; and (f) improved YOLOv5 network models.

As shown in Figure 8, compared with the other networks, the improved network significantly improved the detection of powdery mildew, including obscured diseased leaves. Additionally, the recognition effect of the YOLOX_nano network for powdery mildew is better than that of the original YOLOv5 network. For the detection of anthracnose, the recognition effects of the three networks were similar, with all three effectively detecting anthracnose. Therefore, the effectiveness of the improved network for diseased leaves detection is generally better than those of the original YOLOv5 and the YOLOX_nano networks.

5. Conclusions

The detection and location of plant diseases in the natural environment are of great significance to plant disease control. In this paper, a rubber tree disease recognition model based on the improved YOLOv5 network was established. We replaced the Bottleneck

module with the InvolutionBottleneck module to achieve channel sharing within the group and reduce the number of network parameters. In addition, the SE module was added to the last layer of the Backbone for feature fusion, which improved network performance at a small cost. Finally, the loss function was changed from GIOU to EIOU to accelerate the convergence of the network model. According to the experimental results, the following conclusions can be drawn:

- (1) The model performance verification experiment showed that the rubber tree disease recognition model based on the improved YOLOv5 network achieved 86.5% precision for powdery mildew detection and 86.8% precision for anthracnose detection. In general, the mean average precision reached 70%, which is an increase of 5.4% compared with the original YOLOv5 network. Therefore, the improved YOLOv5 network more accurately identified and classified rubber tree diseases, and it provides a technical reference for the prevention and control of rubber tree diseases.
- (2) A comparison of the detection results showed that the performance of the improved YOLOv5 network was generally better than those of the original YOLOv5 and the YOLOX_nano networks, especially in the detection of powdery mildew. The problem of the missing obscured diseased leaves was improved.

Although the improved YOLOv5 network, as applied to rubber tree disease detection, achieved good results, the detection accuracy still needs to be improved. In future research, the network model structure will be further optimized to improve the network performance of the rubber tree disease recognition model.

Author Contributions: Conceptualization, Z.C. and X.Z.; methodology, Z.C.; software, Z.C.; validation, Z.C.; formal analysis, Z.C. and X.Z.; investigation, Z.C., X.Z., R.W., Y.L., C.L. and S.C. (Siyu Chen); resources, Z.C., X.Z., R.W., Z.Y. and S.C. (Shiwei Chen); data curation, Z.C.; writing—original draft preparation, Z.C.; writing—review and editing, Z.C. and X.Z.; visualization, Z.C.; supervision, X.Z.; project administration, X.Z. All authors have read and agreed to the published version of the manuscript.

Funding: The paper is funded by the No. 03 Special Project and the 5G Project of Jiangxi Province under Grant 20212ABC03A27 and the Key-area Research and Development Program of Guangdong Province under Grant 2019B020223003.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to the privacy policy of the organization.

Acknowledgments: The authors would like to thank the anonymous reviewers for their critical comments and suggestions for improving the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lin, G.; Tang, Y.; Zou, X.; Xiong, J.; Fang, Y. Color-, depth-, and shape-based 3D fruit detection. *Precis. Agric.* **2020**, *21*, 1–17. [[CrossRef](#)]
2. Joshi, R.C.; Kaushik, M.; Dutta, M.K.; Srivastava, A.; Choudhary, N. VirLeafNet: Automatic analysis and viral disease diagnosis using deep-learning in Vigna mungoplant. *Ecol. Inform.* **2021**, *61*, 101197. [[CrossRef](#)]
3. Buja, I.; Sabella, E.; Monteduro, A.G.; Chiriaco, M.S.; De Bellis, L.; Luvisi, A.; Maruccio, G. Advances in Plant Disease Detection and Monitoring: From Traditional Assays to In-Field Diagnostics. *Sensors* **2021**, *21*, 2129. [[CrossRef](#)] [[PubMed](#)]
4. Liu, S.; Liu, D.; Srivastava, G.; Polap, D.; Woźniak, M. Overview and methods of correlation filter algorithms in object tracking. *Complex Intell. Syst.* **2020**, *7*, 1895–1917. [[CrossRef](#)]
5. Tang, Y.; Chen, M.; Wang, C.; Luo, L.; Li, J.; Lian, G.; Zou, X. Recognition and localization methods for vision-based fruit picking robots: A review. *Front. Plant Sci.* **2020**, *11*, 510. [[CrossRef](#)]
6. Li, J.; Tang, Y.; Zou, X.; Lin, G.; Wang, H. Detection of fruit-bearing branches and localization of litchi clusters for vision-based harvesting robots. *IEEE Access* **2020**, *8*, 117746–117758. [[CrossRef](#)]
7. Wu, F.; Duan, J.; Chen, S.; Ye, Y.; Ai, P.; Yang, Z. Multi-Target Recognition of Bananas and Automatic Positioning for the Inflorescence Axis Cutting Point. *Front. Plant Sci.* **2021**, *12*, 705021. [[CrossRef](#)]
8. Wang, C.; Tang, Y.; Zou, X.; Luo, L.; Chen, X. Recognition and matching of clustered mature litchi fruits using binocular charge-coupled device (CCD) color cameras. *Sensors* **2017**, *17*, 2564. [[CrossRef](#)]

9. Luo, L.; Liu, W.; Lu, Q.; Wang, J.; Wen, W.; Yan, D.; Tang, Y. Grape Berry Detection and Size Measurement Based on Edge Image Processing and Geometric morphology. *Machines* **2021**, *9*, 233. [[CrossRef](#)]
10. Gui, J.; Fei, J.; Wu, Z.; Fu, X.; Diakite, A. Grading method of soybean mosaic disease based on hyperspectral imaging technology. *Inf. Process. Agric.* **2021**, *8*, 380–385. [[CrossRef](#)]
11. Luo, L.; Chang, Q.; Wang, Q.; Huang, Y. Identification and Severity Monitoring of Maize Dwarf Mosaic Virus Infection Based on Hyperspectral Measurements. *Remote Sens.* **2021**, *13*, 4560. [[CrossRef](#)]
12. Appeltans, S.; Pieters, J.G.; Mouazen, A.M. Detection of leek white tip disease under field conditions using hyperspectral proximal sensing and supervised machine learning. *Comput. Electron. Agric.* **2021**, *190*, 106453. [[CrossRef](#)]
13. Fazari, A.; Pellicer-Valero, O.J.; Gómez-Sanchis, J.; Bernardi, B.; Cubero, S.; Benalia, S.; Zimbalatti, G.; Blasco, J. Application of deep convolutional neural networks for the detection of anthracnose in olives using VIS/NIR hyperspectral images. *Comput. Electron. Agric.* **2021**, *187*, 106252. [[CrossRef](#)]
14. Shi, Y.; Huang, W.; Luo, J.; Huang, L.; Zhou, X. Detection and discrimination of pests and diseases in winter wheat based on spectral indices and kernel discriminant analysis. *Comput. Electron. Agric.* **2017**, *141*, 171–180. [[CrossRef](#)]
15. Phadikar, S.; Sil, J.; Das, A.K. Rice diseases classification using feature selection and rule generation techniques. *Comput. Electron. Agric.* **2013**, *90*, 76–85. [[CrossRef](#)]
16. Ahmed, N.; Asif, H.M.S.; Saleem, G. Leaf Image-based Plant Disease Identification using Color and Texture Features. *arXiv* **2021**, arXiv:2102.04515.
17. Singh, S.; Gupta, S.; Tanta, A.; Gupta, R. Extraction of Multiple Diseases in Apple Leaf Using Machine Learning. *Int. J. Image Graph.* **2021**, 2140009. [[CrossRef](#)]
18. Gadade, H.D.; Kirange, D.K. Machine Learning Based Identification of Tomato Leaf Diseases at Various Stages of Development. In Proceedings of the 2021 5th International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 8–10 April 2021.
19. Almadhor, A.; Rauf, H.; Lali, M.; Damaševičius, R.; Alouffi, B.; Alharbi, A. AI-Driven Framework for Recognition of Guava Plant Diseases through Machine Learning from DSLR Camera Sensor Based High Resolution Imagery. *Sensors* **2021**, *21*, 3830. [[CrossRef](#)]
20. Kundu, N.; Rani, G.; Dhaka, V.S.; Gupta, K.; Nayak, S.C.; Verma, S.; Ijaz, M.F.; Woźniak, M. IoT and Interpretable Machine Learning Based Framework for Disease Prediction in Pearl Millet. *Sensors* **2021**, *21*, 5386. [[CrossRef](#)]
21. Shrivastava, V.K.; Pradhan, M.K. Rice plant disease classification using color features: A machine learning paradigm. *J. Plant Pathol.* **2020**, *103*, 17–26. [[CrossRef](#)]
22. Alajas, O.J.; Concepcion, R.; Dadios, E.; Sybingco, E.; Mendigoria, C.H.; Aquino, H. Prediction of Grape Leaf Black Rot Damaged Surface Percentage Using Hybrid Linear Discriminant Analysis and Decision Tree. In Proceedings of the 2021 International Conference on Intelligent Technologies (CONIT), Hubli, India, 25–27 June 2021.
23. Kianat, J.; Khan, M.A.; Sharif, M.; Akram, T.; Rehman, A.; Saba, T. A joint framework of feature reduction and robust feature selection for cucumber leaf diseases recognition. *Optik* **2021**, *240*, 166566. [[CrossRef](#)]
24. Mary, N.A.B.; Singh, A.R.; Athisayamani, S. Classification of Banana Leaf Diseases Using Enhanced Gabor Feature Descriptor. In *Inventive Communication and Computational Technologies*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 229–242.
25. Sugiarti, Y.; Supriyatna, A.; Carolina, I.; Amin, R.; Yani, A. Model Naïve Bayes Classifiers For Detection Apple Diseases. In Proceedings of the 2021 9th International Conference on Cyber and IT Service Management (CITSM), Bengkulu, Indonesia, 22–23 September 2021.
26. Mukhopadhyay, S.; Paul, M.; Pal, R.; De, D. Tea leaf disease detection using multi-objective image segmentation. *Multimed. Tools Appl.* **2021**, *80*, 753–771. [[CrossRef](#)]
27. Chen, M.; Tang, Y.; Zou, X.; Huang, K.; Huang, Z.; Zhou, H.; Wang, C.; Lian, G. Three-dimensional perception of orchard banana central stock enhanced by adaptive multi-vision technology. *Comput. Electron. Agric.* **2020**, *174*, 105508. [[CrossRef](#)]
28. Li, Q.; Jia, W.; Sun, M.; Hou, S.; Zheng, Y. A novel green apple segmentation algorithm based on ensemble U-Net under complex orchard environment. *Comput. Electron. Agric.* **2021**, *180*, 105900. [[CrossRef](#)]
29. Cao, X.; Yan, H.; Huang, Z.; Ai, S.; Xu, Y.; Fu, R.; Zou, X. A Multi-Objective Particle Swarm Optimization for Trajectory Planning of Fruit Picking Manipulator. *Agronomy* **2021**, *11*, 2286. [[CrossRef](#)]
30. Anagnostis, A.; Tagarakis, A.C.; Asiminari, G.; Papageorgiou, E.; Kateris, D.; Moshou, D.; Bochtis, D. A deep learning approach for anthracnose infected trees classification in walnut orchards. *Comput. Electron. Agric.* **2021**, *182*, 105998. [[CrossRef](#)]
31. Abbas, A.; Jain, S.; Gour, M.; Vankudothu, S. Tomato plant disease detection using transfer learning with C-GAN synthetic images. *Comput. Electron. Agric.* **2021**, *187*, 106279. [[CrossRef](#)]
32. Xiang, S.; Liang, Q.; Sun, W.; Zhang, D.; Wang, Y. L-CSMS: Novel lightweight network for plant disease severity recognition. *J. Plant Dis. Prot.* **2021**, *128*, 557–569. [[CrossRef](#)]
33. Tan, L.; Lu, J.; Jiang, H. Tomato Leaf Diseases Classification Based on Leaf Images: A Comparison between Classical Machine Learning and Deep Learning Methods. *AgriEngineering* **2021**, *3*, 542–558. [[CrossRef](#)]
34. Atila, Ü.; Uçar, M.; Akyol, K.; Uçar, E. Plant leaf disease classification using EfficientNet deep learning model. *Ecol. Inform.* **2021**, *61*, 101182. [[CrossRef](#)]
35. Mishra, M.; Choudhury, P.; Pati, B. Modified ride-NN optimizer for the IoT based plant disease detection. *J. Ambient Intell. Humaniz. Comput.* **2021**, *12*, 691–703. [[CrossRef](#)]

36. Liu, X.; Li, B.; Cai, J.; Zheng, X.; Feng, Y.; Huang, G. Colletotrichum species causing anthracnose of rubber trees in China. *Sci. Rep.* **2018**, *8*, 10435. [[CrossRef](#)] [[PubMed](#)]
37. Wu, H.; Pan, Y.; Di, R.; He, Q.; Rajaofera, M.J.N.; Liu, W.; Zheng, F.; Miao, W. Molecular identification of the powdery mildew fungus infecting rubber trees in China. *Forest Pathol.* **2019**, *49*, e12519. [[CrossRef](#)]
38. Jocher, G.; Stoken, A.; Borovec, J.; Christopher, S.T.; Laughing, L.C. Ultralytics/yolov5: V4.0-nn.SILU() Activations, Weights & Biases Logging, Pytorch Hub Integration. *Zenodo* **2021**. [[CrossRef](#)]
39. Li, D.; Hu, J.; Wang, C.; Li, X.; She, Q.; Zhu, L.; Zhang, T.; Chen, Q. Involution: Inverting the inherence of convolution for visual recognition. *arXiv* **2021**, arXiv:2103.06255.
40. Hu, J.; Shen, L.; Sun, G. Squeeze-and-excitation networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 7132–7141.
41. Zhang, Y.-F.; Ren, W.; Zhang, Z.; Jia, Z.; Wang, L.; Tan, T. Focal and Efficient IOU Loss for Accurate Bounding Box Regression. *arXiv* **2021**, arXiv:2101.08158.