Article

Detecting Individual Plants Infected with Pine Wilt Disease Using Drones and Satellite Imagery: A Case Study in Xianning, China

Peihua Cai [1,2,†], Guanzhou Chen [2,†], Haobo Yang [2], Xianwei Li [2], Kun Zhu [3], Tong Wang [2], Puyun Liao [2], Mengdi Han [1,2], Yuanfu Gong [2,4], Qing Wang [1] and Xiaodong Zhang [2,*]

1   School of Geosciences, Yangtze University, Wuhan 430010, China
2   State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China; cgz@whu.edu.cn (G.C.)
3   Institute of Geospatial Information, PLA Strategic Support Force Information Engineering University, Zhengzhou 450001, China
4   Hubei Institute of Land Surveying and Mapping, Wuhan 430034, China
*   Correspondence: zxdlmars@whu.edu.cn
†   These authors contributed equally to this work.

Abstract: In recent years, remote sensing techniques such as satellite and drone-based imaging have been used to monitor Pine Wilt Disease (PWD), a widespread forest disease that causes the death of pine species. Researchers have explored the use of remote sensing imagery and deep learning algorithms to improve the accuracy of PWD detection at the single-tree level. This study introduces a novel framework for PWD detection that combines high-resolution RGB drone imagery with free-access Sentinel-2 satellite multi-spectral imagery. The proposed approach includes an PWD-infected tree detection model named YOLOv5-PWD and an effective data augmentation method. To evaluate the proposed framework, we collected data and created a dataset in Xianning City, China, consisting of object detection samples of infected trees at middle and late stages of PWD. Experimental results indicate that the YOLOv5-PWD detection model achieved 1.2% higher mAP compared to the original YOLOv5 model and a further improvement of 1.9% mAP was observed after applying our dataset augmentation method, which demonstrates the effectiveness and potential of the proposed framework for PWD detection.

Keywords: pine wilt disease (PWD); YOLOv5-PWD; deep learning; remote sensing; single-tree level detection; object detection; Sentinel-2

1. Introduction

Pine Wilt Disease (PWD) is a worldwide forest disease caused by the Pine Wood Nematode (PWN), which poses a significant threat to ecological security, biosecurity, and economic development due to its high infectiousness, pathogenicity, and rapid course of disease [1,2]. Since its discovery in Japan in 1905, PWD has caused significant economic losses and reduced ecological service values in Asia and Europe [3,4]. The monitoring of diseased trees infected with PWD is critical for the prevention and control of this disease. By detecting PWD early, forest managers can take steps to prevent the disease from spreading and protect the health of a forest.

Currently, there are three main methods for monitoring PWD-infected trees: field investigation [5–8], satellite remote sensing [3,9–11], and drone-based remote sensing (also known as "unmanned aerial vehicles" or "UAV") [12–18]. Field investigations are often costly and inefficient due to the challenging terrain of many pine forests, since most pine forests are located in areas with high mountains, steep roads and dense forests [19]. Remote sensing offers an alternative by detecting changes in the color of pine needles, which quickly

turn yellow or red after infection with PWD [20]. Satellite-based remote sensing uses image data from satellites such as Landsat and Sentinel-2 to monitor PWD over large areas, but it is limited by spatial resolution, atmospheric interference, and revisit period [9,11,21]. Drone-based remote sensing offers high flexibility, short application period, high temporal and spatial resolution, and ease of operation [22,23]. It can accurately locate individual infected trees and save resources compared to field investigations. Therefore, the use of drones for remote sensing in the detection of PWD-infected tree has thus proven to be an effective method for both detecting individual infected plants and monitoring the spread of this disease [24].

The detection of PWD using drone-based remote sensing and machine learning (ML) is garnering increasing attention. Some scholars extracted color and texture features from the ground objects in images, designed a multi-feature conditional random field (CRF)-based method for UAV image classification to classify trees and identify diseased and dead pine trees in visible light remote sensing images, effectively achieving PWD monitoring [25]. Some scholars utilized an image segmentation algorithm with ultra-green feature factor and the maximum inter-class variance method for PWD monitoring, they conducted a specific analysis of the disease severity of the diseased pine in UAV remote sensing images by effectively extracting geographic information of PWD-infected pines in the image [1]. There are also some scholars who employed both artificial neural network (ANN) and support vector machine (SVM) methods to successfully differentiate between PWD-infected trees and other cover types shown in UAV images [4]. These scholars extracted the geographical information features of ground objects in UAV remote sensing images, enabling ML-based classification methods to accurately identify PWD-infected pine trees in visible light remote sensing images. Thus, these ML-based methods established a foundation for identifying and controlling PWD within forest environments.

In recent years, numerous scholars have introduced deep-learning-based object detection models that utilize remote sensing images for Pine Wilt Disease (PWD) monitoring. The automated feature extraction and accurate detection capabilities make these models the mainstream choice in forest monitoring, disease identification, and individual tree detection [15,24,26]. These models can be categorized into two-stage (proposal-based) algorithms, such as R-CNN, Fast R-CNN, and Faster R-CNN [27–29], and one-stage (proposal-free) algorithms, such as YOLO and SSD [30,31]. The former algorithm generates target object region proposals and performs classification regression in two stages, while the latter locates and classifies target objects directly [32]. Some scholars modified the target box position regression loss function of YOLOv3, replacing the mean-square error bounding box regression loss function with the Complete-IoU loss function. They further constructed a YOLOv3-CIoU PWD diseased trees detection framework based on RGB data from UAV images, significantly enhancing the detection accuracy of YOLOv3 for PWD diseased trees [33]. Other scholars introduced the inverted residual structure and depth-wise separable convolution to enhance the YOLOv4 model. By automatically identifying discolored wood caused by PWD on UAV remote sensing images, they achieved even higher accuracy results [34].

Most of the aforementioned methods employed solely RGB data from UAV imagery, lacking other auxiliary spectral information such as near-infrared bands. Some scholars also use multispectral UAV sensors to obtain high-resolution multispectral UAV data for PWD monitoring related research. For example, some scholars acquired multispectral aerial photographs containing RGB, green, red, NIR, and red edge spectral bands, and used a multichannel convolutional neural network (CNN)-based object detection approach to detect PWD disease trees effectively [15]. Other scholars respectively obtained RGB UAV images and multispectral UAV images including blue, green, red, red edge, and near infrared (NIR) bands, and utilized Faster R-CNN and YOLOv4 deep learning models to achieve PWD-infected pine tree recognition [7]. However, most UAV sensors with dedicated multispectral cameras are expensive, leading to higher data acquisition costs. Free satellite imagery, such as Sentinel-2, can provide medium resolution (10 m) multispectral

data, enhancing the spectral information of RGB UAV imagery. Thus, we hypothesize that combining low-cost UAV RGB images with the spectral information of free satellite multispectral images can facilitate PWD-infected tree detection while further improving the model's accuracy. In addition, the detection accuracy and reliability of deep learning techniques are ultimately influenced by the number and quality of samples in the dataset, which currently remains limited in number [33,35].

Our study aims to investigate the potential of combining free-access medium-resolution multi-spectral Sentinel-2 satellite images and high-resolution drone RGB images for detecting diseased trees affected by pine wilt disease (PWD). To test our hypothesis, we collected Sentinel-2 and high-resolution drone images of a PWD-affected forested area in China, created a dataset containing diseased plants at middle-stage and late-stage of PWD. We proposed an improved PWD-infected tree detection method Yolov5-PWD based on a deep learning algorithm. In our approach, we integrated the Sentinel-2 data with the high-resolution drone images during the preprocessing stage of the UAV object detection sample data, rather than using traditional pixel-level fusion methods. This allowed us to achieve a low-cost but high-quality sample synthesis, which improved the accuracy of the object detector.

The main contributions of this study are as follows:

1. To train an object detection model capable of detecting PWD-infected trees in both middle and late stages of the disease, we constructed a dataset of such trees. This dataset involved categorizing the trees into middle-stage and late-stage categories based on their distinct characteristics at different stages. We then cut and labeled the high-resolution image data from the drone, resulting in a dataset of 1853 images and 51,124 infected trees.
2. In order to increase the accuracy of identifying trees affected by PWD and achieve efficient and accurate detection of individual PWD-infected tree, we propose an improved YOLOv5-PWD model specifically designed for PWD-infected tree identification.
3. To overcome the challenge of limited training samples and further improve the detection accuracy of the model, we propose a cost-effective and efficient sample synthesis method that leverages Sentinel-2 satellite data and UAV images. This approach enables an increase in the size of the dataset and improves detection accuracy without needing to collect more data using drones.

2. Materials and Methods

2.1. Data Collection and Dataset Construction

2.1.1. Study Area

This study was conducted in two areas (Figure 1) located in Xianning City, Hubei Province, China. Xianning City is characterized by its subtropical continental monsoon climate, with moderate temperatures, abundant precipitation, ample sunshine, and clear seasons, including a prolonged frost-free period. The annual average temperature in the city is 16.8 °C, with an average annual precipitation of 1523.3 mm, and 1754.5 h of sunshine per year. The frost-free period in Xianning City lasts for an average of 245–258 days. The forest coverage rate in the city is 53.01%, with major tree species including *Cunninghamia lanceolata*, *Phyllostachys pubescens*, and *Pinus massoniana*. The widespread occurrence of Pine Wilt Disease (PWD) in Xianning City can be attributed to its favorable climate and rich forest resources, which provide a conducive habitat for the vector insect *Monochamus alternatus*.
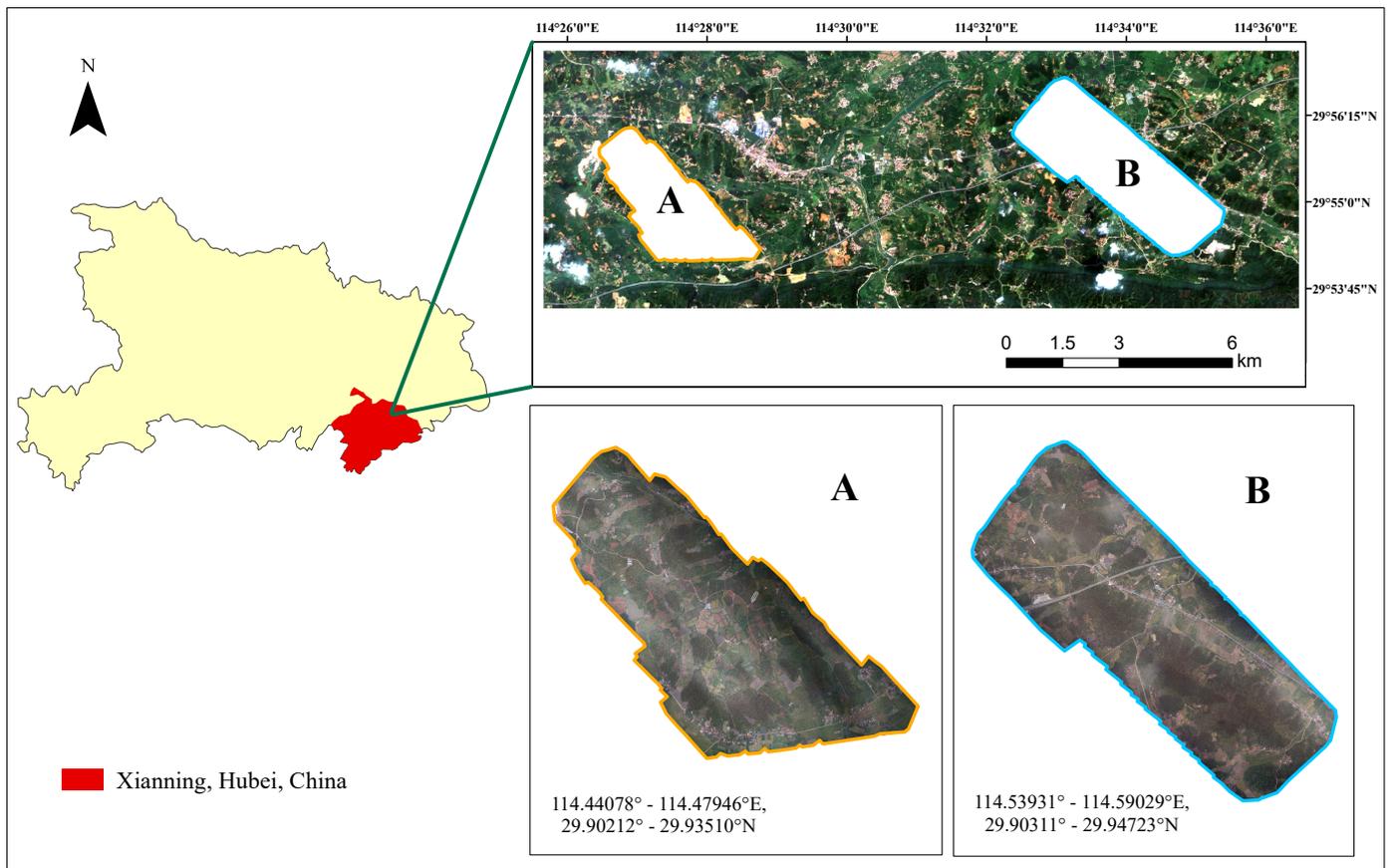
Figure 1. Overview of the study areas in this study. Left and top right are the locations of the study area in Xianning City, Hubei Province, China. (**A**,**B**) are UAV Orthophoto maps of two study areas.

2.1.2. Imagery Source

In this study, high-resolution UAV imagery was collected within the study areas, to construct a training dataset for deep learning models. The Sentinel-2 imagery from the study area was also acquired at the same month as the UAV imagery and prepared for sample synthesis. The UAV images were processed using ContextCapture software to generate high-resolution UAV orthophoto maps for use in the analysis. These high-resolution orthophoto maps acquired by the UAV were partitioned into image patches of size 1024 × 1024 pixels to create a sample dataset for the deep learning model. The UAV remote sensing images consist of three bands of red, green, and blue and were collected in September 2020. The resolution of the UAV images in Area A was 3.975 cm, while the resolution in Area B was 7.754 cm. The Sentinel-2 imagery, consisting of four spectral bands of red, green, blue and near-infrared, was acquired and processed using Google Earth Engine in September 2020. The resolution of the Sentinel-2 imagery was 10 m.

2.1.3. Disease Tree Category System and Labeling Solution

In this study, the classification of Pine Wilt Disease (PWD) infected trees into two categories, the middle-stage and the late-stage, was carried out based on the characteristic color changes of infected trees in the respective stages, as revealed by previous studies [36–38]. A manual annotation strategy (Figure 2) was implemented to produce the dataset, which consisted of three steps: Single Manual Labeling, Object Detection Model Training and Classification and Labeling Update.
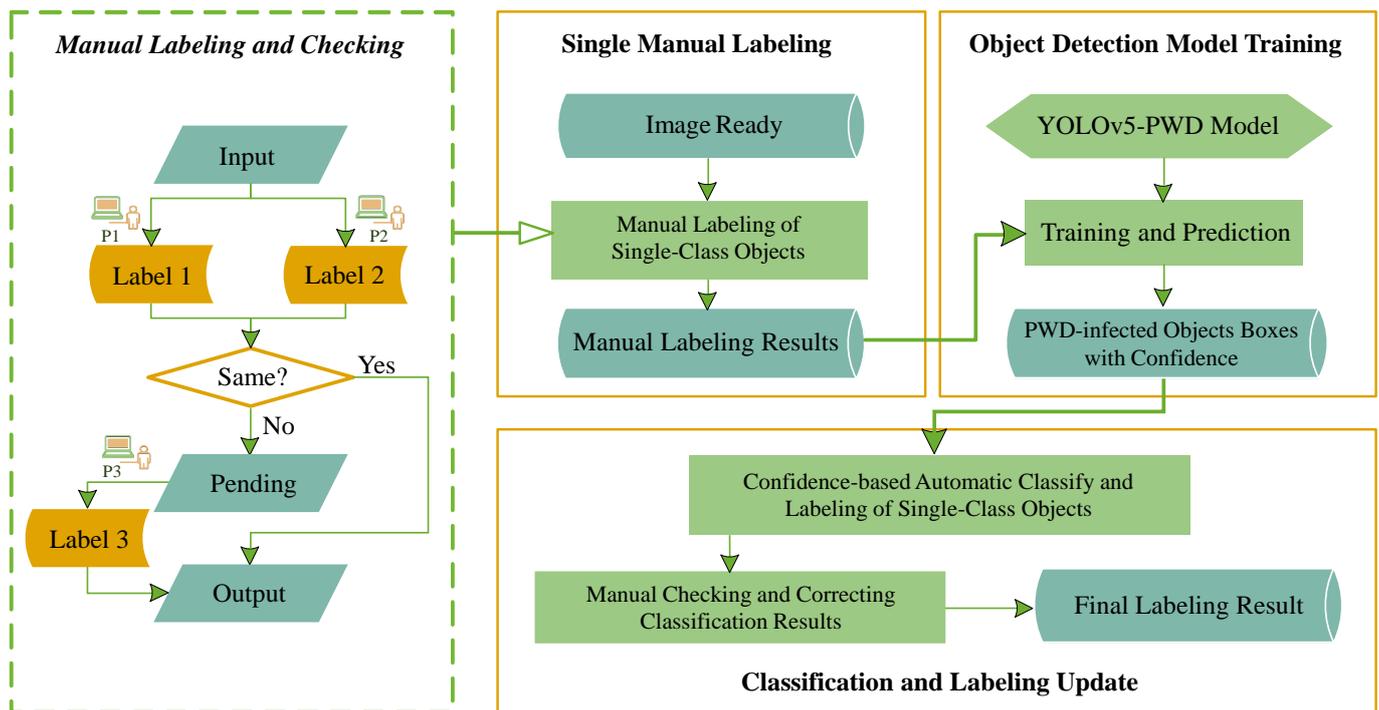
Figure 2. The process of dataset labeling and checking.

In the first step, a single category of PWD-infected tree was manually annotated based on the criterion of "abnormal color". The results of manual labeling were then fed into the YOLOv5-PWD Model (see Section 2.2.1) for training and prediction, producing a result indicating the confidence level. In the final step, annotators based on the confidence level reclassified the PWD-infected object into two categories: middle-stage and late-stage.

The manual annotation and checking process was performed by three personnel to ensure accuracy and consistency. In case of disagreement between Personnel 1 (P1) and Personnel 2 (P2), Personnel 3 (P3) was responsible for making the final decision. This 3-person method was critical to the success of both the Single Manual Labeling and Classification and Labeling Update steps.

2.1.4. Object-Level Labeling

High-resolution UAV aerial photography RGB images were used in this study to create a sample dataset for deep learning models. LabelImg (https://github.com/heartexlabs/labelImg, accessed on 1 April 2023) software was utilized to label the cropped images of the study area. The labeled rectangles contain the PWD-infected tree and their respective categories, including the middle-stage and the late-stage. Figure 3 presents a portion of the calibration results. A total of 1853 annotated images containing 51,124 infected trees were obtained, and the labeled images were randomly divided into training, validation, and test sets, with 889 images in the training set, 370 images in the validation set, and 594 images in the test set.

(**a**) Middle-stage infected trees

(**b**) Late-stage infected trees

Figure 3. Examples of UAV images of middle-stage infected trees and late-stage infected trees. The boxes in the images indicate the PWD-affected trees. (**a**) Middle-stage infected trees are generally yellow-green, yellow and orange. (**b**) Late-stage infected trees are generally orange-red and red.

2.2. Methodology

In this paper, we used a deep learning-based object detection framework to detect the PWD-affected trees. The proposed deep learning-based object detection framework employed a modified YOLOv5-PWD model, which was based on the original YOLOv5 (You Only Look Once v5) object detection algorithm and incorporated the Efficient Intersection over Union (EIoU) loss function [39] instead of the original target box position regression loss function. Furthermore, a high-quality sample synthesis method was proposed to address the issue of an uneven number of categories and a small number of samples by integrating Sentinel-2 satellite multi-spectral data with UAV imagery data.

2.2.1. Yolov5-Pwd Detection Model

In this research, we developed the YOLOv5-PWD model by building on the YOLOv5m algorithm [40], which is the fifth iteration of the YOLO series [30,41]. Compared to proposal-based object detection models [27–29], the YOLO series algorithms are ideal for quickly identifying trees infected by PWD, as they offer fast processing times and cost-effectiveness.

The network structure of the YOLOv5-PWD detection model is shown in Figure 4. YOLOv5-PWD detection model is composed of four parts: Input, Backbone, Neck and Head. The Backbone is a network used to extract the features of the image from the Input, and its function is to extract the information in the image for use by the later network. It introduces the CSP modules (Cross Stage Partial Network) [42] with residual structures into Darknet53, which can enhance the gradient value in the backpropagation between layers and effectively prevents the gradient disappearing while the network deepens. The Neck is

to further extract and fuse the image feature information output by the Backbone. It adopts the structure of FPN [43] combined with PAN [44]. The FPN structure uses up-sampling to improve the network's ability to detect small targets, and the PAN structure makes the underlying positioning information better transmitted to the top layer. The combination of these two structures strengthens the feature fusion ability of the network. In addition, the SPPF (Spatial Pyramid Pooling—Fast) module [45] in the Neck is to serial inputs through multiple MaxPool layers of 5×5 size, and then performs Concat merging, so it can increase the receptive field of the feature network, effectively separate the feature information, and restore the output to be consistent with the input. Finally, the Head component of the network is responsible for predicting the target and applying the anchor box on the target feature map to generate the final output vector containing the classification probability and target box. The Head outputs three scales of target predictions, corresponding to three different target sizes: large, medium, and small.
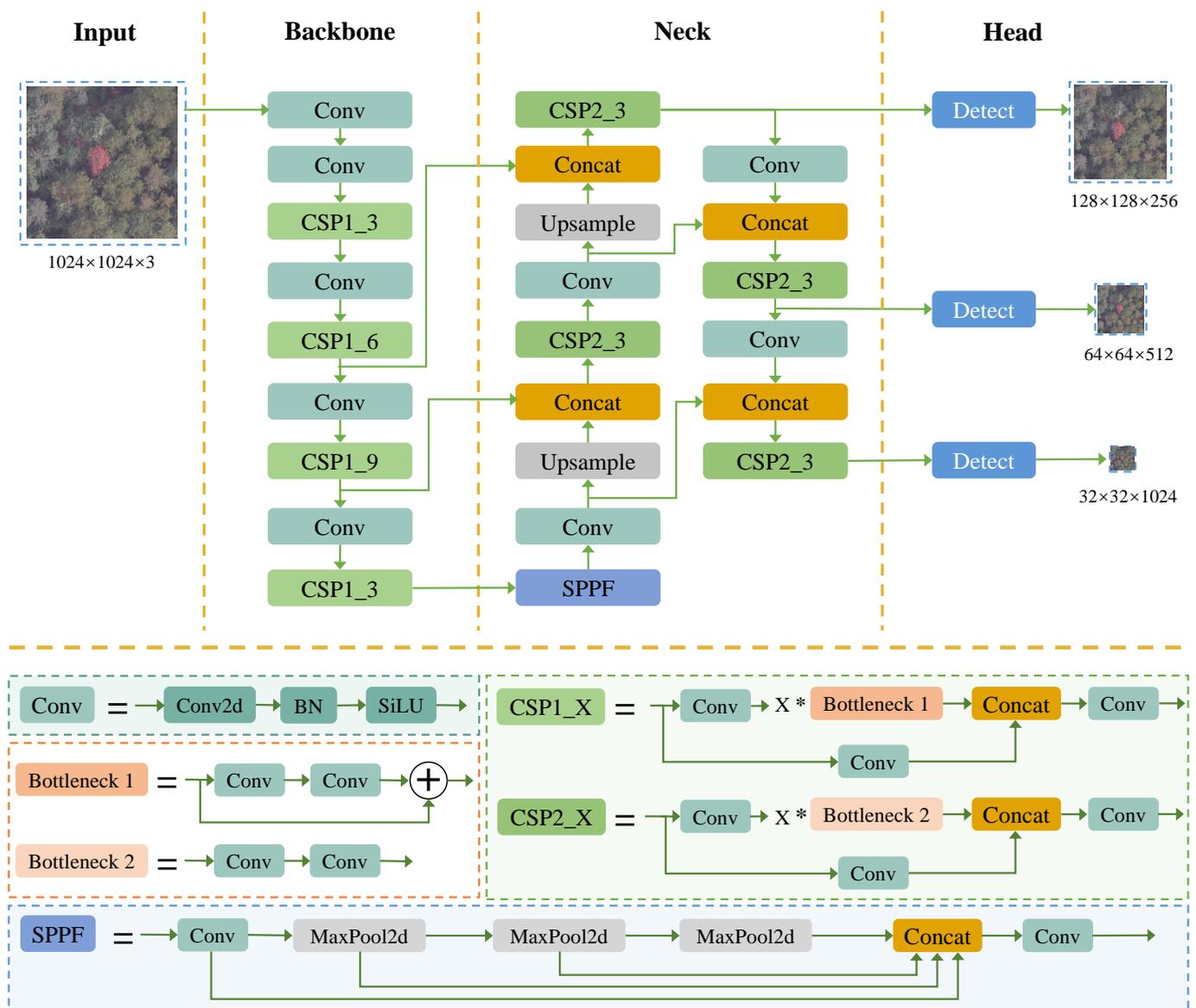

Figure 4. The network architecture of the YOLOv5-PWD model (X * Y represents repeating module Y for X times).

Different models use loss functions to guide their training process by measuring the difference between predicted and actual values, with smaller loss values indicating better

performance. In the YOLOv5-PWD model, we introduce EIoU Loss [39] instead of the Complete-IoU (CIoU) Loss [46] used by the original Yolov5 to calculate the error between the predicted box and the real box. The formula for calculating EIoU loss is as follows:

$$
\begin{aligned}
L_{EIoU} &= L_{IoU} + L_{dis} + L_{wh} \\
&= 1 - IoU + \frac{\rho^2(b, b^{gt})}{c^2} + \frac{\rho^2(w, w^{gt})}{C_w^2} + \frac{\rho^2(h, h^{gt})}{C_h^2}
\end{aligned}
\tag{1}
$$

where, $\rho(b, b^{gt})$ represents the Euclidean distance between two center points. c represents the diagonal distance of the smallest closed convex surface that can contain both the prediction box and the real box. $\rho(w, w^{gt})$ represents the difference between the widths of the real box and the predicted box. C_w represents the width of the smallest closed convex surface that can contain both the prediction box and the real box. $\rho(h, h^{gt})$ represents the difference between the heights of the real box and the predicted box. C_h represents the height of the smallest closed convex surface that can contain both the prediction box and the real box.

The EIoU Loss contains three parts: the overlap loss (L_{IoU}), the center distance loss (L_{dis}), and the width and height loss (L_{wh}). The first two parts of the EIoU Loss are derived from the CIoU Loss. However, the third part in CIoU Loss considers the difference in aspect ratio between the predicted box and the ground truth box, instead of the difference in width and height. This can sometimes hinder the model from effectively optimizing the similarity when the aspect ratio of the predicted box and the ground truth box are linearly proportional in the regression process. The EIoU Loss addresses the limitation of the CIoU Loss by replacing the aspect ratio influence factor in the third part with the width and height loss, which calculates the difference in length and width between the ground truth box and the predicted box, resulting in a direct minimization of the difference in width and height and leading to improved accuracy and faster convergence.

2.2.2. A Method for Synthesizing Samples Using Sentinel-2 Imagery to Augment UAV Image Data

Deep network models are often prone to overfitting due to the limited size of the dataset. While increasing the dataset size is a possible solution, it is not always feasible and can be expensive. To address this issue, we propose a sample synthesis algorithm to augment the UAV image dataset by utilizing the multispectral information from satellite images. This method aims to improve detection accuracy, reduce costs, and prevent model overfitting. The training sets were augmented using this method, while the validation and test sets remained unchanged.

Figure 5 illustrates the procedure of our method. We first extracted single PWD-infected tree targets from UAV images that captured PWD-infected trees. Next, we collected UAV images without PWD-infected trees and synthesized the single PWD-infected tree targets onto these images to create new images containing PWD-infected trees. However, the general synthesis method we used is a simple and direct method that did not consider the degree of fusion between the synthetic target and the background. This led to some PWD-infected trees appearing in non-vegetated areas such as buildings and water bodies, which could negatively impact the accuracy of model detection. Nevertheless, when the background is a forested area or an area with high vegetation coverage, the synthesized image tends to be more realistic, leading to better model training outcomes.
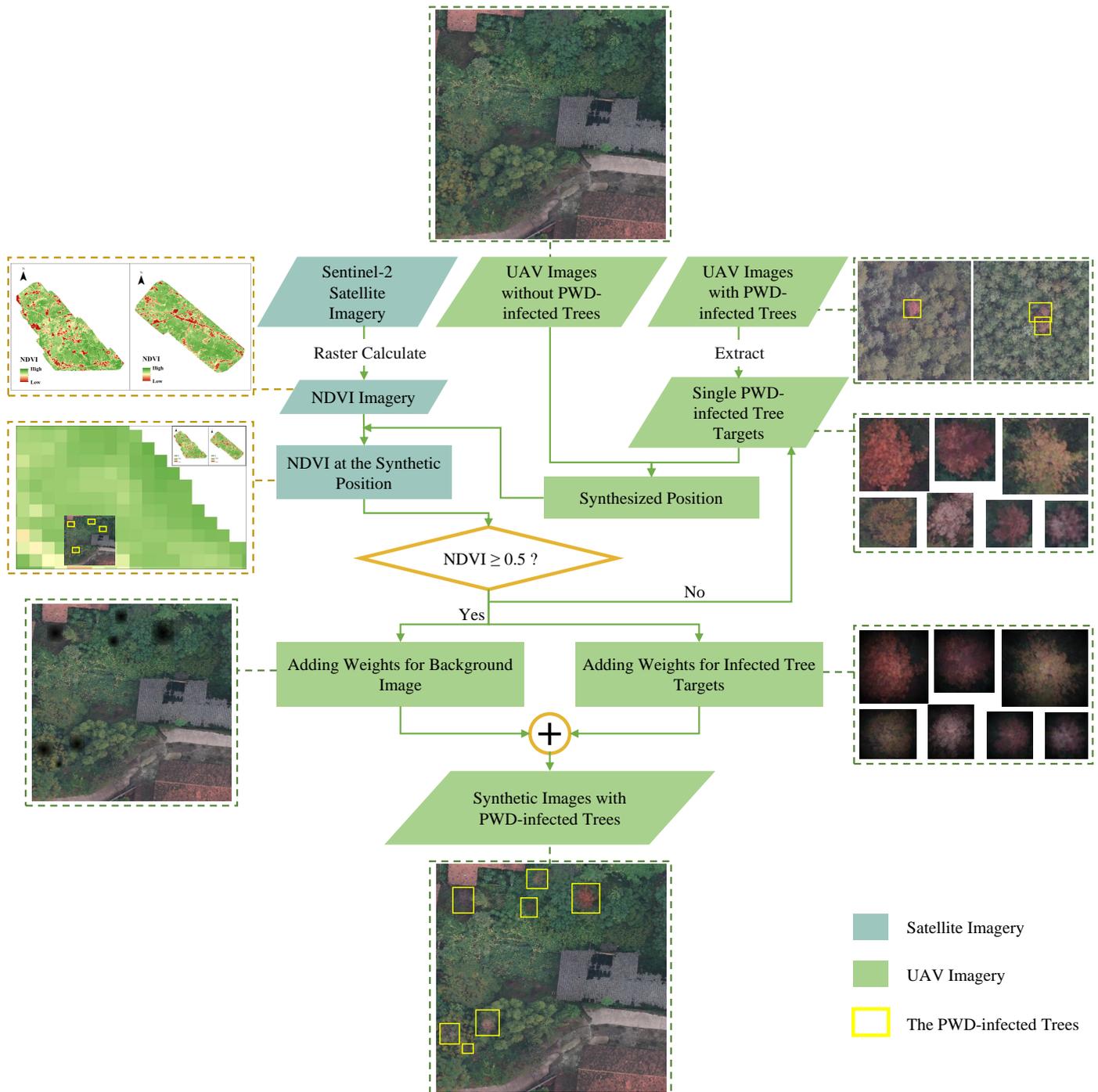
Figure 5. The technical process of synthesizing samples by combining satellite and UAV images.

To address this issue, we developed a more effective synthesis approach by leveraging the rich spectral information from multi-band satellite images and calculating the vegetation index to determine the vegetation coverage of potential synthesis locations. Specifically, we utilized Sentinel-2 satellite imagery of the study area and calculated the NDVI (Normalized Difference Vegetation Index) value. Then, we assessed the NDVI value at the proposed synthesis location. If the NDVI value was greater than or equal to 0.5, we synthesized the single PWD-infected tree target onto the UAV imagery at that position. However, if the NDVI value was less than 0.5, we did not perform the synthesis and instead reselected the single PWD-infected tree targets and the synthesis location. By using this approach, we were able to significantly improve the accuracy of the model detection.

In our sample synthesis approach, we adopted two synthesis strategies for synthesizing single PWD-infected tree targets in UAV images. The first strategy involves directly synthesizing the single PWD-infected tree target in the image, resulting in a slight inconsistency between the synthesized image and the real one (as shown in Figure 6b). The second strategy involves adding weights to the background image and PWD-infected tree targets separately (as illustrated in the bottom part of Figure 5). This approach ensures that the prominent parts of the target frame are preserved, while the edges of the target frame are weakened, and this makes the PWD-infected tree targets blend more seamlessly with the background. As a result, the synthesized image is more realistic (as depicted in Figure 6c).
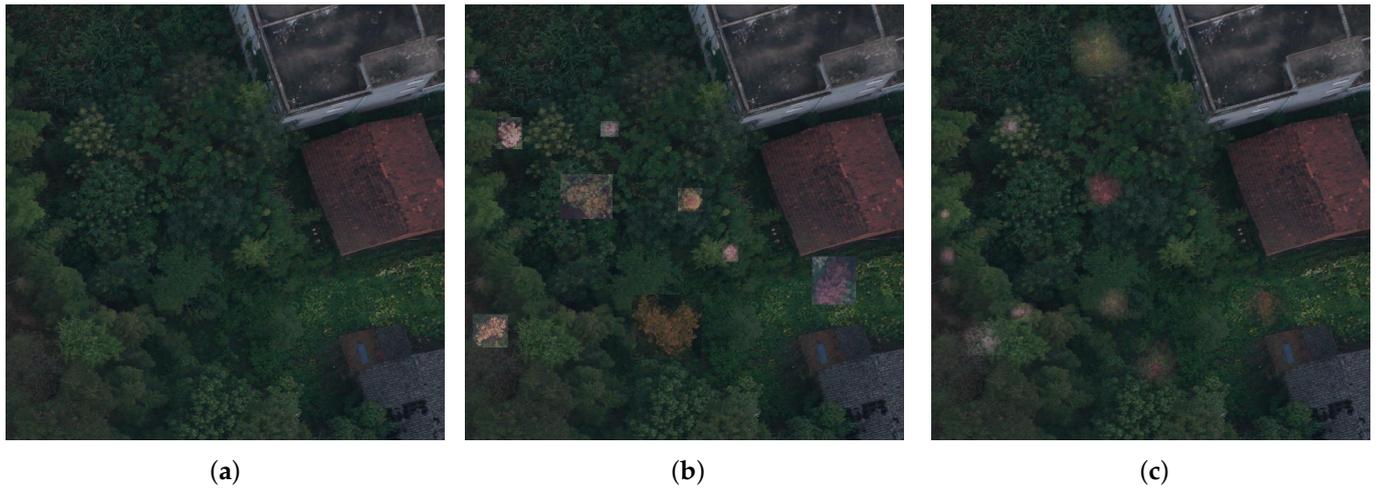

(a)   (b)   (c)

Figure 6. An example of synthetic images with different synthesis strategies. (**a**) Without infected trees; (**b**) Synthesis by Direct Synthesis; (**c**) Synthesis by Weighted synthesis.

The following are the pixel-wise calculation formulas for the second synthesis strategy that involves adding weights:

$$T_{new} = \alpha T_{old} + (1 - \alpha) T_{background}$$

$$\alpha = 1 - \frac{\sqrt{\left(w_i - \frac{W}{2}\right)^2 + \left(h_i - \frac{H}{2}\right)^2}}{\sqrt{\left(\frac{W}{2}\right)^2 + \left(\frac{H}{2}\right)^2}} \tag{2}$$

where T_{new} represents the pixel value of point i in the weighted infection tree target frame image; T_{old} represents the pixel value of point i in the unweighted original infection tree target frame image; $T_{background}$ represents the pixel value of point i in the background image; α is the weight value to be increased; W represents the width of the target frame image of the infected tree; H represents the height of the target frame image of the infected tree; (w_i, h_i) represents the coordinate at point i when the top left vertex of the target frame image of the infected tree is $(0, 0)$. These are shown in Figure 7a, where r represents half of the diagonal length of the target frame image of the infected tree, which is the distance from the center point of the target frame image to the four vertices; d_i represents the distance from point i to the center point in the target frame image of the infected tree. When the point i is closer to the center point, the value of d_i is smaller and the value of α is larger, the feature information of the target frame image of the infected tree retained at the point i is more, and the feature information of the background image retained at the point i is less. When point i is at the edge of the target frame image of the infected tree, the value of d_i is larger, the value of α is smaller and the value of $(1 - \alpha)$ is larger, the feature information of the background image retained at the point i is more, and the feature information of the target frame image of the infected tree retained at the point i is less.
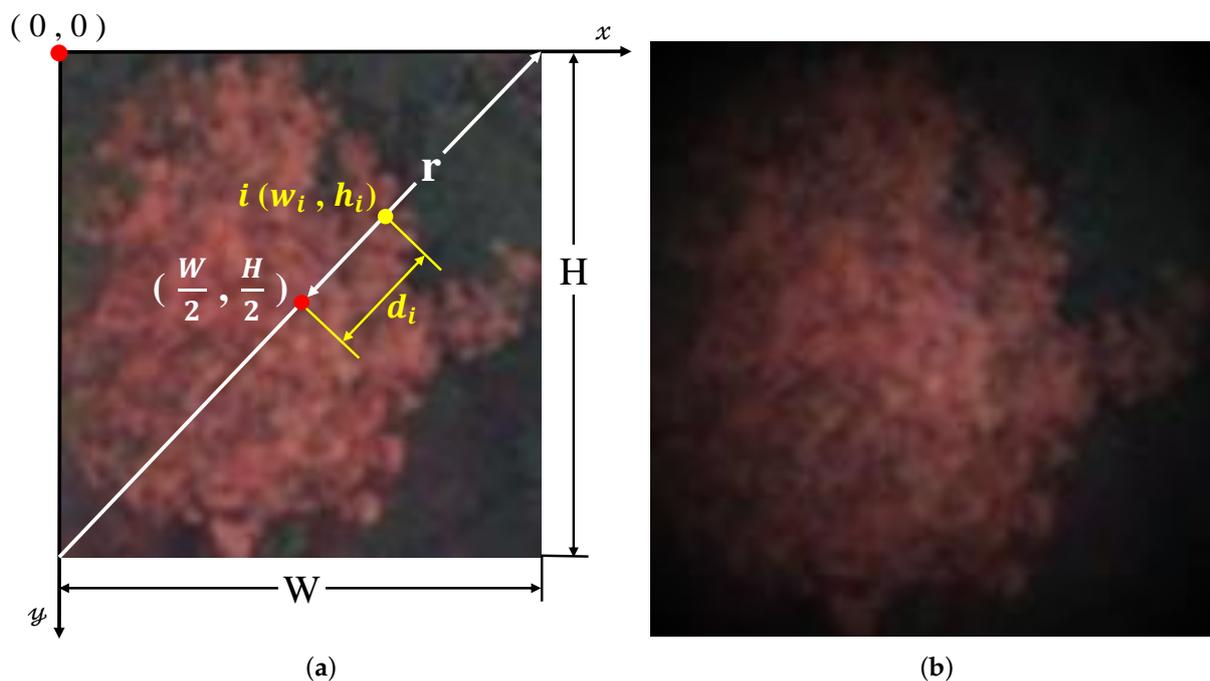
Figure 7. Diagram of adding weights. (**a**) Schematic diagram of the relevant points in Formule (2). (**b**) The target frame image after adding weights.

2.3. Experiment Settings

2.3.1. Design of Experiments

In this paper, we conduct external and internal experiments to validate our model and method. For external experiments, we use the PWD-infected tree dataset we created (mentioned in Section 2.1) to train, verify, and test several commonly used object detection models and our YOLOv5-PWD model. We then compare the detection accuracy and speed of these models to demonstrate the superior detection performance of our YOLOv5-PWD model in recognizing PWD-infected trees.

For internal experiments, we train both the YOLOv5 model and the YOLOv5-PWD model using the original training set (889 images), the training set without NDVI synthesis (General synthesis), and the NDVI training sets that use two different synthesis strategies (Direct synthesis and Weighted synthesis). We then compare the detection accuracy of the models trained on different training sets using the same test set to demonstrate the effectiveness of the proposed sample synthesis method.

2.3.2. Evaluation and Metrics

To evaluate the performance of our PWD-infected tree detection model, we use four indicators: the P-R curve, average precision, mean average precision (mAP), and Frames Per Second (FPS). The first three indicators are used to evaluate the accuracy of model detection, while FPS is used to assess the speed of the object detection model, specifically, the number of pictures that the model can process per second. A higher FPS indicates faster model detection speed.

In object detection, Precision measures the proportion of the detected targets that are real, while Recall measures the proportion of all real targets that are detected. The P-R curve (Precision-Recall curve) plots Precision on the *y*-axis and Recall on the *x*-axis, illustrating the relationship between the two metrics [32]. We aim to maximize both Precision and Recall, resulting in a rightward convex PR curve.

Average Precision (AP) refers to the average precision value for all Recall values between 0 and 1, which represents the area under the P-R curve. A higher AP score

indicates better detection performance. The mean average precision (mAP) is the average AP score across all categories [47].

2.3.3. Experimental Settings

We performed the experiments on Wuhan University's Supercomputing Center using PyTorch deep learning framework and CUDAToolkit 10.2. We adapted YOLOv5 and YOLOv5-PWD from their official code repository (https://github.com/ultralytics/yolov5, accessed on 1 April 2023). We implemented Faster R-CNN, DetectoRS Cascade RCNN, RetinaNet, and Cascade RCNN using MMDetection framework [48]. We trained all models for 72 epochs using Adam optimizer with a learning rate of 0.01 for YOLOv5 and YOLOv5-PWD, and 0.0025 for others. We used an NMS (non-maximum suppression) threshold of 0.6.

3. Results

3.1. Results of External Experiments

Table 1 displays the results of external experiments. In each column, the bold number indicates the best detection result, and the other tables are the same. The YOLOv5-PWD detection model achieved the highest detection accuracy, followed by YOLOv5. Meanwhile, YOLOv5 and RetinaNet outperformed the others in terms of FPS, with YOLOv5-PWD following closely. Hence, the YOLOv5-PWD detection model is best suited for epidemic prevention, as it enables quick and precise detection.

Table 1. Comparison of detection performance of different models.

Model	Middle-Stage AP	Late-Stage AP	mAP@0.5	FPS
Faster R-CNN	43.8%	74.0%	58.9%	7.35
DetectoRS Cascade RCNN	44.5%	74.1%	59.3%	5.41
RetinaNet	37.8%	72.4%	55.1%	12.20
Cascade RCNN	43.5%	74.2%	58.8%	10.99
YOLOv5	57.4%	84.4%	70.9%	12.20
YOLOv5-PWD	**59.4%**	**84.8%**	**72.1%**	11.49

Table 1 shows that the detection accuracy of several models for late-stage infected trees is significantly better than that of middle-stage infected trees. Additionally, there is a substantial difference in AP values between the two types of infected trees in the initial four models, with a difference of more than 29%. However, the YOLOv5 and YOLOv5-PWD models demonstrated better performance in detecting middle-stage infected trees, with differences of 27% and 25.4% in AP values, respectively, for the two types of infected trees. This indicates that the YOLOv5-PWD model can achieve better detection of middle-stage infected trees while maintaining high-precision detection of late-stage infected trees. Moreover, the YOLOv5-PWD model outperformed other models in detecting PWD infection in trees.

3.2. Results of Internal Experiments

The results of the internal experiments are shown in Table 2 and the P-R curves of the YOLOv5 and YOLOv5-PWD models are shown in the Figure 8. The table demonstrates that the model's performance has improved with the three augmentation strategies applied to the training sets. We trained the YOLOv5-PWD model with different training sets and found that the weighted synthesis with NDVI augmentation achieved the best results. The model's mAP reached 74%, which was 1.9% higher than that of the original training set. The mAP of the general synthesis without NDVI augmentation was 0.6% higher than that of the original training set. The mAPs of the direct synthesis and weighted synthesis with NDVI augmentation were 0.8% and 1.3% higher than that of the general synthesis without NDVI augmentation, respectively. When using satellite imagery NDVI for sample synthesis, the weighted synthesis outperformed the direct synthesis by 0.5%. In summary,

the weighted synthesis with NDVI augmentation had the best effect on the model, followed by the direct synthesis with NDVI augmentation, then by the general synthesis without NDVI augmentation, and finally by the original dataset. This result was also consistent for the original YOLOv5 model. Therefore, we conclude that our proposed UAV image data augmentation sample synthesis method based on satellite imagery can improve the detection accuracy of the model.

Table 2. Detection results of YOLOv5 and YOLOv5-PWD models on dataset with different data augmentation policies.

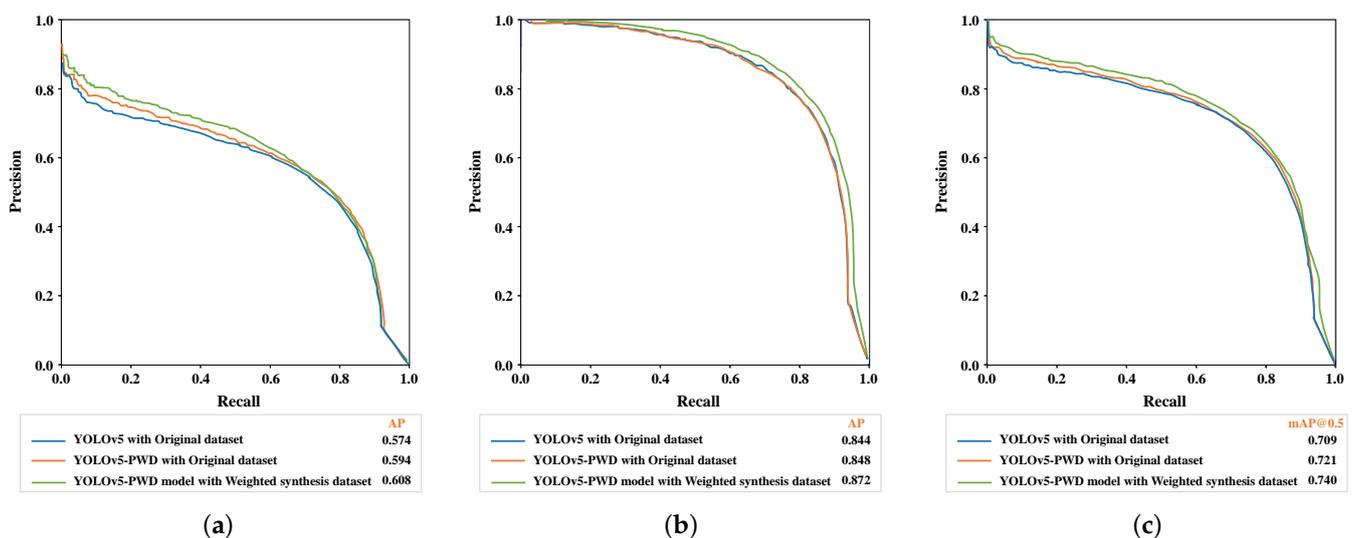| Training Set	Number of Images	Middle-Stage AP (%)		Late-Stage AP (%)		mAP@0.5 (%)	
		YOLOv5	YOLOv5-PWD	YOLOv5	YOLOv5-PWD	YOLOv5	YOLOv5-PWD
Original	889	57.4	59.4	84.4	84.8	70.9	72.1
General synthesis	5480	57.9	59.6	84.6	85.8	71.2	72.7
Direct synthesis	5471	59.1	60.4	85.6	86.5	72.3	73.5
Weighted synthesis	5470	59.7	**60.8**	86.6	**87.2**	73.1	**74.0**

Figure 8. The P-R curves of the YOLOv5 and YOLOv5-PWD models on the test set. (**a**) The detection results of middle-stage infection trees; (**b**) The detection results of late-stage infection trees; (**c**) The detection results of two classes infection trees.

In addition, the YOLOv5-PWD model always achieved higher detection accuracy than the YOLOv5 model when trained with the same dataset. Compared with the YOLOv5 model trained with the original dataset, the YOLOv5-PWD model trained with the same dataset showed a 1.2% improvement in mAP for PWD-infected tree detection. The YOLOv5-PWD detection model trained with the weighted synthesis with NDVI augmentation further increased the mAP of PWD-infected tree detection by 1.9%. Our YOLOv5-PWD model can better detect PWD-infected tree, and our UAV image data augmentation sample synthesis method based on satellite imagery can further improve its performance.

4. Discussion

In this paper, we present a PWD-infected tree sample dataset. This dataset consists of 1853 images and 51,124 infected trees. We divided the infected trees into middle-stage and late-stage categories, with 28,400 and 22,724 targets, respectively. In order to observe the distribution of samples of different categories in the dataset, we additionally randomly

cropped 15,000 targets as background, which together with the middle-stage infection tree targets and late-stage infection tree targets constitute a sample feature dataset. We use the t-Distributed Stochastic Neighbor Embedding algorithm (t-SNE) [49] to reduce the dimensionality of the sample features of this dataset, and finally obtain a visualized data distribution result (as shown in Figure 9). As Figure 9 shows, there is a clear distinction between the background targets and the infected tree targets in our sample data set, which can make the model better classify the infected trees and the non-infected trees. However, it also reveals that the distinction between categories of middle-stage infected trees and late-stage infected trees is not obvious. Therefore, effectively distinguishing between middle-stage infected trees and late-stage infected trees has become one of the challenges of the model for detecting PWD-infected trees.
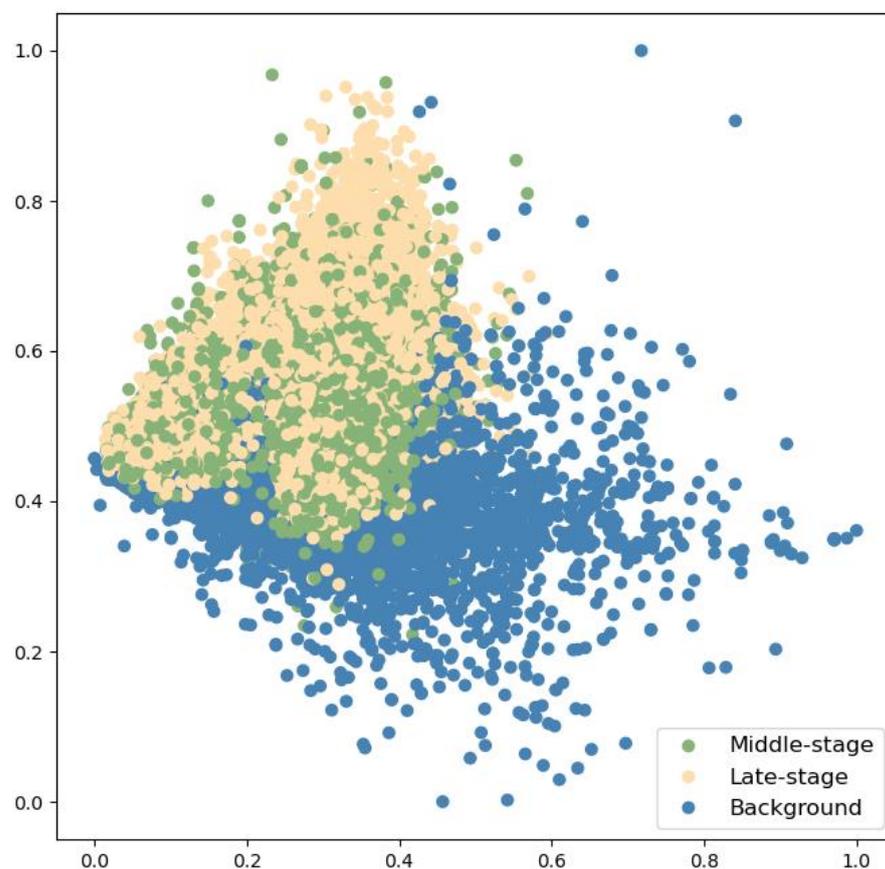

Figure 9. Visualization of sample data distribution of our dataset in two-dimensional space by t-SNE.

Experiments on our constructed dataset show that the YOLOv5-PWD disease tree detection model we built can efficiently and accurately identify PWD-infected trees. It has achieved certain results in the classification and identification of middle-stage and late-stage infected trees, but it needs to be improved in more refined identification of early- and middle-stage infected trees. Compared with several other commonly used object detection models, our proposed YOLOv5-PWD model shows better detection accuracy and acceptable speed in the detection of single PWD-infected trees, achieving the purpose of rapid and accurate detection of PWD. In addition, as shown in Table 1, our model shows better detection accuracy in identifying middle-stage and late-stage infected trees, respectively, and it has a more obvious advantage in distinguishing between them. However, it can also be seen from Table 1 that although our model outperforms other models in detecting PWD middle-stage infected trees, all of these models have lower accuracy in detecting middle-stage PWD-infected trees than late-stage PWD-infected trees. The reason for this may be that the features of the PWD middle-stage infected trees are more

complex, compared with the PWD late-stage infected trees. Therefore, further improving the recognition accuracy of the model for early and middle stage PWD-infected trees is one of our future challenges.

The sample synthesis method we propose combines satellite images and high-resolution UAV images. It can increase the number of samples in the dataset while ensuring the quality of the synthesized samples, saving costs and improving model detection accuracy. We used this sample synthesis method to increase the number of samples in the training dataset by over six times, so that the deep learning model can be more fully trained, and the detection accuracy of the model has also been improved. During the synthesis process, if we directly pasted the infected tree targets onto the image, the synthesized image would be too inconsistent and distorted (as shown in Figure 6b), which was not conducive to training the model. Therefore, we increased the weight of both the infected tree targets and the background image, which can retain the obvious features of the center and weaken the indistinct features of the edge, making the final composite image more realistic (as shown in Figure 6c). It can also be seen from Table 2 that more realistic images really help to improve the effect of model training and detection.

5. Conclusions

In this paper, we took the forested area known to be affected by PWD in Xianning City, China as the study area, and we collected the UAV images and Sentinel-2 satellite imagery of the study area as the experimental data for our research. Firstly, we classify the PWD-infected tree of high-resolution UAV RGB images into middle-stage infected trees and late-stage infected trees, and construct a PWD-infected tree sample dataset, which consists of 889, 370, and 594 images in training sets, validation sets and test sets, respectively. Secondly, we improve the YOLOv5 deep learning object detection model and build the YOLOv5-PWD detection model. The YOLOv5-PWD model has obvious advantages over other object detection models in terms of detection effect, achieving an AP value of 59.4% for detecting the middle-stage class, an AP value of 84.8% for detecting the late-stage class, an mAP of 72.1%, and an FPS of 11.49, thus enabling our model to quickly and accurately identify single PWD-infected trees. Finally, to increase the amount of training data, improve the recognition efficiency of the infected tree object detection model, and reduce the cost of data acquisition, we propose a sample synthesis method that combines large-scale free-access Sentinel-2 satellite imagery and high-resolution UAV imagery. After sample synthesis, the number of images in the training set increased by more than six times, and our model's detection effect was improved. The model achieved an AP value of 60.8% for detecting the middle-stage infected trees and an AP value of 87.2% for detecting the late-stage infected trees, and an mAP of 74% for the overall model detection, achieving a relatively high detection accuracy.

For future work, we plan to extend our research in the following directions: (1) We will clarify the classification system of PWD-infected trees, construct a larger dataset containing early-, mid-, and late-stage infected trees, evaluate its accuracy and quality using authoritative methods, and further explore models and methods to improve the identification of early and mid-stage infected trees. (2) We will explore methods to fuse multi-source data such as satellite and drone data to provide the model with more comprehensive training and improve its ability in detecting PWD-infected trees. (3) During the model training process, we will consider more auxiliary data such as terrain, meteorological data, and humidity, and combine these factors with remote sensing image data to explore methods for improving detection accuracy.

Author Contributions: Conceptualization, P.C. and G.C.; methodology, P.C. and G.C.; investigation, P.C. and H.Y.; data curation, P.C., H.Y., X.L., T.W. and M.H.; writing—original draft preparation, P.C., G.C. and H.Y.; writing—review and editing, K.Z., P.L. and X.Z.; supervision, Q.W. and Y.G.; project administration, G.C.; funding acquisition, G.C. and X.Z. All authors have read and agreed to the published version of the manuscript.

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

References

1. Li, H.; Xu, H.; Zheng, H.; Chen, X. Research on pine wood nematode surveillance technology based on unmanned aerial vehicle remote sensing image. *J. Chin. Agric. Mech.* **2020**, *41*, 170–175.
2. Zhang, X.; Yang, H.; Cai, P.; Chen, G.; Li, X.; Zhu, K. Research progress on remote sensing monitoring of pine wilt disease. *Trans. Chin. Soc. Agric. Eng.* **2022**, *38*, 184–194.
3. Huan, T.; Cunjun, L.; Cheng, C.; Liya, J.; Haitang, H. Progress in remote sensing monitoring for pine wilt disease induced tree mortality: A review. *For. Res.* **2020**, *33*, 172–183.
4. Syifa, M.; Park, S.J.; Lee, C.W. Detection of the pine wilt disease tree candidates for drone remote sensing using artificial intelligence techniques. *Engineering* **2020**, *6*, 919–926. [CrossRef]
5. Kim, S.R.; Lee, W.K.; Lim, C.H.; Kim, M.; Kafatos, M.C.; Lee, S.H.; Lee, S.S. Hyperspectral analysis of pine wilt disease to determine an optimal detection index. *Forests* **2018**, *9*, 115. [CrossRef]
6. Yu, R.; Ren, L.; Luo, Y. Early detection of pine wilt disease in Pinus tabuliformis in North China using a field portable spectrometer and UAV-based hyperspectral imagery. *For. Ecosyst.* **2021**, *8*, 44. [CrossRef]
7. Yu, R.; Luo, Y.; Zhou, Q.; Zhang, X.; Wu, D.; Ren, L. Early detection of pine wilt disease using deep learning algorithms and UAV-based multispectral imagery. *For. Ecol. Manag.* **2021**, *497*, 119493. [CrossRef]
8. Yu, R.; Luo, Y.; Zhou, Q.; Zhang, X.; Wu, D.; Ren, L. A machine learning algorithm to detect pine wilt disease using UAV-based hyperspectral imagery and LiDAR data at the tree level. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *101*, 102363. [CrossRef]
9. Zhang, Y.; Dian, Y.; Zhou, J.; Peng, S.; Hu, Y.; Hu, L.; Han, Z.; Fang, X.; Cui, H. Characterizing Spatial Patterns of Pine Wood Nematode Outbreaks in Subtropical Zone in China. *Remote Sens.* **2021**, *13*, 4682. [CrossRef]
10. Zhou, H.; Yuan, X.; Zhou, H.; Shen, H.; Ma, L.; Sun, L.; Fang, G.; Sun, H. Surveillance of pine wilt disease by high resolution satellite. *J. For. Res.* **2022**, *33*, 1401–1408. [CrossRef]
11. Zhou, Z.; Li, R.; Ben, Z.; He, B.; Shi, Y.; Dai, W. Automatic identification of Bursaphelenchus xylophilus from remote sensing images using residual network. *J. For. Eng.* **2022**, *7*, 185–191.
12. Liu, S.; Wang, Q.; Tang, Q.; Liu, L.; He, H.; Lu, J.; Dai, X. High-resolution image identification of trees with pinewood nematode disease based on multi-feature extraction and deep learning of attention mechanism. *J. For. Eng.* **2022**, *7*, 177–184.
13. Wu, B.; Liang, A.; Zhang, H.; Zhu, T.; Zou, Z.; Yang, D.; Tang, W.; Li, J.; Su, J. Application of conventional UAV-based high-throughput object detection to the early diagnosis of pine wilt disease by deep learning. *For. Ecol. Manag.* **2021**, *486*, 118986. [CrossRef]
14. Li, F.; Liu, Z.; Shen, W.; Wang, Y.; Wang, Y.; Ge, C.; Sun, F.; Lan, P. A remote sensing and airborne edge-computing based detection system for pine wilt disease. *IEEE Access* **2021**, *9*, 66346–66360. [CrossRef]
15. Park, H.G.; Yun, J.P.; Kim, M.Y.; Jeong, S.H. Multichannel Object Detection for Detecting Suspected Trees with Pine Wilt Disease Using Multispectral Drone Imagery. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 8350–8358. [CrossRef]
16. Xia, L.; Zhang, R.; Chen, L.; Li, L.; Yi, T.; Wen, Y.; Ding, C.; Xie, C. Evaluation of Deep Learning Segmentation Models for Detection of Pine Wilt Disease in Unmanned Aerial Vehicle Images. *Remote Sens.* **2021**, *13*, 3594. [CrossRef]
17. Qin, J.; Wang, B.; Wu, Y.; Lu, Q.; Zhu, H. Identifying pine wood nematode disease using UAV images and deep learning algorithms. *Remote Sens.* **2021**, *13*, 162. [CrossRef]
18. Yu, R.; Luo, Y.; Li, H.; Yang, L.; Huang, H.; Yu, L.; Ren, L. Three-Dimensional Convolutional Neural Network Model for Early Detection of Pine Wilt Disease Using UAV-Based Hyperspectral Images. *Remote Sens.* **2021**, *13*, 4065. [CrossRef]
19. Huang, H.; Ma, X.; Hu, L.; Huang, Y.; Huang, H. The preliminary application of the combination of Fast R-CNN deep learning and UAV remote sensing in the monitoring of pine wilt disease. *J. Environ. Entomol.* **2021**, *43*, 1295–1303.
20. Vollenweider, P.; Günthardt-Goerg, M.S. Diagnosis of abiotic and biotic stress factors using the visible symptoms in foliage. *Environ. Pollut.* **2005**, *137*, 455–465. [CrossRef]
21. Liu, C.; Zhang, X.; Wang, T.; Chen, G.; Zhu, K.; Wang, Q.; Wang, J. Detection of vegetation coverage changes in the Yellow River Basin from 2003 to 2020. *Ecol. Indic.* **2022**, *138*, 108818. [CrossRef]
22. Zhang, X.; Chen, G.; Wang, W.; Wang, Q.; Dai, F. Object-based land-cover supervised classification for very-high-resolution UAV images using stacked denoising autoencoders. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 3373–3385. [CrossRef]
23. Zhang, X.; Tan, X.; Chen, G.; Zhu, K.; Liao, P.; Wang, T. Object-based classification framework of remote sensing images with graph convolutional networks. *IEEE Geosci. Remote Sens. Lett.* **2021**, *19*, 1–5. [CrossRef]
24. Xu, X.; Tao, H.; Li, C.; Cheng, C.; Guo, H.; Zhou, J. Detection and Location of Pine Wilt Disease Induced Dead Pine Trees Based on Faster R-CNN. *Trans. Chin. Soc. Agric. Mach.* **2020**, *51*, 228–236.
25. Jincang, L.; Chengbo, W.; Yuanfei, C. Monitoring method of bursaphelenchus xylophilus based on multi-feature CRF by UAV image. *Bull. Surv. Mapp.* **2019**, *25*, 78. [CrossRef]
26. You, J.; Zhang, R.; Lee, J. A Deep Learning-Based Generalized System for Detecting Pine Wilt Disease Using RGB-Based UAV Images. *Remote Sens.* **2021**, *14*, 150. [CrossRef]

27. Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Washington, DC, USA, 23–28 June 2014; pp. 580–587.

28. Girshick, R. Fast r-cnn. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 1440–1448.

29. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. In Proceedings of the Advances in Neural Information Processing Systems, Montreal, QC, Canada, 7–12 December 2015; Volume 28.

30. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 779–788.

31. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.; Berg, A.C. SSD: Single shot multibox detector. In Proceedings of the Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, 11–14 October 2016; pp. 21–37.

32. Zhang, X.; Zhu, K.; Chen, G.; Tan, X.; Zhang, L.; Dai, F.; Liao, P.; Gong, Y. Geospatial object detection on high resolution remote sensing imagery based on double multi-scale feature pyramid network. *Remote Sens.* **2019**, *11*, 755. [CrossRef]

33. Li, F.; Shen, W.; Wu, J.; Sun, F.; Xu, L.; Liu, Z.; Lan, P. Study on the Detection Method for Pinewood Wilt Disease Tree Based on YOLOv3-CloU. *J. Shandong Agric. Univ.* **2021**, *52*, 224–233.

34. Huang, L.; Wang, Y.; Xu, Q.; Liu, Q. Recognition of abnormally discolored trees caused by pine wilt disease using YOLO algorithm and UAV images. *Trans. Chin. Soc. Agric. Eng* **2021**, *37*, 197–203.

35. Chen, W.; Huihui, Z.; Jiwang, L.; Shuai, Z. Object Detection to the Pine Trees Affected by Pine Wilt Disease in Remote Sensing Images Using Deep Learning. *J. Nanjing Norm. Univ* **2021**, *44*, 84–89.

36. Xu, H.; Luo, Y.; Zhang, Q. Changes in water content, pigments and antioxidant enzyme activities in pine needles of Pinus thunbergii and Pinus massoniana affected by pine wood nematode. *Sci. Silvae Sin.* **2012**, *48*, 140–143.

37. dos Santos, C.S.S.; de Vasconcelos, M.W. Identification of genes differentially expressed in Pinus pinaster and Pinus pinea after infection with the pine wood nematode. *Eur. J. Plant Pathol.* **2012**, *132*, 407–418. [CrossRef]

38. Xu, H.; Luo, Y.; Zhang, T.; Shi, Y. Changes of reflectance spectra of pine needles in different stage after being infected by pine wood nematode. *Spectrosc. Spectr. Anal.* **2011**, *31*, 1352–1356.

39. Zhang, Y.; Ren, W.; Zhang, Z.; Jia, Z.; Wang, L.; Tan, T. Focal and efficient IOU loss for accurate bounding box regression. *Neurocomputing* **2022**, *506*, 146–157. [CrossRef]

40. Jocher, G. YOLOv5 by Ultralytics. 2020. Available online: https://doi.org/10.5281/zenodo.3908559 (accessed on 1 April 2023).

41. Redmon, J.; Farhadi, A. YOLO9000: Better, faster, stronger. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 7263–7271.

42. Wang, C.Y.; Liao, H.Y.M.; Wu, Y.H.; Chen, P.Y.; Hsieh, J.W.; Yeh, I.H. CSPNet: A new backbone that can enhance learning capability of CNN. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Seattle, WA, USA, 14–19 June 2020; pp. 390–391.

43. Lin, T.Y.; Dollár, P.; Girshick, R.; He, K.; Hariharan, B.; Belongie, S. Feature pyramid networks for object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 2117–2125.

44. Liu, S.; Qi, L.; Qin, H.; Shi, J.; Jia, J. Path aggregation network for instance segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 8759–8768.

45. He, K.; Zhang, X.; Ren, S.; Sun, J. Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2015**, *37*, 1904–1916. [CrossRef]

46. Zheng, Z.; Wang, P.; Liu, W.; Li, J.; Ye, R.; Ren, D. Distance-IoU loss: Faster and better learning for bounding box regression. In Proceedings of the AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 February 2020; Volume 34, pp. 12993–13000.

47. Zhu, K.; Zhang, X.; Chen, G.; Li, X.; Cai, P.; Liao, P.; Wang, T. Multi-Oriented Rotation-Equivariant Network for Object Detection on Remote Sensing Images. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 1–5. [CrossRef]

48. Chen, K.; Wang, J.; Pang, J.; Cao, Y.; Xiong, Y.; Li, X.; Sun, S.; Feng, W.; Liu, Z.; Xu, J.; et al. MMDetection: Open MMLab Detection Toolbox and Benchmark. *arXiv* **2019**, arXiv:1906.07155.

49. Van der Maaten, L.; Hinton, G. Visualizing Data using t-SNE. *J. Mach. Learn. Res.* **2008**, *9*, 2579–2605.