

Article

Aerial Image Segmentation of Nematode-Affected Pine Trees with U-Net Convolutional Neural Network

Jiankang Shen ¹, Qinghua Xu ¹, Mingyang Gao ¹, Jicai Ning ² , Xiaopeng Jiang ² and Meng Gao ^{1,*}

¹ School of Mathematics and Information Sciences, Yantai University, Yantai 264005, China; 13280062738@163.com (J.S.)

² Yantai Institute of Coastal Zone Research, Chinese Academy of Sciences, Yantai 264003, China; xjiang@yic.ac.cn (X.J.)

* Correspondence: mgao@ytu.edu.cn

Abstract: Pine wood nematode disease, commonly referred to as pine wilt, poses a grave threat to forest health, leading to profound ecological and economic impacts. Originating from the pine wood nematode, this disease not only causes the demise of pine trees but also casts a long shadow over the entire forest ecosystem. The accurate identification of infected trees stands as a pivotal initial step in developing effective prevention and control measures for pine wilt. Nevertheless, existing identification methods face challenges in precisely determining the disease status of individual pine trees, impeding early detection and efficient intervention. In this study, we leverage the capabilities of unmanned aerial vehicle (UAV) remote sensing technology and integrate the VGG classical small convolutional kernel network with U-Net to detect diseased pine trees. This cutting-edge approach captures the spatial and characteristic intricacies of infected trees, converting them into high-dimensional features through multiple convolutions within the VGG network. This method significantly reduces the parameter count while enhancing the sensing range. The results obtained from our validation set are remarkably promising, achieving a Mean Intersection over Union (MIoU) of 81.62%, a Mean Pixel Accuracy (MPA) of 85.13%, an Accuracy of 99.13%, and an F1 Score of 88.50%. These figures surpass those obtained using other methods such as ResNet50 and DeepLab v3+. The methodology presented in this research facilitates rapid and accurate monitoring of pine trees infected with nematodes, offering invaluable technical assistance in the prevention and management of pine wilt disease.

Keywords: UAV remote sensing; pine wood nematode; machine learning; U-Net



Citation: Shen, J.; Xu, Q.; Gao, M.; Ning, J.; Jiang, X.; Gao, M. Aerial Image Segmentation of Nematode-Affected Pine Trees with U-Net Convolutional Neural Network. *Appl. Sci.* **2024**, *14*, 5087. <https://doi.org/10.3390/app14125087>

Academic Editor: Emanuel Guariglia

Received: 8 April 2024

Revised: 7 June 2024

Accepted: 10 June 2024

Published: 11 June 2024



Copyright: © 2024 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

Pine wood nematode disease, also known as pine wilt, is a destructive and infectious forest disease caused by the pine wood nematode vector insects, specifically beetles of the genus *Monochamus*, and their host pine plants of the *Pinus* species. The pine wood nematode originated from North America [1,2], was introduced to Japan in the early 20th century, and was discovered in 1982 in Nanjing, China [3]. It has become a major foreign invasive species in China and has been listed as a forest plant quarantine object internally and externally by China. By the end of 2017, pine wood nematode disease had affected 19 provinces, 742 counties, and an area of more than 4667 km² in China, causing significant ecological and economic losses [4]. Therefore, monitoring the pine nematode is crucial for the control of pine nematode disease.

Pine nematode disease leads to substantial damage to pine trees, with infected pines typically dying within a short period of time. The progression of the disease results in a variety of symptoms in the pine tree, initially manifesting as subtle changes in the needles. As the disease worsens, the needles gradually turn yellowish-brown or even reddish-brown, and the pine tree wilts and eventually dies. The xylem of dead pines often exhibits a bluish-grey coloration due to the presence of the blue-variegated fungus.

Pine wood nematode disease primarily arises from the parasitism of pine wood nematodes, which are transmitted within pine trees by vector insects like the pine beetles. The disease typically progresses through four distinct stages: an initial stage where trees appear normal, followed by needle discoloration, then wilting of the tree, and ultimately, the death of the entire plant. When it comes to preventing and controlling pine wilt disease, monitoring the epidemic is crucial. In China, initial monitoring efforts primarily relied on manual surveys. However, due to limitations posed by roads, terrain, and various other complex factors, these methods often resulted in inefficient work productivity and limited effectiveness in epidemic prevention and control.

As science and technology progress, the employment of remote sensing technology to detect pine wood nematode-infested trees has emerged as a crucial technical approach for monitoring the spread of pine wood nematode disease. The primary foundation for identifying pine wilt disease-affected trees through remote sensing lies in the alterations in the spectral properties of plants, stemming from changes in water and chlorophyll content within infected specimens [5]. Studies focusing on remote sensing monitoring of pine wood nematode disease began to emerge in the 1990s [6], encompassing the utilization of spectroradiometers to measure spectral curves of diverse pine tree types infected with pine wilt disease, conducting differential analyses on these curves, as well as measuring reflectance spectra at various stages of infection and analyzing the associated spectral characteristics and chlorophyll variations [7]. Numerous investigations into the spectral features of pine wilt disease-infected trees have confirmed that analyzing remotely sensed data can effectively lead to the identification of diseased trees [8]. However, the relatively low resolution of early remote sensing satellites often leads to reduced accuracy in results and the potential omission of diseased trees [9].

Machine learning-based image segmentation techniques offer remarkable speed and precision, making them a popular choice for target recognition and segmentation in high-resolution remote sensing images [10]. Previous research has validated the effectiveness of these methods. For instance, traditional support vector machines (SVM) [11,12] have achieved enhanced results, while a study focusing on individual pine tree identification employed a modified LeNet architecture [13]. Furthermore, a stochastic deep forest algorithm was implemented to monitor pine nematode-infested trees [14]. Another approach for detecting pine diseases involves utilizing orthophoto-corrected 5-band multispectral images from unmanned aerial vehicles (UAVs), paired with a ResNet18 backbone network augmented with a modified DenseNet module for classification [15]. The VNet network, a fusion of VGG-16 and dilated convolution (DC), was also developed for pine disease detection. This network leverages the initial 13 layers of VGG-16 to capture fundamental features, coupled with 6 DC layers for extracting more intricate characteristics [16]. Furthermore, a recognition algorithm for pine wilt-affected wood has been crafted around the YOLOv5 framework, integrating Mosaic data augmentation, Focus unit, and Spatial Pyramid Pooling (SPP) unit [17]. Additionally, a multi-channel convolutional neural network (CNN)-based target detection method has been employed for detecting pine wilt-diseased trees, comparing detection accuracy between the combination of each channel and the utilization of vegetation indices [18].

When dealing with image-related tasks, traditional deep learning methods encounter significant challenges due to the sheer volume of data and the necessity to consider intricate spatial information. Key among these challenges are slow training speeds and the difficulty in achieving high-quality segmentation results, especially for fine segmentation tasks. Although deep learning methods have demonstrated their effectiveness in pine nematode tree monitoring in recent years [19–23], traditional approaches can still lead to the loss of crucial details during the feature extraction process. This loss of details can significantly impact the accuracy of segmentation results, especially in the case of pine nematode infestation, where preserving key information is vital for subsequent segmentation tasks. Therefore, it is imperative to optimize and enhance these methods to address these issues.

Fractal-wavelet analysis provides a new theoretical foundation and methodological support for the segmentation of pine wood nematode disease in images. This analysis encompasses various aspects, including Weierstrass–Mandelbrot fractal functions [24], fractional wavelet analysis, wavelet-based hyperspectral image classification, and adaptive multi-scale wavelet decomposition [25–27]. These theoretical studies offer fresh perspectives and methods for in-depth exploration in the field of image segmentation.

The introduction of fractional wavelet analysis allows for a more precise capture of local features and texture information in images, significantly enhancing the accuracy and stability of pine wood nematode disease image segmentation. Furthermore, the method of hyperspectral image classification based on wavelet transforms offers new insights for feature extraction and classification of pine wood nematode disease images. Additionally, the framework of adaptive multi-scale wavelet decomposition provides an efficient solution for dealing with complex structures and textures in images.

When addressing the challenges posed by small targets in UAV images, along with the issues of extensive training parameters and lengthy training cycles, this paper employs a modified U-Net architecture fused with VGG. This approach differs from the traditional U-Net by augmenting the network's layers to extract higher-dimensional feature information. By leveraging the properties of VGG's classical small convolution kernels, multiple convolution operations are performed, which significantly reduces the number of model parameters while achieving an excellent fitting effect. It has better performance than ResNet50 and DeepLab v3+. The key contributions of this paper are as follows:

- (1) Increase the depth of the U-Net network to facilitate high-dimensional feature extraction.
- (2) Using the characteristics of a small convolutional kernel of the VGG backbone network to significantly reduce the number of parameters, thereby speeding up training and improving segmentation accuracy.

This study, with a focus on Yantai in the Shandong Peninsula, aimed to devise a precise and efficient method for the rapid identification of pine trees infected with nematodes. The goal was to prevent the spread of infection to other healthy trees.

2. Materials and Methods

2.1. Materials

2.1.1. Study Area

In this study, forested areas with reported outbreaks of pine wood nematode disease were chosen as our test areas. Recognizing the importance of coastal zone protection forests as vital ecological barriers on the Shandong Peninsula, we specifically chose the coastal area in Laishan District, Yantai City, as our primary data collection location. To capture aerial imagery, we utilized an 8-rotor unmanned aerial vehicle (UAV) equipped with a GaiaSky-mini camera (DUALIX, Wuxi, China). We employed multispectral and visible light data to classify the various feature types within the surveyed region. Images showing non-vegetation areas, such as rocks, bare ground, or water bodies, as well as non-pine vegetation like broad-leaved forests, shrubs, or grasslands, were excluded. This filtering process ensured that only images of black pine trees were retained for analysis. Furthermore, we verified the accuracy of feature type identification through a combination of visual interpretation of the visible light data and on-site survey data. The GaiaSky-mini camera offers a spectral coverage range of 450–1000 nm, a spectral resolution of 3.5 nm, and a spatial resolution of 0.3 m when flown at an altitude of 100 m. The data are stored in hypercube format, enabling the capture of up to 380 bands of imagery. The main specifications of UAV and camera used in this study are summarized in Table 1.

We utilized Pix4DMapper 4.5.6 software for the orthorectification of UAV aerial images. Orthorectification involves selecting ground control points on the image and leveraging the inherent digital elevation model (DEM) data to simultaneously correct image tilt and projection discrepancies. Subsequently, the image is resampled into an orthophoto. After mosaicing multiple orthophotos and performing color balance adjustments, the final orthophoto image is cropped to a specific range. Ultimately, we obtained orthophoto maps

and their corresponding sparse landmark models for three distinct regions. Additionally, we calculated the number of overlapping images per pixel of the orthophoto map. Take area (b) as an example in Figures 1 and 2.

Table 1. Drone models and cameras.

UAV & Camera	Parameters	Value
	Name	DJI_M600 (DJI, Shenzhen, China)
	Maximum take-off weight	11 kg
	Weights	4.4 kg
	Endurance	15 min
	Effective working hours	12 min
	Name	GaiaSky-mini
	Aperture	f/3.5
	Exposure time	1/240 s
	ISO speed	100~200
	Exposure compensation	0-stop aperture
	Monitoring range	15 mm
	Maximum aperture	1.7 mm
	Spectral coverage	450~1000 nm
	Spectral resolution	3.5 nm

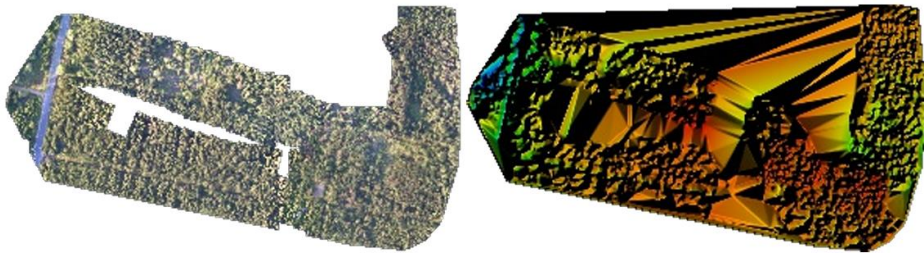


Figure 1. Orthomosaic and the corresponding sparse Digital Surface Model (DSM) before densification.

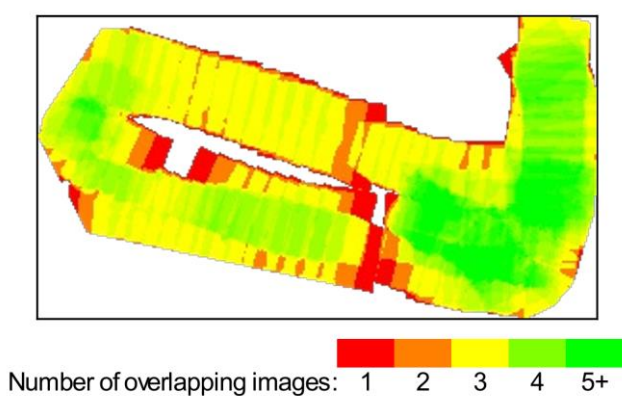


Figure 2. Number of overlapping images computed for each pixel of the orthomosaic. Red and yellow areas indicate low overlap for which poor results may be generated. Green areas indicate an overlap over 5 images for every pixel. Good quality results will be generated as long as the number of keypoint matches is also sufficient for these areas.

The three study areas are located in (a) Binhai East Road, Yantai City, Shandong Province, (b) Jingliu Road, and (c) Guanzhuang Road. The geographic locations of these images on the map are in Figure 3.

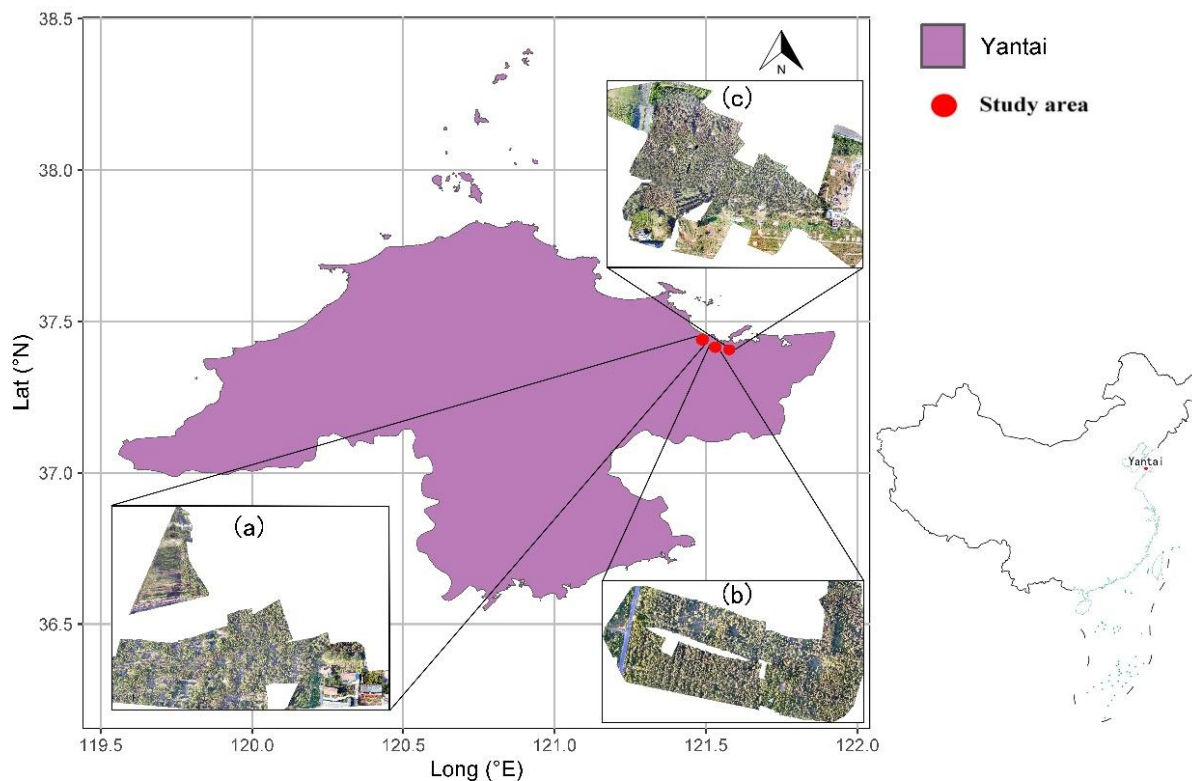


Figure 3. Location of Yantai City and its study area. (a) Binhai East Road, Yantai City, Shandong Province, (b) Jingliu Road, and (c) Guanzhuang Road.

Because of the considerable geographical overlap in some captured photos, it was advantageous to omit certain images from the dataset. This helped avoid repetitive learning of analogous image features or geographical conditions during model training, ultimately improving training efficiency and reducing overall training time. The samples excluded due to these issues are designated as invalid samples. Following their elimination, 500 drone images were organized. These 500 images were then processed, converted, and calibrated to ENVI format using ArcMap v. 10.0.

2.1.2. Input Data

In this study, a single aerial image captured by a UAV was employed for training purposes. Visual interpretation, a prevalent method in remote sensing image analysis, was used. An image analyst visually interpreted the image to discern diverse land surface features. The criteria for labeling typically encompass shape features, texture features, contextual information, scale, and direction. Leveraging these features, various elements were accurately identified and labeled through visual interpretation during the image annotation process. As a result, the corresponding labels for infected trees were obtained. The image boasts a resolution of 4608×3456 pixels and is composed of three bands: red (central wavelength: 650 nm), green (552 nm), and blue (475 nm).

To evaluate the model's effectiveness, the entire image depicting pine nematode-infested trees underwent manual labeling utilizing LabelMe. Subsequently, all sample data were randomly allocated into training, validation, and test sets at a 3:1:1 ratio. Validation was performed on training samples collected from three distinct locations, encompassing 100 UAV images in the validation set. Some example datasets are in Figure 4.

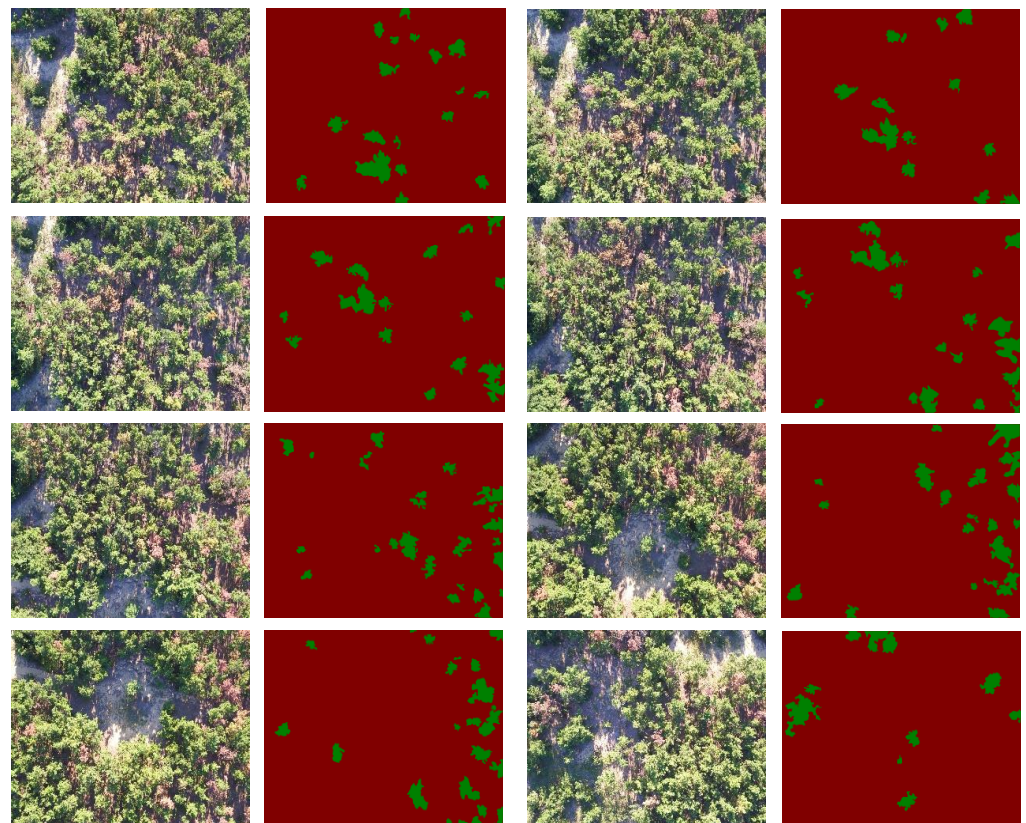


Figure 4. UAV images and corresponding labels.

2.2. Methods

2.2.1. Image Segmentation Dataset

To annotate the pine trees with disease, we utilized the LabelMe tool to outline the affected trees and saved the annotation data as a JSON file. In small-scale datasets, neural networks are prone to weak generalization, and using data augmentation methods such as mirroring, rotation, flipping, and scaling can increase sample diversity, making the network more stable during training [28]. Therefore, to address the problem of limited sample size in this experiment, we choose to use mirroring and rotation to augment the training set, in order to improve the performance of the model in segmenting diseased pine trees infected with wilting disease.

2.2.2. U-Net Construction

In semantic segmentation tasks, the receptive field size is paramount for achieving precise segmentation outcomes. Existing techniques, including the Atrous Spatial Pyramid Pooling (ASPP) model [29,30], Pyramid Pooling Model [31], and the employment of large convolutional kernels [32], strive to enlarge the receptive field. However, these approaches often involve a substantial computational cost for determining model parameters, consuming significant memory and consequently leading to slower training speeds. To tackle this challenge, this paper leverages VGG as the backbone network and integrates it with the U-Net architecture. The depth of the U-Net network enables the extraction of high-dimensional features from the input image. By bridging corresponding layers between the encoder and decoder paths, the proposed method effectively fuses low-level features with high-level features, ultimately enhancing segmentation accuracy.

Using VGG has several advantages. Firstly, the number of parameters in the multi-layer small convolution kernel is significantly smaller than that of a large convolution kernel. Secondly, the multi-layer small convolution kernel performs nonlinear activation only once, followed by multiple linear combinations and nonlinear activations, thereby capturing

higher-order functions and enhancing the expressive ability of the network. For example, a 3×3 convolutional kernel commonly used in VGG is equivalent to performing two convolutions with a 5×5 convolutional kernel while reducing the number of parameters to 72% of the original size. This comparison highlights the efficiency of the 3×3 convolutional kernel in terms of parameter reduction while convolving on the same image.

By leveraging the fusion of VGG and U-Net, this proposed method achieves an effective combination of low-level and high-level features, leading to improved segmentation accuracy without the need for computationally intensive operations or excessive memory consumption.

The feature information extracted from the VGG backbone network is fused with the up-sampled feature information to obtain more profound feature insights. Unlike the conventional U-Net structure, this approach involves increasing the number of layers in the U-Net to extract higher-dimensional features. The left portion of the network, representing the encoder, utilizes the VGG backbone network to extract input image features. VGG16 comprises a total of 16 layers structured into five convolutional blocks, each containing three microstructures. Within these microstructures, 3×3 small convolutional kernels are chosen due to their superior fitting effect and lower computational complexity.

On the other hand, the right portion of the network acts as the decoder and consists of the up-sampling module and skip connections. The up-sampling module expands the input tensor by a factor of two through up-sampling operations and connects the feature tensor from the previous encoder layer to the up-sampled result. A series of convolution and ReLU activation function operations are then applied to generate the final output.

In the U-Net architecture, the concatenation splicing layer plays a crucial role by connecting the encoder feature tensor in each up-sampling layer to the corresponding higher-resolution feature. This layer maintains spatial information, thereby enhancing segmentation accuracy, and facilitates faster learning of mapping relationships between features. Additionally, U-Net employs skip connections to pass encoder features directly to the decoder. This strategy enables the decoder to leverage semantic information from different levels, ultimately improving segmentation performance. By facilitating the direct transfer of underlying feature information to the decoder, skip connections help alleviate information loss issues. The U-Net structure of this paper is in Figure 5.

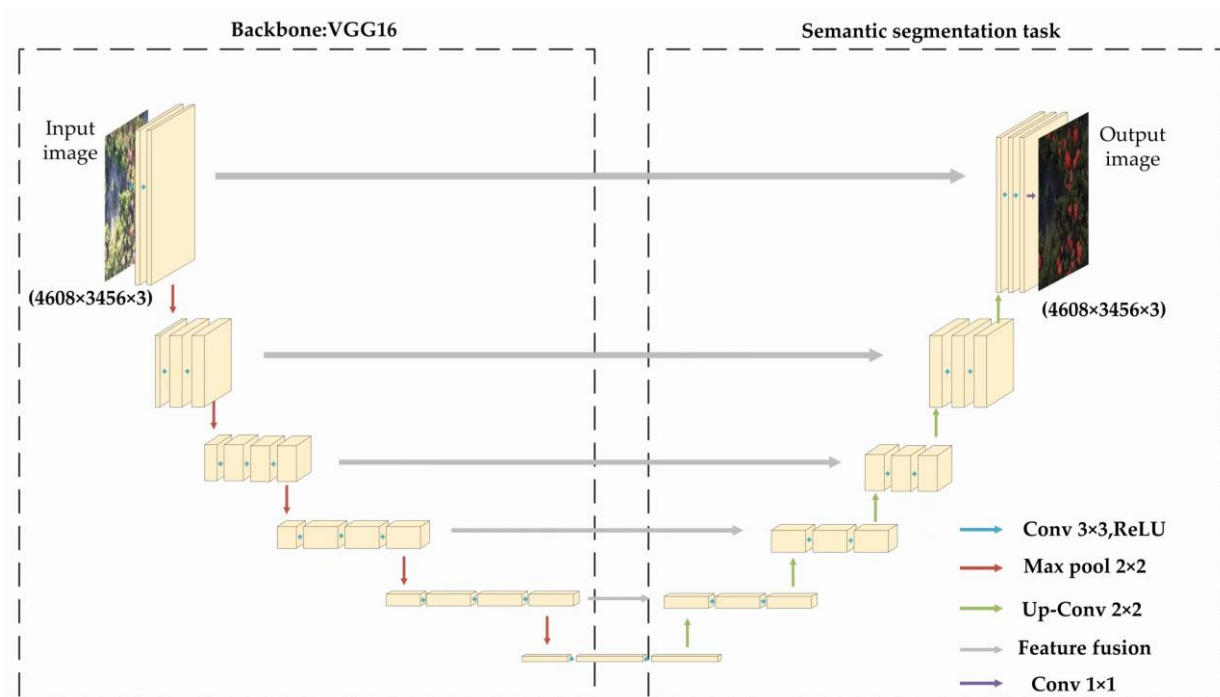


Figure 5. U-Net structure.

2.2.3. Loss Function

For the model's loss function, we utilize the cross-entropy loss, which is a commonly employed loss function in machine learning. It is primarily utilized to quantify the variance between a model's output and the actual labels and subsequently refine its performance. When making binary classifications for each pixel, different pixels carry varying degrees of importance. Pixels corresponding to the foreground hold greater significance, thus necessitating the use of larger weights to accentuate their influence, denoted by α . Additionally, for the edge pixels of the object to be segmented, the computational challenges are greater, leading to the introduction of weight Y . Consequently, the loss function for an individual pixel can be expressed as:

$$\text{loss} = -\alpha(1 - y_{\text{pred}})^Y \cdot y_{\text{true}} \log_2(y_{\text{pred}}) - (1 - \alpha)y_{\text{pred}}^Y \cdot (1 - y_{\text{true}}) \log_2(1 - y_{\text{pred}}) \quad (1)$$

$$y_{\text{true}} \in \{0, 1\} \quad (2)$$

$$y_{\text{pred}} \in [0, 1] \quad (3)$$

where y_{true} represents binary value (0 or 1), indicating the real label of the sample. y_{pred} represents the probability that the model predicts the sample as a disease tree.

This loss function offers the advantage of considering the impact of background factors on the target pixel points while also improving the handling of target edges. This aspect plays a crucial role in the subsequent segmentation of UAV images.

2.2.4. Evaluation Systems

In the image segmentation task, the objective is to ascertain whether the characteristics of a pixel in the image mimic those of a pine nematode-infested tree, while considering the influence of neighboring pixels. To address the challenge of segmenting remote sensing UAV images, we utilize four key metrics to evaluate segmentation accuracy. Mean Intersection over Union (*MIoU*) gauges the precision of image segmentation by quantifying the overlap between the model's predicted segmentation and the ground truth. Mean Pixel Accuracy (*MPA*) reflects the average proportion of pixels accurately predicted by the model. Accuracy signifies the percentage of correctly categorized pixels across the entire image. The *F1 Score*, being the harmonic mean of Precision and Recall, is invaluable for evaluating the model's performance in handling both positive and negative instances. Additionally, the Confusion Matrix, a 2×2 structure, aids in computing the *F1 Score* and outlines four possible outcomes for monitoring the pine nematode tree model during the prediction phase.

True Positive (TP): a true example, indicating that the model correctly judges diseased tree pixel points (Positive) as diseased tree pixel points.

True Negative (TN): true negative example, indicating that the model correctly judged other pixel points (Negative) as other pixel points.

False Positive (FP): a False Positive example, indicating that the model incorrectly judges other pixel points as diseased tree pixel points.

False Negative (FN): false negative example, indicating that the model incorrectly judged the diseased tree pixel point as other pixel points.

To evaluate the segmentation accuracy of the model, we selected four metrics: *MIoU*, *MPA*, *Accuracy*, and *F1 Score*. The mathematical expressions for these metrics are as follows:

$$\text{Accuracy} = \frac{P_{tp} + P_{tn}}{P_{tp} + P_{fp} + P_{tn} + P_{fn}} \quad (4)$$

$$\text{F1 Score} = \frac{2 \cdot P_{tp}}{(2P_{tp} + P_{fp} + P_{fn})} \quad (5)$$

$$MIoU = \frac{1}{N} \sum \frac{P_{tp}}{P_{tp} + P_{fp} + P_{fn}} \quad (6)$$

$$MPA = \frac{1}{N} \sum \frac{P_{tp}}{P_{tp} + P_{fp}} \quad (7)$$

2.2.5. Model Training

The model underwent 82 epochs of training in total. During the first 50 epochs, a frozen training approach was employed, which made use of pre-trained weights. By freezing the training, the backbone network (VGG16) of the model was kept locked, permitting fine-tuning of only selected layers. This method accelerated the training process, preserved effective feature extraction abilities, and precluded the introduction of extraneous parameters in a small-scale dataset. For the remaining 32 epochs, the model underwent unfrozen training, allowing all parameters to become trainable and adjusted according to the present task. Unfreezing the model's parameters permitted gradient backpropagation across the entirety of the model, enabling thorough tuning and optimization. Below, we present the loss function curves for the iterations and the Mean Intersection over Union (*MIoU*) curves derived from the validation set.

3. Results

3.1. Model Training Results

Upon analysis of Figure 6, it is evident that the loss stabilizes for both the training and validation sets after approximately 20 epochs. Specifically, the minimum total loss achieved for the training set is 0.013, occurring in the 81st epoch, with a corresponding validation set loss of 0.018. The validation set attains its minimum loss of 0.016 when the total loss for the training set stands at 0.019. Moreover, the *MIoU* demonstrates stable performance after 60 epochs, culminating in a peak value of 82.79 in the 82nd epoch. At this point, the *MPA* registers at 88.0, and *Accuracy* stands at 99.15. Following the 60th epoch, the changes flatten out, resulting in the *MIoU* achieving its highest value of 82.79 in the 82nd epoch, along with an *MPA* of 88.0 and *Accuracy* of 99.15. To evaluate the trained model, we tested the training weights from the 82nd epoch on the test samples, yielding the following partial images.

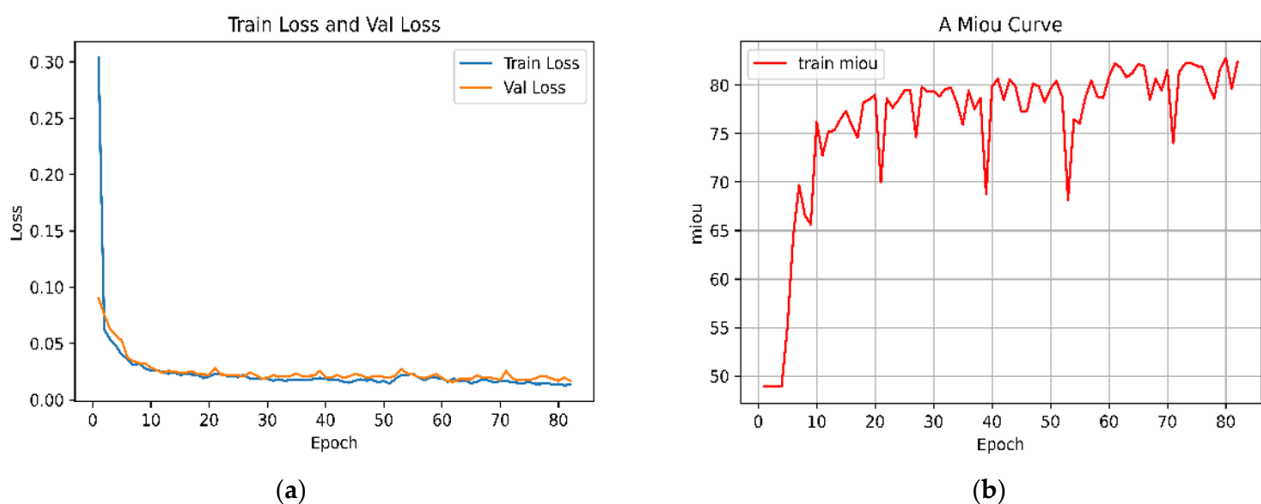


Figure 6. (a) Training set loss and validation set loss; (b) *MIoU* curves.

3.2. Results of Monitoring Pine Nematode Disease Trees

After training, this model detects and outlines the infected trees with pine wilt disease in the test image set, as shown in Figure 7.

In the predicted result images, red pixels highlight the presence of trees predicted to be infected with pine wood nematode disease, while the darkened background serves to

accentuate the visualization of the prediction outcomes. These predicted images offer a straightforward visual assessment of the model's accuracy. Notably, the model exhibits improved precision in identifying pine wilt trees across diverse scenarios, including those featuring rivers, open spaces, and other potential disturbances in drone imagery. Furthermore, for drone images devoid of pine wilt trees, the prediction results accurately reflect the absence of diseased trees. The evaluation indicators for different generations of training models are shown in Table 2.

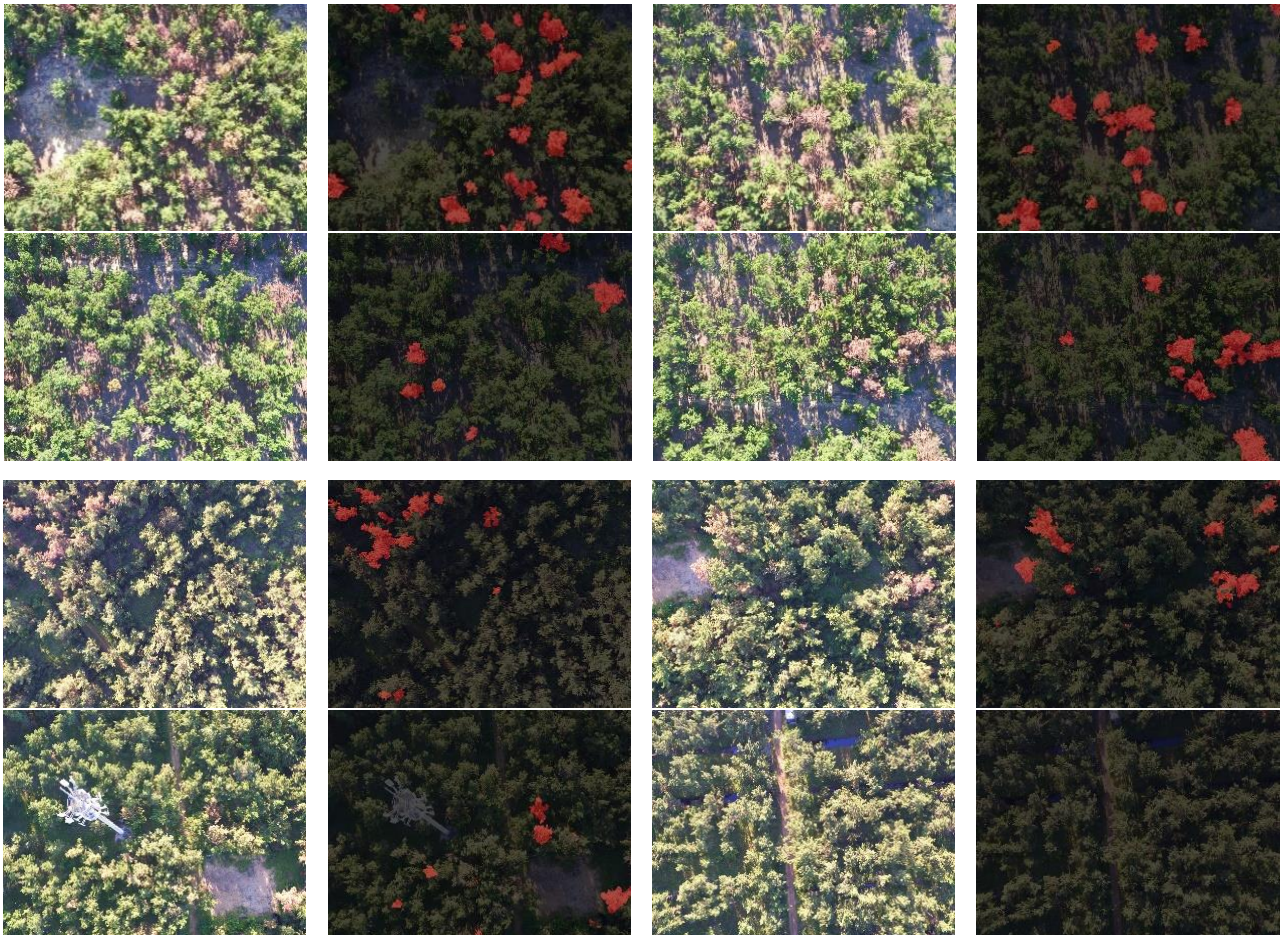


Figure 7. The UAV images and predicted results.

Table 2. Corresponding epochs at the maximum value of the indicator.

Target Epoch	Index			
	<i>MIoU</i>	<i>MPA</i>	<i>Accuracy</i>	<i>F1 Score</i>
82	82.79	88.0	99.15	0.881
77	81.83	90.83	99.01	0.877
81	81.62	85.13	99.13	0.885

For each epoch, the maximum values of the different performance indicators are summarized in a relevant data table. This compilation offers a thorough perspective on the model's performance throughout multiple epochs, emphasizing its achievements in predictive accuracy and disease detection. Notably, it is apparent that both *MIoU* and *F1 Score* have attained notably high values at the 81st epoch. Consequently, we have selected the weights from this epoch as optimal for predicting pine nematode-infested trees in UAV imagery.

3.3. Comparison of Model Results

To fully understand the performance of the trained model, we compared it with other image segmentation models, ResNet50 and DeepLab v3+, in terms of training time, convergence speed, and related segmentation performance evaluation metrics, including *MIoU*, *PMA*, *Accuracy*, and *F1 Score* in the Table 3.

Table 3. Performance comparison of different models.

Model	Training Time	Convergence Rate (Time)	<i>MIoU</i>	<i>MPA</i>	<i>Accuracy</i>	<i>F1 Score</i>
VGG-U-Net	16 min, 32 s	9 epochs	81.62	85.13	99.13	0.885
ResNet50	32 min, 53 s	25 epochs	71.5	74.6	98.78	0.765
DeepLab v3+	15 min, 46 s	18 epochs	48.9	50	73.5	0.497

Upon reviewing the table, it is evident that the VGG-U-Net network surpasses the other two image segmentation models in terms of both training efficiency and accuracy on the validation set. Across all evaluation metrics, VGG-U-Net consistently outperforms its counterparts. To offer a more tangible illustration of the disparities between the three models, predictions were conducted using untrained images, and a comparative analysis of the prediction differences was undertaken. The results are depicted in Figure 8.

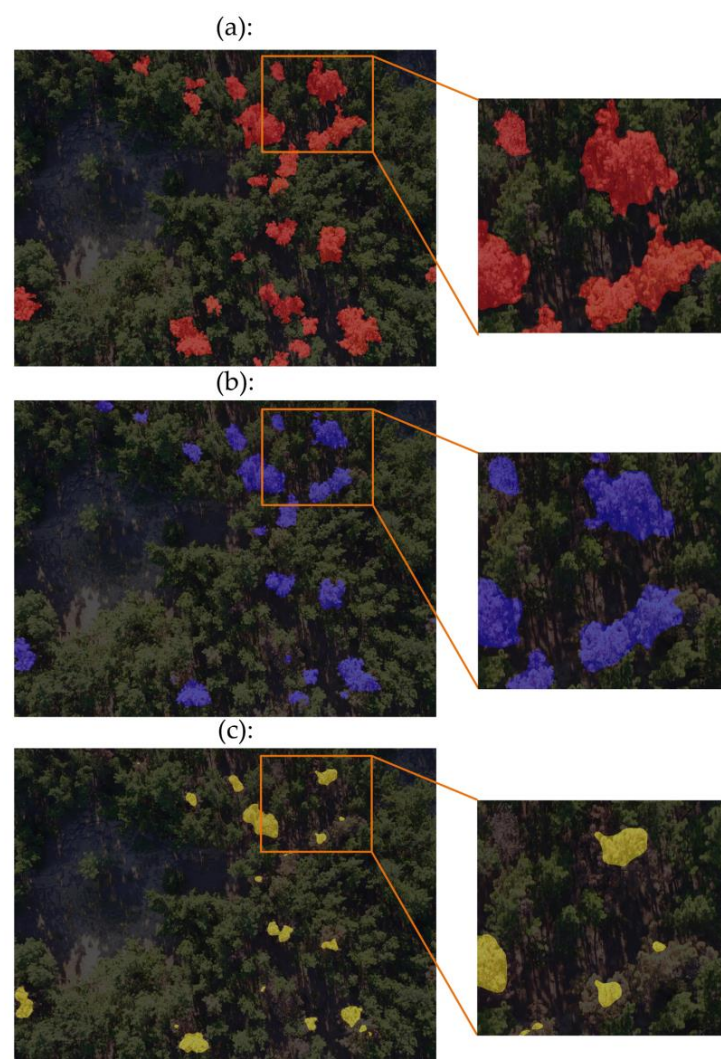


Figure 8. The performance of different models on the test set: (a) VGG-U-Net, (b) ResNet50, and (c) DeepLab v3+.

Upon careful analysis of the graph, it becomes apparent that the VGG-U-Net model demonstrates superior accuracy and precision in disease localization and tree edge segmentation. This signifies that the model is capable of more accurately pinpointing the disease location within the image and delivering clearer segmentation of tree edges. In contrast, while the other two models can also detect the disease position in the image, their segmentation of the disease and tree boundaries appears relatively blurred, potentially causing inaccuracies in localization and identification. Consequently, this finding underscores the remarkable proficiency of the U-Net model in managing intricate edge details, especially excelling in delicate aspects such as the contours of infected trees.

4. Conclusions

Originating from the pine wood nematode, pine wilt disease poses a significant threat, not only causing the demise of pine trees but also having profound impacts on the overall forest ecosystem. The accurate identification of infected trees stands as a vital first step in formulating strategies to prevent and control this disease. Nevertheless, existing identification techniques face challenges in precisely determining the disease status of individual pine trees, impeding early detection and efficient management measures.

In this study, we harness the power of unmanned aerial vehicle (UAV) remote sensing technology and integrate the VGG classical small convolutional kernel network with U-Net to identify diseased pine trees. This cutting-edge method skillfully captures the spatial and characteristic complexities of infected trees, converting them into high-dimensional features through multiple convolutions of the VGG network. This approach notably decreases the parameter count while expanding the sensing range. Our validation set results are remarkably encouraging, achieving a Mean Intersection over Union (*MIoU*) of 81.62%, a Mean Pixel Accuracy (*MPA*) of 85.13%, an overall *Accuracy* of 99.13%, and an *F1 Score* of 88.50%. This method surpasses the other two approaches, highlighting its efficacy. The research technique introduced here facilitates swift and accurate monitoring of nematode-infected pine trees, offering essential technical backing for the prevention and control of pine wilt disease.

Author Contributions: Conceptualization, M.G. (Meng Gao) and J.N.; methodology, J.S. and J.N.; software, J.S., J.N., X.J. and M.G. (Mingyang Gao); validation, J.S., Q.X. and M.G. (Mingyang Gao); formal analysis, X.J., J.S., J.N. and M.G. (Meng Gao); investigation, J.S., Q.X. and M.G. (Mingyang Gao); resources, X.J., M.G. (Meng Gao) and J.N.; data curation, X.J., M.G. (Meng Gao) and J.N.; writing—original draft preparation, J.S., Q.X. and M.G. (Mingyang Gao); writing—review and editing, J.S., Q.X. and M.G. (Meng Gao); visualization, J.S., Q.X. and M.G. (Meng Gao); supervision, M.G. (Meng Gao); project administration, M.G. (Meng Gao); funding acquisition, M.G. (Meng Gao). All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Key Program of Shandong Natural Science Foundation, grant number ZR2020KF031.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author, M.G., upon reasonable request. The data are not publicly available due to privacy.

Acknowledgments: The authors are grateful to Weitao Shang and Yueqi Wang in collecting the datasets, and Yuke Gan in revising the manuscript.

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

References

1. Ikegami, M.; Jenkins, T.A.R. Estimate global risks of a forest disease under current and future climates using species distribution model and simple thermal model—Pine Wilt disease as a model case. *For. Ecol. Manag.* **2018**, *409*, 343–352. [\[CrossRef\]](#)
2. Ichihara, Y.; Fukuda, K.; Suzuki, K. Early symptom development and histological changes associated with migration of *Bursaphelenchus xylophilus* in seedling tissues of *Pinus thunbergii*. *Plant Dis.* **2000**, *84*, 675–680. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Sun, H.; Zhou, Y.; Li, X.; Zhang, Y.; Wang, Y. Occurrence of major forest pests in 2020 and prediction of occurrence trend in 2021 in China. *For. Pest Dis.* **2021**, *40*, 45–48.
4. Jiang, M.; Huang, B.; Yu, X.; Zheng, W.T.; Lin, Y.L.; Liao, M.N.; Ni, J. Distribution, damage and control of pine wilt disease. *J. Zhejiang For. Sci. Technol.* **2018**, *38*, 83–91.
5. 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. [\[CrossRef\]](#)
6. Li, X. Use satellite remote sensing data to grasp the surgery of forests. *World For. Res.* **1992**, *50*. [\[CrossRef\]](#)
7. Wang, Z.; Zhang, X.L.; An, S.J. spectral characteristics analysis of pinus massoniana suffered by *Bursaphelenchus xylophilus*. *Remote Sens. Technol. Appl.* **2007**, *22*, 367–370.
8. Xu, H.C.; Luo, Y.Q.; Zhang, T.T.; Shi, Y.J. Changes of reflectance spectra of pine needles in different stage after being infected by pine wood nematode. *Spectrosc. Spectr. Anal.* **2011**, *31*, 1352–1356.
9. 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 Diseases. *IEEE Access* **2021**, *9*, 66346–66360. [\[CrossRef\]](#)
10. Lee, D.S.; Choi, W.I.; Nam, Y.; Park, Y.S. Predicting potential occurrence of pine wilt disease based on environmental factors in South Korea using machine learning algorithms. *Ecol. Inform.* **2021**, *64*, 101378. [\[CrossRef\]](#)
11. Xiong, Y.; Zhang, Z.; Chen, F. Comparison of Artificial Neural Network and Support Vector Machine Methods for Urban Land Use/Cover Classifications from Remote Sensing Images. In Proceedings of the 2010 International Conference on Computer Application and System Modeling (ICCSM 2010), Taiyuan, China, 22–24 October 2010.
12. Zhang, S.; Huang, H.; Huang, Y.; Cheng, D.; Huang, J. A GA and SVM Classification Model for Pine Wilt Disease Detection Using UAV-Based Hyperspectral Imagery. *Appl. Sci.* **2022**, *12*, 6676. [\[CrossRef\]](#)
13. 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. [\[CrossRef\]](#)
14. Zhang, Y.; Feng, W.; Quan, Y.; Zhong, X.; Song, Y.; Li, Q.; Dauphin, G.; Wang, Y.; Xing, M. A Novel Spatial-Spectral Random Forest Algorithm for Pine WILT Monitoring. In Proceedings of the IGARSS 2022—2022 IEEE International Geoscience and Remote Sensing Symposium, Kuala Lumpur, Malaysia, 17–22 July 2022; pp. 6045–6048. [\[CrossRef\]](#)
15. Zhang, R.; You, J.; Lee, J. Detecting Pine Trees Damaged by Wilt Disease Using Deep Learning Techniques Applied to Multi-Spectral Images. *IEEE Access* **2022**, *10*, 39108–39118. [\[CrossRef\]](#)
16. Zhang, L.; Huang, W.; Wang, J. Counting of Pine Wood Nematode Based on VNet Convolutional Neural Network. In Proceedings of the 2022 4th International Conference on Robotics and Computer Vision (ICRCV), Wuhan, China, 25–27 September 2022; pp. 164–168. [\[CrossRef\]](#)
17. Gong, H.; Ding, Y.; Li, D.; Wang, W.; Li, Z. Recognition of Pine Wood Affected by Pine Wilt Disease Based on YOLOv5. In Proceedings of the 2022 China Automation Congress (CAC), Xiamen, China, 25–27 November 2022; pp. 4753–4757. [\[CrossRef\]](#)
18. 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. [\[CrossRef\]](#)
19. Huang, J.; Lu, X.; Chen, L.; Sun, H.; Wang, S.; Fang, G. Accurate Identification of Pine Wood Nematode Disease with a Deep Convolution Neural Network. *Remote Sens.* **2022**, *14*, 913. [\[CrossRef\]](#)
20. Qin, B.; Sun, F.; Shen, W.; Dong, B.; Ma, S.; Huo, X.; Lan, P. Deep learning-based pine nematode trees' identification using multispectral and visible UAV imagery. *Drones* **2023**, *7*, 183. [\[CrossRef\]](#)
21. Deng, X.; Tong, Z.; Lan, Y.; Huang, Z. Detection and Location of Dead Trees with Pine Wilt Disease Based on Deep Learning and UAV Remote Sensing. *AgriEngineering* **2020**, *2*, 294–307. [\[CrossRef\]](#)
22. Li, H.; Chen, L.; Yao, Z.; Li, N.; Long, L.; Zhang, X. Intelligent Identification of Pine Wilt Disease Infected Individual Trees Using UAV-Based Hyperspectral Imagery. *Remote Sens.* **2023**, *15*, 3295. [\[CrossRef\]](#)
23. Lee, M.-G.; Cho, H.-B.; Youm, S.-K.; Kim, S.-W. Detection of Pine Wilt Disease Using Time Series UAV Imagery and Deep Learning Semantic Segmentation. *Forests* **2023**, *14*, 1576. [\[CrossRef\]](#)
24. Berry, M.V.; Lewis, Z.V.; Nye, J.F. On the Weierstrass-Mandelbrot fractal function. *Proc. R. Soc. Lond. A Math. Phys. Sci.* **1980**, *370*, 459–484.
25. Guariglia, E.; Silvestrov, S. Fractional-Wavelet Analysis of Positive definite Distributions and Wavelets on $D'(C)$ $D'(C)$. In *Engineering Mathematics II: Algebraic, Stochastic and Analysis Structures for Networks, Data Classification and Optimization*; Springer International Publishing: Berlin/Heidelberg, Germany, 2016; pp. 337–353.
26. Yang, L.; Su, H.; Zhong, C.; Meng, Z.; Luo, H.; Li, X.; Tang, Y.Y.; Lu, Y. Hyperspectral image classification using wavelet transform-based smooth ordering. *Int. J. Wavelets Multiresolut. Inf. Process.* **2019**, *17*, 1950050. [\[CrossRef\]](#)
27. Zheng, X.; Tang, Y.Y.; Zhou, J. A framework of adaptive multiscale wavelet decomposition for signals on undirected graphs. *IEEE Trans. Signal Process.* **2019**, *67*, 1696–1711. [\[CrossRef\]](#)

28. Li, W.; An, B.; Kong, Y. Data Augmentation Method on Pine Wilt Disease Recognition. In Proceedings of the International Conference on Intelligence Science, Xi'an, China, 28–31 October 2022; Springer International Publishing: Cham, Switzerland, 2022; pp. 458–465.
29. Chen, L.C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A.L. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE Trans. Pattern Anal. Mach. Intell.* **2016**, *40*, 834–848. [[CrossRef](#)] [[PubMed](#)]
30. Chen, L.C.; Papandreou, G.; Schroff, F.; Adam, H. Rethinking Atrous Convolution for Semantic Image Segmentation. *arXiv* **2017**, arXiv:1706.05587v2.
31. Zhao, H.; Shi, J.; Qi, X.; Wang, X.; Jia, J. Pyramid scene parsing network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017.
32. Peng, C.; Zhang, X.; Yu, G.; Luo, G.; Sun, J. Large Kernel Matters—Improve Semantic Segmentation by Global Convolutional Network. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 1743–1751.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.