



## Article

# Three-Dimensional Convolutional Neural Network Model for Early Detection of Pine Wilt Disease Using UAV-Based Hyperspectral Images

Run Yu <sup>1</sup>, Youqing Luo <sup>1,2,\*</sup> , Haonan Li <sup>1</sup>, Liyuan Yang <sup>1</sup>, Huaguo Huang <sup>3</sup> , Linfeng Yu <sup>1</sup> and Lili Ren <sup>1,2</sup>

- <sup>1</sup> Key Laboratory for Forest Pest Control, College for Forestry, Beijing Forestry University, Beijing 100083, China; yurun@bjfu.edu.cn (R.Y.); Lhn3190115@bjfu.edu.cn (H.L.); Yr0801@bjfu.edu.cn (L.Y.); ylf1801@bjfu.edu.cn (L.Y.); lily\_ren@bjfu.edu.cn (L.R.)
- <sup>2</sup> Sino-French Joint Laboratory for Invasive Forest Pests in Eurasia, Beijing Forestry University—French National Research Institute for Agriculture, Food and Environment (INRAE), Beijing 100083, China
- <sup>3</sup> Research Center of Forest Management Engineering of State Forestry and Grassland Administration, Beijing Forestry University, Beijing 100083, China; huaguo\_huang@bjfu.edu.cn
- \* Correspondence: youqingluo@126.com; Tel.: +86-10-6233-6840

**Abstract:** As one of the most devastating disasters to pine forests, pine wilt disease (PWD) has caused tremendous ecological and economic losses in China. An effective way to prevent large-scale PWD outbreaks is to detect and remove the damaged pine trees at the early stage of PWD infection. However, early infected pine trees do not show obvious changes in morphology or color in the visible wavelength range, making early detection of PWD tricky. Unmanned aerial vehicle (UAV)-based hyperspectral imagery (HI) has great potential for early detection of PWD. However, the commonly used methods, such as the two-dimensional convolutional neural network (2D-CNN), fail to simultaneously extract and fully utilize the spatial and spectral information, whereas the three-dimensional convolutional neural network (3D-CNN) is able to collect this information from raw hyperspectral data. In this paper, we applied the residual block to 3D-CNN and constructed a 3D-Res CNN model, the performance of which was then compared with that of 3D-CNN, 2D-CNN, and 2D-Res CNN in identifying PWD-infected pine trees from the hyperspectral images. The 3D-Res CNN model outperformed the other models, achieving an overall accuracy (OA) of 88.11% and an accuracy of 72.86% for detecting early infected pine trees (EIPs). Using only 20% of the training samples, the OA and EIP accuracy of 3D-Res CNN can still achieve 81.06% and 51.97%, which is superior to the state-of-the-art method in the early detection of PWD based on hyperspectral images. Collectively, 3D-Res CNN was more accurate and effective in early detection of PWD. In conclusion, 3D-Res CNN is proposed for early detection of PWD in this paper, making the prediction and control of PWD more accurate and effective. This model can also be applied to detect pine trees damaged by other diseases or insect pests in the forest.



**Citation:** 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. <https://doi.org/10.3390/rs13204065>

Academic Editors: Liangxiu Han, Wenjiang Huang, Yanbo Huang, Jiali Shang and Stefano Pignatti

Received: 2 September 2021  
Accepted: 8 October 2021  
Published: 11 October 2021

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

**Keywords:** pine wilt disease; early detection; UAV-based hyperspectral imagery; 3D-CNN; 3D-Res CNN



**Copyright:** © 2021 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 wilt disease (PWD, also known as “cancer” of pine trees), caused by the pine wood nematode (PWN; *Bursaphelenchus xylophilus*), is one of the most harmful and potential international quarantine forest diseases [1]. PWD originated in North America but now widely occurs worldwide (Figure 1) [2–5], causing tremendous damages to the global forest ecosystems. In a natural environment, the pathogenic mechanism of PWD is as follows. When vector insects that carry the PWN emerge from the pine tree, they locate and feed on the bark of young shoots of pine tree branches, creating wounds to the pine tree [6]. Then, the PWN invades the wound and eats the xylem of the pine tree [7,8], resulting in blockage of the tree's vessel. Finally, the transpiration of the pine tree gradually loses its function,

and the water absorbed by the root cannot reach the crown; thus, the pine tree needles wither, and eventually the whole pine tree dies. The detailed process of PWN infection is shown in Figure 2.

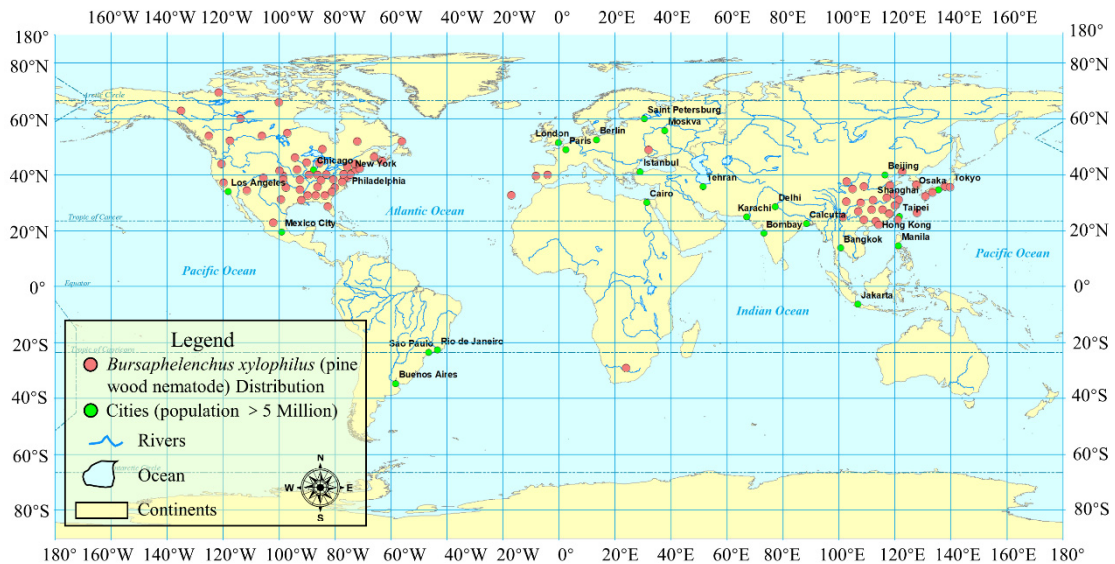


Figure 1. The global distribution of pine wilt disease (data source: <https://www.cabi.org/ISC>; accessed on 16 August 2021).

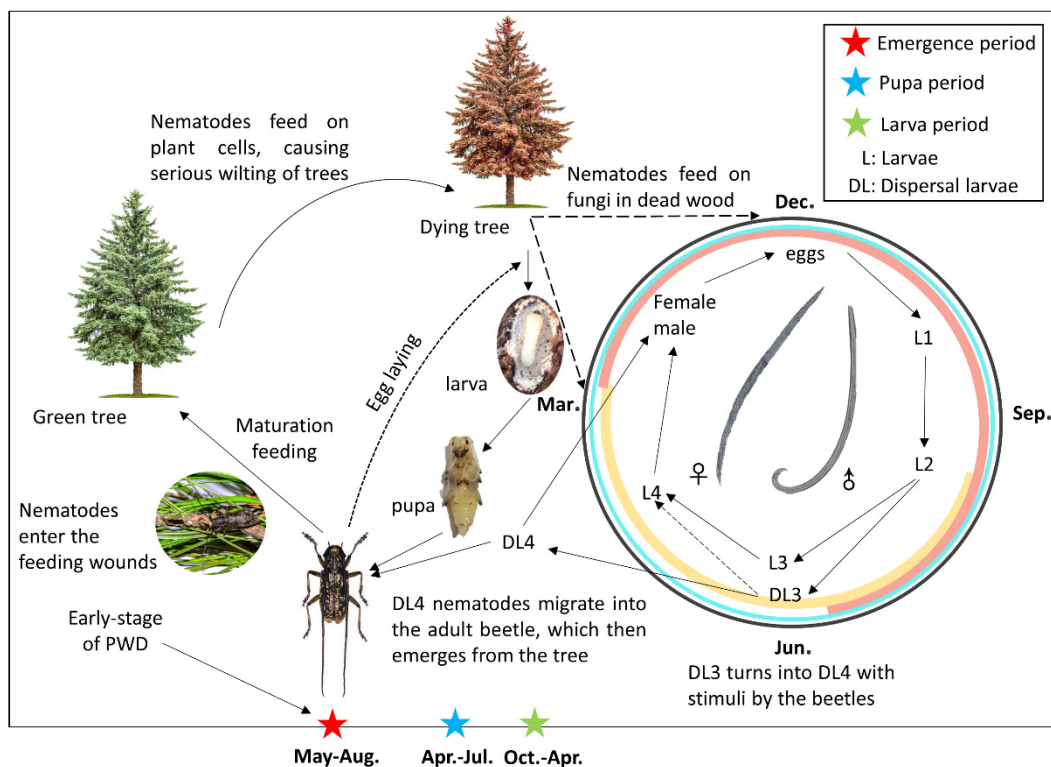


Figure 2. The process of pine wood nematode infection.

Since the first detection of PWD in China in 1982, it has been observed in many provinces across the country, with a northward trend spreading to North China (Figure 3) [9]. At the same time, *Monochamus saltuarius* (Gebler, 1830) has been confirmed as a new insect vector of PWD in North China (Figure 4) [9,10]. PWD has caused severe damage to *Pinus massoniana*, *P. tabulaeformis*, *P. koraiensis*, and other pine tree species in the process of spreading northward (Figure 3) [5,11]. PWD has become one of the most devastating

diseases to pine forests in China and has resulted in devastating disasters and tremendous economic losses to Chinese pine forests [12]. Furthermore, global warming also shows a great impact on forest ecosystems [13,14]. Due to the higher temperature, there would be an exponential increase in the population density of forest pests when the range of suitable habitats becomes wider [15].

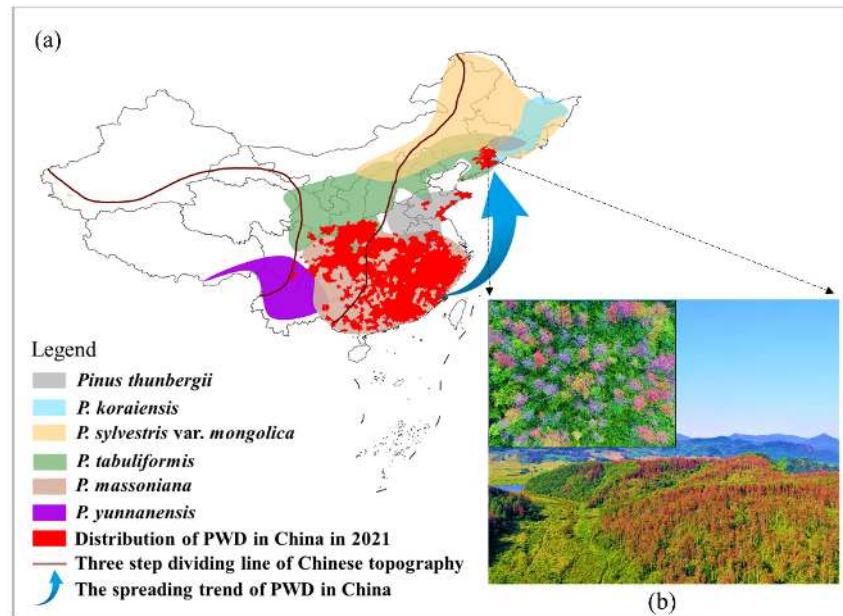


Figure 3. (a) Distribution of pine wilt disease (PWD) and major pine species susceptible to PWD in China. (b) PWD has caused a large area of death of pine forests in northern China.

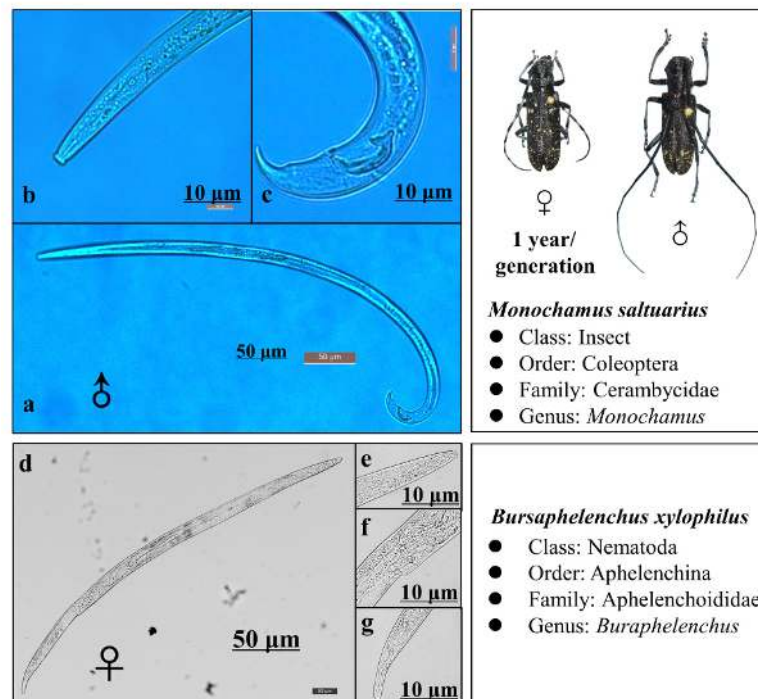


Figure 4. The morphology of *Monochamus saltuarius* and the pine wood nematode (PWN). (a) Male adult PWN; (b) head of the male PWN; (c) tail of the male PWN; (d) female adult; (e) head of the female PWN; (f) vulva of the female PWN; (g) tail of the female PWN.

To control and monitor PWD effectively, detecting the early infected pine trees by PWD is of great significance. However, it is an arduous assignment to achieve the goal of early monitoring of PWD because it only takes five weeks for pine trees to develop from the early stage of PWD infection to the late stage [16]. Currently, the main management practice to control PWD is to remove the dead trees infected by PWD through felling and burning [11,17]. To achieve the goal of early detection of PWD, a rapid and effective approach for monitoring pine forests is urgently needed. Another obstacle in the countermeasures of PWD is that the pine forest community is very large, which makes traditional ground investigations impractical.

To solve these problems, remote sensing (RS), as a potential detection method, is employed to monitor PWD. By reducing the space and time constraints, RS technology becomes more and more suitable for large-scale applications.

Hyperspectral remote sensing (HRS) features narrow bandwidths and can express both spatial and spectral information. HRS can capture continuous spectral data of targets; thus, it can be applied to detect minor changes in the spectral features of pine tree needles at the early stage of PWD infection during the process of discoloration (which is difficult to detect with the naked eye). Kim et al. [17] investigated the hyperspectral analysis of PWD, finding that within two months after PWN inoculation, the reflectance of red and mid-infrared wavelengths changed in most infected pine trees. Iordache et al. [18] collected unmanned aerial vehicle (UAV)-based hyperspectral images and applied random forest (RF) algorithms to detect PWD, achieving good results in distinguishing the healthy, PWD-infected, and suspicious pine trees. In another study, Yu et al. [11] combined ground hyperspectral data and UAV-based hyperspectral images, and found that the hyperspectral data performed well in discriminating the early infected pine trees by PWD using red edge parameters. These results demonstrate that HRS has great potential in monitoring PWD. However, the above studies employed traditional machine learning methods, which cannot directly recognize the spatial and spectral information from the images [19,20]. The three-dimensional data need to be flattened into one-dimensional vector data when a traditional machine learning algorithm is used on the whole image.

Due to the limitation of traditional machine learning models, the employment of deep learning algorithms in hyperspectral imagery (HI) classification has been attracting increasingly more attention, which provides a feasible solution for PWD detection. Deep learning algorithms can directly and effectively extract the information of deep features from the raw imagery data with an end-to-end mode [21]. Additionally, it can better explain the complicated architecture of high-dimensional data and obtain better accuracies through multi-layer neural network operations [22]. Over the last few years, deep learning has accomplished good performance in the field of computer vision and image processing, and has been widely applied in image classification and object detection [21,23,24]. At present, deep belief network (DBN) [25], stacked autoencoder (SAE) [26], convolutional neural network (CNN) [27], and other models have been applied in HI classification, and CNN is significantly superior to the other models in classification and target detection tasks [28–30].

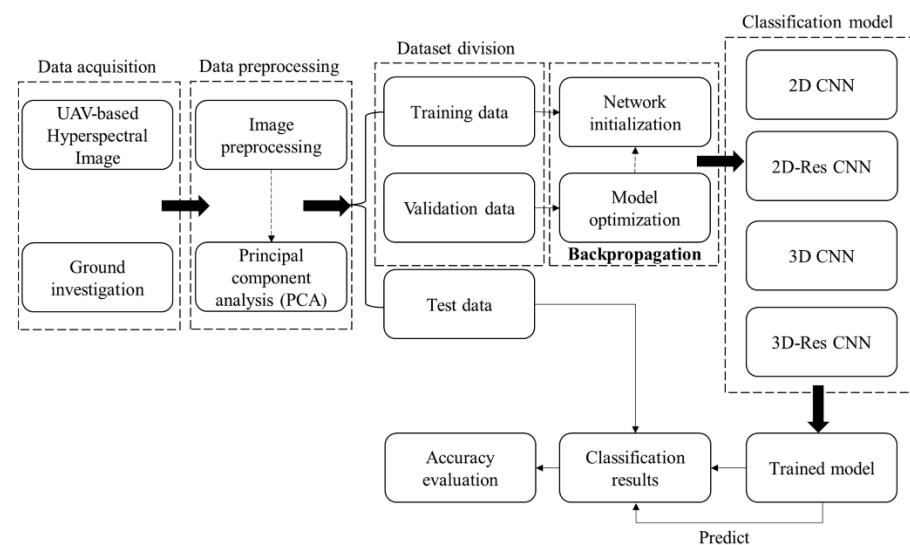
Consequently, the CNN model has been widely employed in PWD studies in recent years. In a study, two advanced object detection models, namely You Only Look Once version 3 (YOLOv3) and Faster Region-based Convolutional Neural Network (Faster R-CNN), were used in early diagnosis of PWD infection, obtaining good results and proposing an effective and rapid approach for the early diagnosis of PWD [19]. In another research, Yu et al. [20] employed Faster R-CNN and YOLOv4 to identify early infected pine trees by PWD, revealing that early detection of PWD can be optimized by regarding broad-leaved trees. Qin et al. [31] proposed a new framework, namely spatial-context-attention network (SCANet), to recognize PWD-infected pine trees using UAV images. The study obtained an overall accuracy (OA) of 79% and provided a valuable method to monitor and manage PWD. Tao et al. [32] applied two CNN models (i.e., AlexNet and GoogLeNet) and a traditional template matching (TM) approach to predict the distribution of dead pine

trees caused by PWD, revealing that the detection accuracy of CNN-based approaches was better than that of the traditional TM method. The above studies are all based on two-dimensional CNN (2D-CNN). Here, 2D-CNN [27] can obtain spatial information from the original raw images, whereas it cannot effectively extract spectral information. When 2D-CNN is applied to HI classification, it is necessary to operate 2-D convolution on the original data of all bands; the convolution operation would be very complicated because each band requires a group of convolution kernels to be trained. Different from the images with RGB bands, the input hyperspectral data in the network usually harbor hundreds of spectral dimensions, which requires numerous convolution kernels. This will cause over-fitting of the model, greatly increasing the computational cost.

To solve this difficulty, three-dimensional CNN (3D-CNN) is thus introduced to HI classification [33–35]. Here, 3D-CNN uses 3-D convolution to work simultaneously in three dimensions to directly extract the spectral and spatial information from the hyperspectral images. The 3-D convolution kernel is capable of extracting 3-D information, of which two represent spatial dimensions and the other one represents the spectral dimension. The HRS image is a 3-D cube, thus 3D-CNN can directly extract spatial and spectral data at the same time. These advantages enable 3D-CNN to serve as a more suitable model for HI classification. For example, Mäyrä et al. [21] collected hyperspectral and LiDAR data (LiDAR data can obtain canopy height model, which was used to match ground reference data to aerial imagery), and employed the 3D-CNN model for individual tree species classification from hyperspectral data, showing that 3D-CNNs were efficient in distinguishing coniferous species from each other, and at the same time showed high accuracy in classifying aspen. In another study, Zhang et al. [24] used hyperspectral images and proposed a 3D-1D convolutional neural network model for tree species classification, turning the captured high-level semantic concept (a joint spatial spectral feature representation) into a one-dimensional feature as a new input to learn a more abstract level of expression, and realized large area, high precision, high speed multi-tree species classification. In addition, the use of residual learning in the CNN model can optimize the performance of the model by solving the degradation problem of the network [36,37]. Residual learning can also be used in 3D-CNN. For example, Zhong et al. [38] designed an end-to-end spectral spatial residual network (SSRN), which selected 3-D cubes with a size of  $7 \times 7 \times 200$  as input data and did not require feature engineering for HI classification. In SSRN, spectral and spatial features were extracted by constructing spectral and spatial residual blocks, which further improved the recognition accuracy. Lu et al. [39] proposed a new 3-D channel and spatial attention-based multi-scale spatial spectral residual network (CSMS-SSRN). CSMS-SSRN used a three-layer parallel residual network structure to constantly learn spatial and spectral features from their respective residual blocks by using different 3-D convolution kernels, and then superimposed the extracted multi-scale features and input them into the 3-D attention module. The expressiveness of image features was enhanced from two aspects of the channel and spatial domain, enhancing the performance of the classification model.

Hyperspectral images and 3D-CNN models have also been employed in the forestry field, including tree species classification [21,24,40]. The principles for classifying PWD-infected pine trees at different stages are consistent with those of tree species classification. Therefore, 3D-CNN has the potential to be an ideal and feasible technology to precisely monitor PWD, which has not been explored in previous PWD research. Inspired by the aforementioned studies, the main objective of this study was to explore the ability to use 3D-CNN and residual blocks to identify pine trees at different stages of PWD infection.

The remainder of this paper is structured as follows: (1) construct 2D-CNN and 3D-CNN models to accurately detect PWD-infected pine trees; (2) compare the performance of 2D-CNN and 3D-CNN models for identifying pine trees at different stages of PWD infection; (3) explore the potential of adding the residual blocks to 2D-CNN and 3D-CNN models for an improvement in the accuracy; and (4) explore the impact of reducing training samples on model accuracies. The overall workflow of the study is shown in Figure 5.

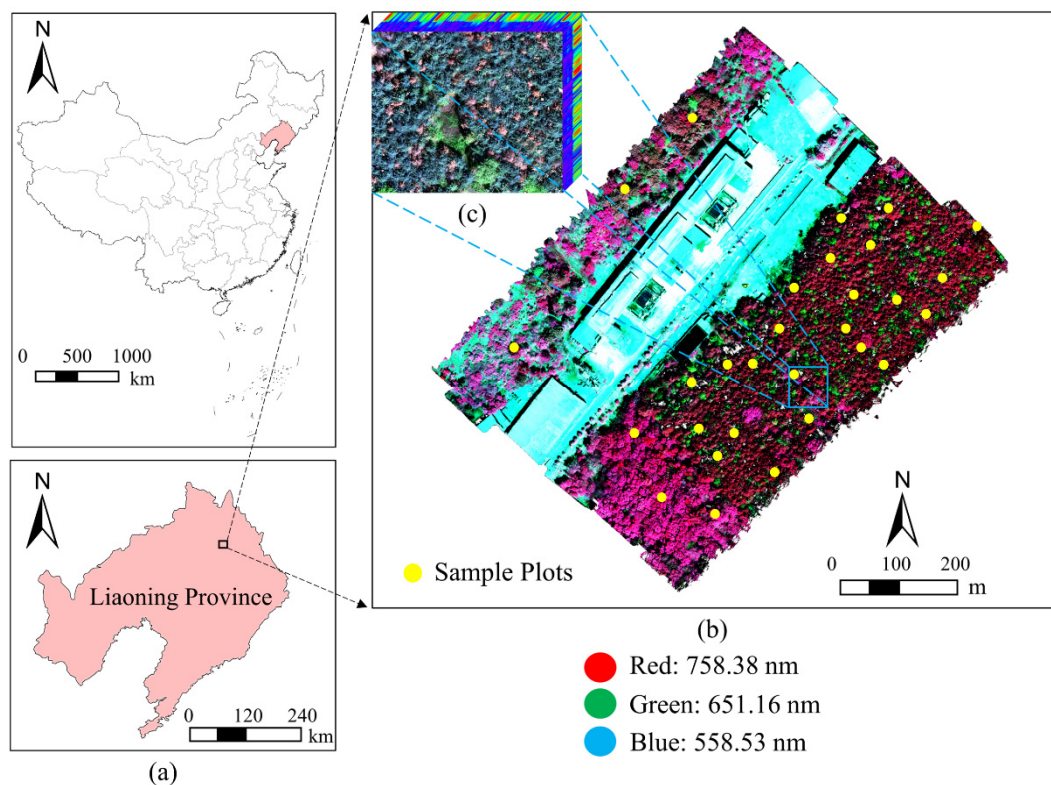


**Figure 5.** The overall workflow of the study.

## 2. Materials and Methods

### 2.1. Study Area and Ground Survey

The study area is located in Dongzhou District of Fushun City ( $124^{\circ}12'36''$ – $124^{\circ}13'48''$  E,  $41^{\circ}56'53''$ – $41^{\circ}57'46''$  N; Figure 6), Liaoning Province, China. This area has a continental monsoon climate and is located in the middle temperate zone. The mean annual air temperature of this area is approximately  $5$ – $7$  °C, and the average annual precipitation is  $760$ – $790$  mm. The plantation forests in the study area are dominated by *Pinus koraiensis*; broad-leaved tree species mainly include *Quercus acutissima* and *Q. mongolica*.



**Figure 6.** Study area. (a) Location of the study area; (b) the hyperspectral image (in false color composition) and location of sample plots; (c) the hyperspectral cube of the corresponding area from the hyperspectral image.

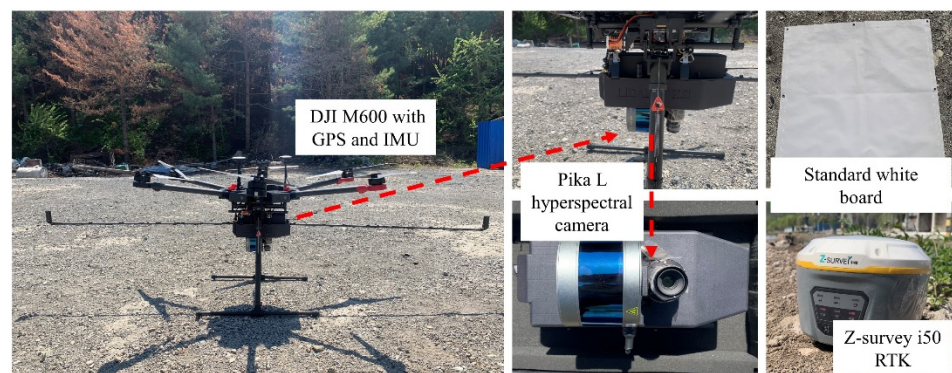
A ground survey was implemented from 8–14 May 2021. Twenty-eight sample plots (15 m × 15 m) were established in the study area (Figure 6). In each plot, tree species, the color of needles, tree vigor, and the PWD infection stage of each tree were recorded. The location of sampled trees was measured using a hand-held differential global positioning system (DGPS, South Surveying & Mapping Technology Co., Ltd., Guangzhou, China, Version S760) with an accuracy of sub-meter. A total of 1152 trees were measured. At the same time, 157 pine trees (42 trees with discolored needles and 115 trees showing no discoloration) were randomly sampled to confirm whether each tree carried the PWN or not by morphological and molecular identification [41,42]. The results showed that trees with discolored needles all carried the PWN, and more than 80% (93 out of 115) of trees showing no discoloration were verified to be infested by the PWN. Therefore, we confirmed that the study area was indeed infected by PWD.

## 2.2. Airborne Data Collection and Preprocessing

A DJI Matrice 300 drone (DJI, Shenzhen, China) carrying a Pika L hyperspectral sensor (to acquire the HI data; Resonon, USA) and a LR1601-IRIS LiDAR system (to collect the LiDAR data; IRIS Inc., Beijing, China) were used in this study. The main parameters of the hyperspectral sensor are shown in Table 1. The LiDAR system is uncalibrated, and the laser wavelength, pulse repetition frequency, and returns per pulse are 905 nm, 5–20 Hz, and 2, respectively. An inertial measurement unit (IMU) and a global position system (GPS) were mounted on the UAV, the horizontal and vertical position errors of which were 2.0 and 5.0 m, respectively. In addition, Z-survey i50 RTK (Shanghai Huace Navigation Technology Ltd., Shanghai, China) was applied to improve the POS (position and orientation system) accuracy (Figure 7). The whole UAV-based system is displayed in Figure 7.

**Table 1.** Main parameters of the Pika L hyperspectral camera.

Parameters	Values	Parameters	Values
Field of view	17.6°	Spectral resolution	2.10 nm
Focal length	17 mm	Spectral channels	281
Wavelength range	400–1000 nm	Sampling interval	1.07 nm



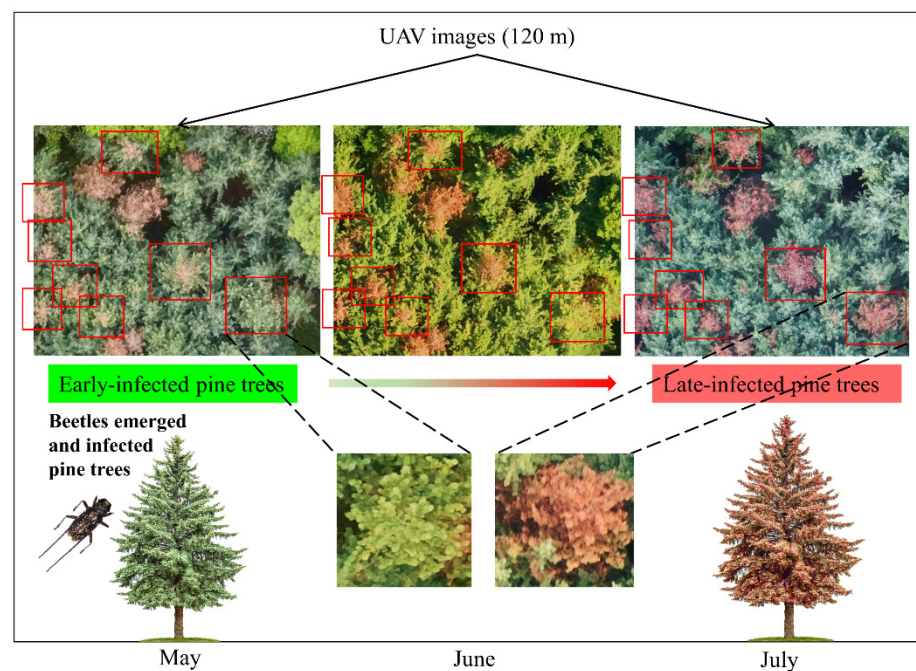
**Figure 7.** The UAV-based hyperspectral and LiDAR system.

UAV-based HI and LiDAR data acquisition was conducted from 11:40–12:20, on 11 May 2021. The RGB images collection were carried out from 12:00–12:20, on 11 May, 10 June, and 12 July 2021. The weather was sunny and windless during the flight. The flight height was set to 120 m, the overlap of the front and side was set to 60%, and the flight speed was 3 m/s. A standard whiteboard was set up in the flying area. The HI consisted of 281 spectral channels ranging from 400–1000 nm. The radiometric calibration and reflectance correction were performed with SpectronPro software (Resonon, Bozeman, MT, USA), using a 3-m<sup>2</sup> carpet as the reference. Geometric correction of images was conducted by applying six ground control points, the location of which was measured

using a DGPS device with an accuracy of sub-meter. The spatial resolution of HI was 0.44 m. The synchronously collected LiDAR data provided accurate DEM data for the hyperspectral data preprocessing. LiDAR data were georeferenced in the Universal Transverse Mercator 51N, and WGS1984 datum was used as the coordinate system. We classified ground, above-ground, and understory points from the raw LiDAR data for hyperspectral data preprocessing using the LiDAR360 software (version 4.1, GreenValley Inc., Beijing, China).

### 2.3. Division of PWD Infection Stage and Data Labeling

In this study, the early infected pine trees by PWD were confirmed by multi-temporal observations, that is, using images collected in July and June to distinguish early infected pine trees in the images collected in May when most insect vectors emerge and infect pine trees (Figure 8). The HI collected on May 11 was taken as input to the models (Figure 6b). Pine trees with a red crown were defined as late infected pine trees. Additionally, we also identified the broad-leaved trees in the images based on the ground survey. According to the above criteria, we labeled the trees in the hyperspectral images with these three categories.



**Figure 8.** UAV images of pine trees from the early (May) to the late stage (July) of PWD infection.

### 2.4. Model Construction

#### 2.4.1. 3D-CNN

In the traditional 2D-CNN, the convolution operation is only used to capture the spatial features of 2-D feature map. However, both spatial and temporal information need to be captured in processing 3-D data (e.g., video data). To overcome this obstacle, 3D-CNN was proposed; the 3D convolution operation is performed on the 3-D feature map to measure the spatiotemporal characteristics of the 3-D data [43]. In HI classification, it is essential to retain abundant spectral information of targets. The 2-D convolution operation is to convolute the input data in the spatial dimension, and the output is the 2-D feature map regardless of whether the input is 2-D or 3-D data, resulting in the loss of spectral information of the hyperspectral data. For the 3-D convolution operation, the input data can be convoluted in both spatial and spectral dimensions. Moreover, the output of the 3-D convolution operation is 3-D cubes, which can retain the spectral information of hyperspectral images.



HI includes both spatial and spectral information. In this study, spatial and spectral information from the images were integrated to build a CNN classification model. Furthermore, the 3D-CNN used for HI classification was not based on image-level features but pixel-level features, and the input data were a set of spatial-spectral neighboring cubes around pixels rather than the whole image [24].

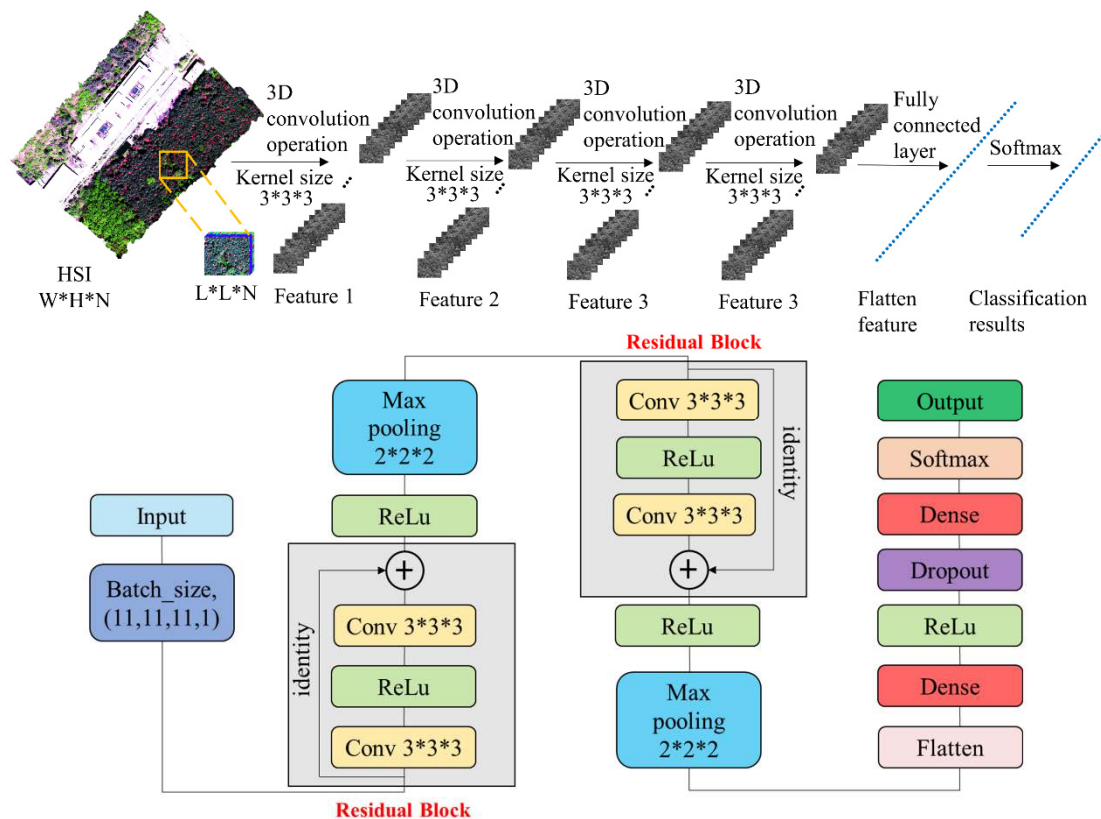
Different from the image-level classification task, the space size of the input information applied in RS image classification is smaller, and that of the feature map is further decreased after convolution. Generally, the convolution kernel with smaller space can be used to evade immoderate loss of the input data. Based on previous studies [24,35,40,44], the 3D-CNN model with a convolution kernel size of  $3 \times 3$  exhibited the best performance in HI classification, and the 3D-CNN model with a  $3 \times 3 \times 3$  convolution kernel achieved good results in spatiotemporal feature learning [44]. Therefore, in this study, the kernel size was set to  $3 \times 3 \times 3$ . In addition, the window size was set to  $11 \times 11$ , and according to Zhang et al. [24], the stride of the convolution layer was 1. The kernel size of the pooling layer was  $2 \times 2 \times 2$ , and the stride of the pooling layer was set to 2.

#### 2.4.2. Construction of the 3D-Res CNN Classification Model

The 3D-Res CNN model consists of four convolution layers, two pooling layers, and two residual blocks. Figure 9 shows the model architecture, and the details are described as follows:

- (1) Data collection from HI. Here, 3D-CNN can use raw data without dimensionality reduction or feature filtering, but the data collected in this study were enormous and contained a lot of redundant information. Therefore, to make our model more rapid and lightweight, the dimensionality of the raw data was reduced through a principal component analysis (PCA), and 11 principal components (PCs) were extracted for further analyses. The objective pixel was set as the center, and the spatial-spectral cubes with a size of  $L \times L \times N$  as well as their category information were extracted. Here,  $L \times L$  stands for the space size, and  $N$  is the number of bands in the image.
- (2) Feature extraction after 3-D convolution operation. The model includes four convolution layers and two fully connected layers. The spatial-spectral cubes ( $L \times L \times N$ ) obtained from the previous step were used as input of the model. The first convolutional layer (Conv1) contains 32 convolution kernels with a size of  $3 \times 3 \times 3$ , a step size of  $1 \times 1 \times 1$ , and a padding of 1. The 32 output 3-D cubes (cubes-Conv1) had a size of  $(L - \text{kernel size} + 2 \times \text{padding}) / \text{stride} + 1$ . The 32 cubes-Conv1 were input to the second convolution layer (Conv2), and 32 output 3-D cubes (cubes-Conv2) were obtained. The add operation was performed on the output of the input and cubes-Conv2, and the activation function and pooling layer ( $k = 2 \times 2 \times 2$ , stride =  $2 \times 2 \times 2$ ) were applied for down-sampling. As a result, the length, width, and height of these cubes were reduced to half of the original values; the 32 output 3-D cubes were denoted as cubes-Pool1. After two more rounds of convolution operation, cubes-Conv4 were obtained; the add operation was performed to cubes-Pool1 and cubes-Conv4. After applying the activation function and the pooling layer, the length, width, and height were again reduced to half of the original values, and the 32 output cubes were denoted as cubes-Pool2.
- (3) Residual blocks. The residual structure consists of two convolution layers. The data were input to the first convolution layer (Conv1R), and the rectified linear unit (ReLU) activation function was used. The output of Conv1R was input to the second convolution layer (Conv2R), and the ReLU activation function was used to obtain the output of Conv2R. The add operation was performed on the output of Conv1R and Conv2R, and the ReLU activation function was then employed to obtain the output of the whole residual structure.
- (4) Fully connected layers. The features of cubes-Pool2 were flattened, and by applying the fully connected layers, the cubes-Pool2 were transformed into feature vectors with a size of  $1 \times 128$ .

- (5) Logistic regression. A logistic regression classifier was added after the fully connected layers. Softmax was applied for multiple classification. After flattening the features of the input data, the probability of these features can be attached to each category of trees.



**Figure 9.** The architecture of the 3D-Res CNN model, which includes four convolution layers, two max pooling layers, and two residual blocks. Conv stands for convolutional layer, ReLu stand for the rectified linear unit.

The parameters of the model were initialized randomly and optimized by backpropagation to minimize network loss and complete model training. Before setting the weight update rule, a suitable loss function is required. This study adopted the mini-batch update strategy, which is suitable for processing large datasets. The calculation of the loss function is based on the mini-batch input, and the formula is as follows:

$$L_{CE}(\hat{y}, y) = - \sum_{i=1}^{n\_classes} y_i * \log(\hat{y}_i) \quad (1)$$

where  $y$  is the true label and  $\hat{y}$  is the predicted label.

The first fully connected layer and the convolution layers in the network use a linear correction unit (i.e., ReLU) as the activation function, where the formula is:  $f(x) = \max(0, x)$  [27]. ReLU is a widely used unsaturated activation function. In terms of gradient descent and training time, the efficiency of ReLU is higher than other saturated activation functions. The last fully connected layer uses the softmax activation function, and the sum of the probability values of all neuron activation is 1.

The network adds dropout to the two fully connected layers. According to the probability, the output of the neuron was set to 0 to limit the interaction of hidden units, enable the network to learn more robust features, and reduce the impact of noise and overfitting, ultimately enhancing the performance of the neural network [45].

After the network was constructed, the training process was configured to renew the parameters of the 3-D convolution kernel through the backpropagation loss function gradient. The batch size was 64, and the Adam optimizer was used to complete the training process. Adam introduces momentum and exponential weighted average strategies, which can adaptively adjust the learning rate and converge the model faster. Among them, the hyperparameters were set as follows: learning rate = 0.001, beta\_1 = 0.9, beta\_2 = 0.999, epsilon =  $1 \times 10^{-8}$ , and decay = 0.0. The model was trained for 300 epochs.

Table 2 shows the architecture of the 3D-Res CNN model.

**Table 2.** Structure of the 3D-Res CNN model.

Layer (Type)	Output Shape (Height, Width, Depth, Numbers of Feature Map)	Parameter Number	Connected to
input_1 (InputLayer)	(11, 11, 11,1)	0	
conv3d (conv3D)	(11, 11, 11, 32)	896	input_1
conv3d_1 (conv3D)	(11, 11, 11, 32)	27680	conv3d
add (Add)	(11, 11, 11, 32)	0	conv3d input_1
re_lu (ReLU)	(11, 11, 11, 32)	0	add
max_pooling3d (MaxPooling3D)	(5, 5, 5, 32)	0	re_lu
conv3d_2 (conv3D)	(5, 5, 5, 32)	27680	max_pooling3d
conv3d_3 (conv3D)	(5, 5, 5, 32)	27680	conv3d_2
add_1 (Add)	(5, 5, 5, 32)	0	conv3d_3max_pooling3d
re_lu_1 (ReLU)	(5, 5, 5, 32)	0	add_1
max_pooling3d_1 (MaxPooling3D)	(2, 2, 2, 32)	0	re_lu_1
flatten (Flatten)	(256)	0	max_pooling3d_1
dense (Dense)	(128)	32896	flatten
dropout (Dropout)	(128)	0	dense
dense_1 (Dense)	(3)	387	dropout

### 2.5. Comparison between the 3D-Res CNN and Other Models

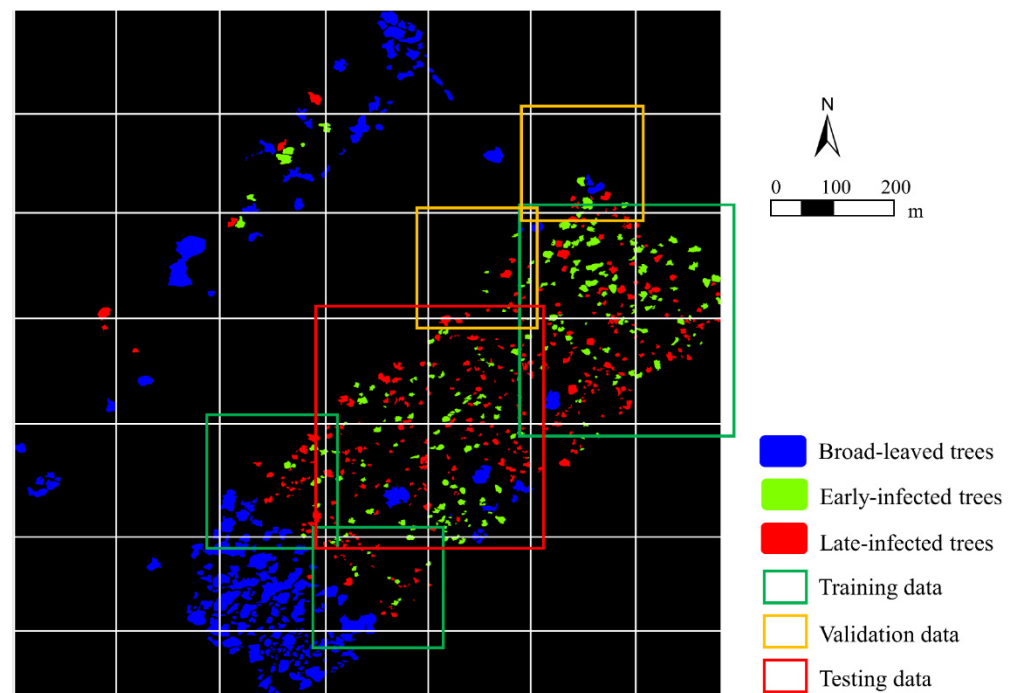
To test the performance of the 3D-Res CNN model in identifying PWD-infected pine trees based on hyperspectral data, the 3D-CNN, 2D-CNN, and 2D-Res CNN models were used for comparative analysis.

For 2D-CNN, the PCA generated 11 PCs from 150 bands of the original hyperspectral data, and  $11 \times 11 \times 11$  data were extracted as the original features. The network included four convolution layers, two pooling layers, and two fully connected layers. The size of the convolution kernel was  $3 \times 3$ , and each layer had 32 convolution kernels. The structure of 3D-Res CNN was similar to that of 2D-Res CNN. Although 3D-Res CNN shared the same parameters as 2D-CNN, it had five convolution layers since adding residuals requires an additional convolutional layer.

The 2D-CNN, 2D-Res CNN, 3D-CNN, and 3D-Res CNN modeling were implemented in Python using the Tensorflow framework. The operation platform incorporated Intel(R) Xeon (R) CPU E5-2620 v4 @2.10 GHZ, and NVIDIA GeForce RTX 2080Ti GPUs.

### 2.6. Dataset Division and Evaluation Metrics

We divided the whole hyperspectral image into 49 small pieces (Figure 10), and stitched the resulting maps together after the analyses. At the same time, we selected 6 pieces as training data, 2 pieces as validation data, and 4 pieces as testing data (Figure 10). Each tree category was divided into training, validation, and testing datasets at a ratio of 5:1:4. The specific pixel number for each category is shown in Table 3.



**Figure 10.** Training, validation, and testing samples of each tree category with the true labels.

**Table 3.** Pixels of training, validation, and testing for each tree category.

Categories	Sample's Pixel Number			
	Training	Validation	Testing	Total
Early infected pine trees	163,628	32,726	130,902	327,256
Late infected pine trees	242,107	48,421	193,685	484,213
Broad-leaved trees	100,163	20,033	80,130	200,326
Total	505,898	101,180	404,717	1,011,795

The classification accuracy was assessed by calculating the producer accuracy (PA), average accuracy (AA), overall accuracy (OA), and the Kappa coefficient value [46]. The formulas are as follows:

$$PA = \text{correct classification pixel number of each class} / \text{total pixel number of each class} \quad (2)$$

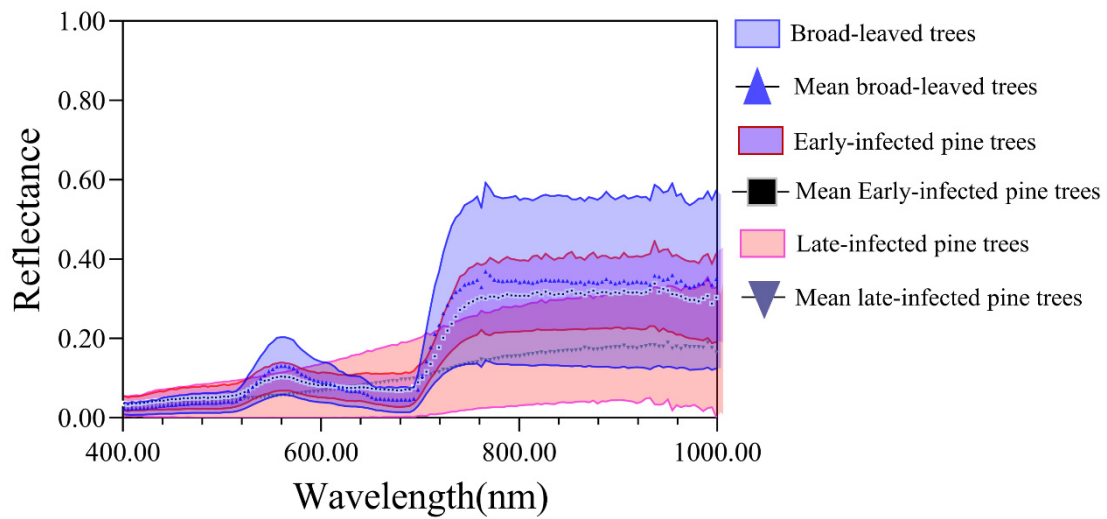
$$Kappa = (OA - eAccuracy) / (1 - eAccuracy) \quad (3)$$

$$eAccuracy = (\sum_{i=1}^k Vp \times Vm) / S^2 \quad (4)$$

where OA is overall accuracy, k is the number of categories,  $Vp$  is the predicted value,  $Vm$  is the measured value, and  $S$  is the sample number.

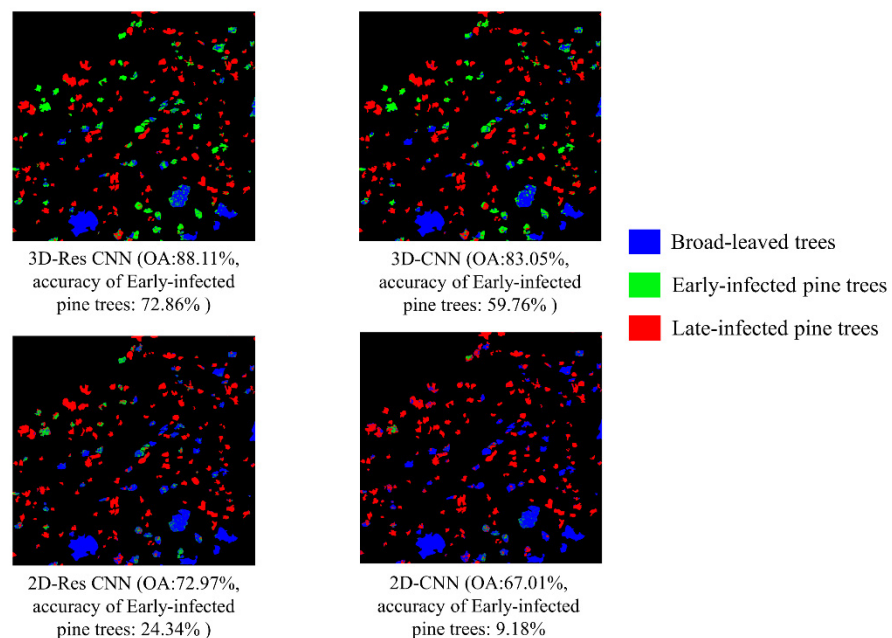
### 3. Results

The reflectance curves of broad-leaved trees, early infected pine trees, and late infected pine trees within 400–1000 nm are depicted in Figure 11. Of the broad-leaved trees and two stages of infected pines, the difference in the spectral reflectance was most obvious in the green peak (520–580 nm), red edge (660–780 nm), and NIR (720–900 nm). Furthermore, the models we used still incorrectly classified early infected pine trees into broad-leaved trees because the spectrum of early infected pine trees is similar to that of broad-leaved trees (Figure 11).



**Figure 11.** The reflectance curve of broad-leaved trees, early infected pine trees, and late infected pine trees.

Further, 2D-CNN did not achieve satisfactory results in the classification task (OA: 67.01%; Figure 12 and Table 4). Moreover, it barely recognized the early infected pine trees in the hyperspectral image with relatively low resolution, which could be disturbed by the similar color, contour, or texture of the crown as those of broad-leaved trees. Additionally, the accuracies were improved by adding the residual block in the CNN model. The OA was improved from 67.01% to 72.97%, and the accuracy for identifying the early infected pine trees was increased from 9.18% to 24.34% when applying the 2D-Res CNN model (Figure 12 and Table 4).



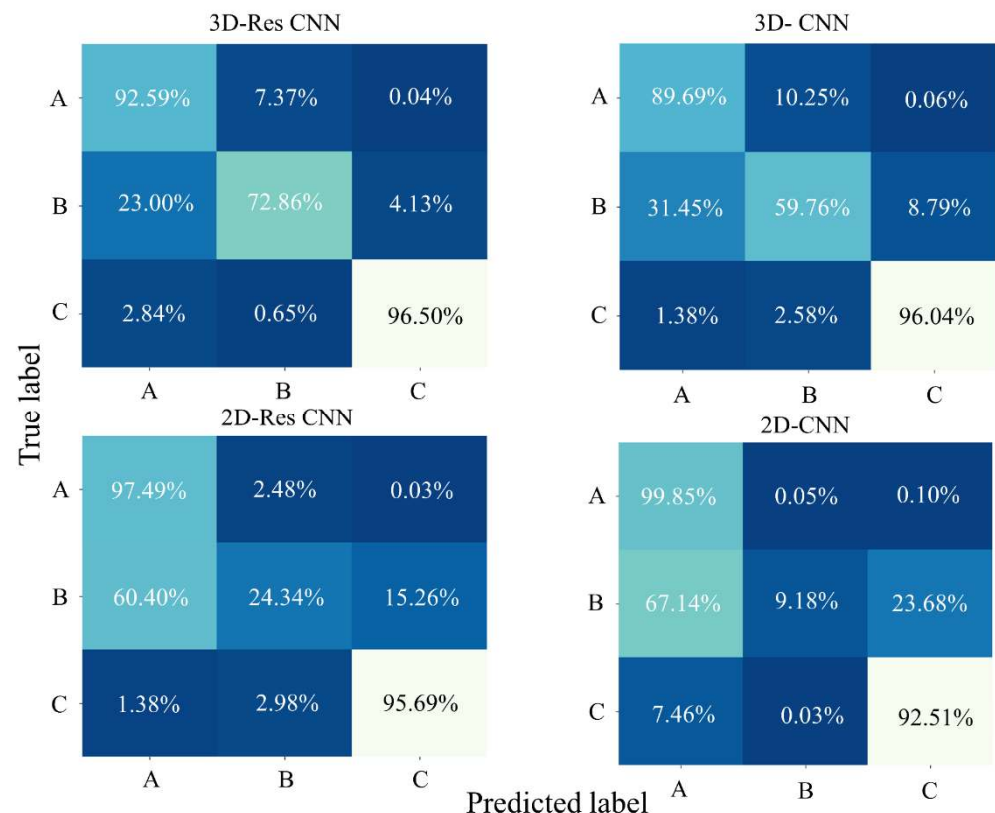
**Figure 12.** The classification results of three tree categories in the study area using the four models.

**Table 4.** Classification accuracy of three classes using different approaches.

Model	2D-CNN	2D-Res CNN	3D-CNN	3D-Res CNN
OA (%)	67.01	72.97	83.05	88.11
AA (%)	67.18	72.51	81.83	87.32
Kappa $\times$ 100	49.44	58.25	73.37	81.29
Early infected pine trees (PA%)	9.18	24.34	59.76	72.86
Late infected pine trees (PA%)	92.51	95.69	96.04	96.51
Broad-leaved trees (PA%)	99.85	97.49	89.69	92.58
Trainable parameters	47,843	47,843	117,219	117,219
Trainable time (minute)	34 min	35 min	100 min	115 min
Prediction time (second)	14.3 s	14.8 s	20.1 s	20.9 s

The performance of 3D-CNN was better than that of 2D-CNN in distinguishing the three tree categories. Specifically, the OA of 3D-CNN was 83.05%, and the accuracy of identifying early infected pine trees was 59.76% (Figure 12 and Table 4). When the residual block was applied to the 3D-CNN model, 3D-Res CNN obtained better results, with an OA of 88.11% and an accuracy of 72.86% for identifying early infected pine trees.

The classification performance of all the models is shown in Figure 13. In summary, all the models successfully identified the broad-leaved trees and late infected pine trees. Furthermore, the results demonstrated that the classification accuracy can be greatly enhanced by adding the residual block, and the training parameters and overall training time were almost unchanged compared to those of models without the residual block (Table 4). More importantly, the OA was dramatically improved when switching 2D-CNN to 3D-CNN, and the training parameters and training time were greatly increased as well, indicating that the increase of the training time is a valuable trade-off.

**Figure 13.** Confusion matrices for the three tree categories using different models, where A, B, and C respectively represent broad-leaved trees, early infected, and late infected pine trees.

In our study, the ratio of the training, validation, and testing samples was 5:1:4; the relatively sufficient training samples enabled us to achieve good results. Hyperspectral data are enormous and complex, and in the training process of CNN, a great quantity of samples is needed to better grasp the valuable features of the classification model. The CNN model may not achieve satisfactory accuracy without enough training samples. However, in actual forestry management, especially in large-scale applications, it is difficult to collect sufficient training samples, which consumes manpower and material resources. Thus, it is of great importance to ensure good accuracies of the model even with a smaller number of training samples in practical forestry applications.

To verify whether the proposed 3D-Res CNN model can maintain a relatively good accuracy when given a smaller size of training samples, we reduced the training samples to 40%, 30%, 20%, and 10% of the total sample size, and calculated its respective accuracies. The number of the testing samples remained unchanged, and the remaining samples were added to the validation samples.

Figure 14 shows the classification accuracy and time consumption under different training dataset conditions. The results indicated that the classification accuracies of the 3D-Res CNN model slightly decreased when the training sample size was reduced from 50% to 20%. When the training sample size was 10%, the accuracy for identifying early infected pine trees was abnormal due to the smaller size of the training dataset. The 3D-Res CNN model performed almost as well as or even better than the 2D-CNN and 2D-Res CNN models when the training sample size was reduced to 20%. When the training sample size was set to 20%, the OA and the Kappa value of the 3D-Res CNN model were 81.06% and 70.29%, respectively, and the accuracy for identifying early infected pine trees was 51.97%, which were still better than those of 2D-CNN. In general, the accuracies of the 3D-Res CNN model decreased with the reduction of the training sample size, but the accuracies still meet the requirement of forestry applications in a large area. Additionally, the training time for the 3D-Res CNN model using a smaller training sample size was shorter than that using the full set of training samples (e.g., 22.61% less when using 30% training samples), which accelerated the training process of the classification task. In general, it is feasible for our 3D-Res CNN model to be employed in practical forestry applications using a smaller number of samples.

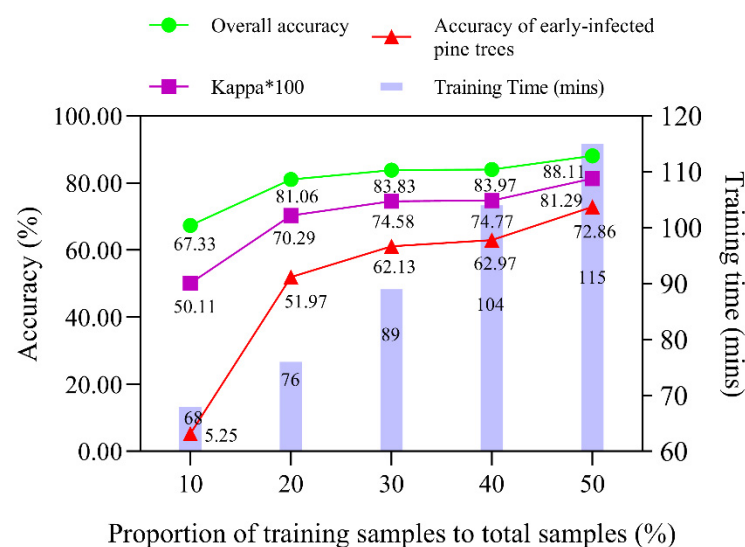


Figure 14. Classification performance of the 3D-Res CNN model using different training sample sizes.

#### 4. Discussion

##### 4.1. Comparison of Different Models and the Contribution of Residual Learning

In this study, 2D-CNN and 3D-CNN models were applied to identify the PWD-infected pine trees. The classification method based on spatial features (e.g., 2D-CNN)

exhibits some limitations in classifying hyperspectral data [47]. The dimensionality of the original hyperspectral image needs to be reduced prior to data processing, converting the hyperspectral image into an RGB-like image. On the one hand, if dimensionality reduction is not carried out, the number of parameters would be very large, which is prone to over-fitting. On the other hand, dimensionality reduction may destroy the spectral structure of hyperspectral images that contain hundreds of bands, resulting in a loss of spectral information and a waste of some specific properties of the HI data. Moreover, the spatial resolution of hyperspectral image is often inferior to that of the RGB image, thus it is difficult for 2D-CNN to accurately distinguish early infected pine trees from the crowns with close color, contour, or texture.

Different from 2D-CNN, which requires dimensionality reduction of the original image, 3D-CNN directly and simultaneously extracts spatial and spectral information from the original hyperspectral images. In this study, 3D-CNN models achieved better accuracies compared with the other models (Table 4 and Figure 12). Although the training parameters and training time were increased, the classification accuracy was also greatly improved. It is worth trading off 70 min of training time for more than a 20% increase in accuracy. The overall training time (115 min) of 3D-Res CNN can fully meet the requirement of practical forestry applications in a large area.

In our work, the model accuracy was greatly improved by adding the residual block. For 2D-CNN, after adding the residual block (i.e., 2D-Res CNN), the OA increased from 67.01% to 72.97%, and the accuracy for identifying early infected pine trees also increased by 15.16%. For the 3D-Res CNN model, both the OA (from 83.05% to 88.11%) and the accuracy for identifying early infected pine trees (from 59.76% to 72.86%) were greatly improved compared to those of 3D-CNN. Furthermore, the training time of the 3D-Res CNN model increased by only 15 min (15% of the training time of 3D-CNN), while that of 2D-Res CNN remained unchanged compared to 2D-CNN. This is because the degradation problem of the network is solved through residual learning in the residual CNN model, thus a better accuracy can be achieved [36]. Therefore, residual learning can indeed improve the classification accuracy of our model and only increased the relatively short training time.

#### 4.2. Early Monitoring of PWD

PWD has destroyed billions of pine trees in China, leading to countless ecological and economic losses [5,11]. Therefore, it is imperative to detect PWD at the early stage and take preventive measures as soon as possible. In recent years, “early monitoring” has been a hot topic in forest pest research [18,48–50]. Nevertheless, the precise definition of “early stage” is difficult to determine, especially in the PWD research. In this study, we determined the early infected pine trees by PWD by continuously observing the specific pine trees at equal intervals over a period of time. For one thing, in addition to the discoloration of pine tree crowns caused by PWD, phenology can also lead to the discoloration of pine trees, which will affect the judgment of “early stage”. For another thing, multitemporal observations are particularly time-consuming, as several months or even years were taken in some experiments [18,19].

Some scholars inoculated healthy pine trees with PWN and defined these trees to be at the early stage of PWD infection [17]. First, this approach is only suitable for small sample sizes and cannot be employed to actual large-scale forestry applications. Second, artificial injection of PWN is different from its infection mechanism in the natural environment (by vector insects). More importantly, it is difficult to carry out such an operation and the rate of inoculation cannot be guaranteed [51]. Thus, this method is not suitable for practical forestry applications.

In the actual control of forest pests, it is usually required to detect PWD at a single time point and take control measures at this very time, rather than long-term observations. Detecting PWD at a single time point has already met the requirement of actual forestry management. Therefore, a rapid and easy method should be presented to confirm the occurrence of PWD in the practical forestry application. On this basis, the UAV-based



RS images should be obtained at the optimal monitoring time of PWD infection (under investigation) and the stage of PWD infection should be preliminarily estimated through the color of tree crowns. In addition, a feasible attractant for PWN should be designed and applied to determine whether the pine trees carry PWD in the large-scale area. Combining these two processes, it is feasible to prevent and control PWD in large-scale forestry applications in a timely fashion.

#### 4.3. Existing Deficiencies and Future Prospects

In this work, we applied 3D CNN and residual blocks to construct a 3D-Res CNN and used it in the study of forest pest detection (PWD in this study, but it can be used for other forest disease and pest detection), which has not been studied in previous works. In our work, the proposed 3D-Res CNN is the best model in the detection of PWD. Compared with 2D CNN, it can directly extract spatial and spectral information from hyperspectral images at the same time, and make us more accurate in identifying PWD-infected pine trees. Additionally, using only 20% of the training samples, the OA and EIP accuracy of the 3D-Res CNN can still achieve 81.06% and 51.97%, which is superior to the state-of-the-art method in the early detection of PWD based on hyperspectral images [11,19,20,31,48]. However, in our study, when we used the proposed model, we performed PCA first instead of directly using the raw data (because the raw data is too enormous), which made our classification process less convenient. In addition, the enormous hyperspectral data have higher requirements on GPUs, and the training time is relatively long. Therefore, a lightweight and fast convergence 3D CNN classification model should be designed in the future. Furthermore, in this work, we divided the whole hyperspectral image into 49 small pieces, and different pieces were used for training, validation, and test purposes. Although each piece is different, and the input data of the model can be reduced by this method, they still belong to a single image on a single date, which will affect the generalization capacities of the models. In order to make our model more generalized, we will use multitemporal hyperspectral images for PWD detection in the next study.

Additionally, there are several effective methods to improve the performance of classification models, which can also be used for PWD and other forest damage monitoring. First, the layers of the CNN model can be increased, and more rounds of residual learning can be performed to optimize the accuracies of the model. He et al. [36] put forward a deep residual network (ResNet) with 152 layers, greatly reducing the error of the CNN model. Second, the split-transform-merge strategy can also be employed in processing enormous hyperspectral data, which would reduce the training time and computational cost. Szegedy et al. [52] introduced a residual structure, proposed Inception-Resnet-v1 and Inception-Resnet-v2, and modified the inception module to propose the Inception-v4 structure. Moreover, Inception used a split-transform-merge strategy: the input data were first divided into several parts, then different operations were separately performed, and finally the results were merged. In this way, the computational cost can be reduced while maintaining the expressive ability of the model [30]. Based on the split-transform-merge strategy of Inception, Xie et al. [53] designed a ResNeXt model, which is simpler and more efficient than Inception and ResNet.

In recent studies, Yin et al. [54] combined 3D CNN and a band grouping-based bidirectional long short-term memory (Bi-LSTM) network for HSI classification. In the network, the extracted spectral features were regarded as a procedure of processing sequence data, and the Bi-LSTM network acted as the spectral feature extractor to fully use the relationships between spectral bands. Their results showed that the proposed method performed better than the other HSI classification methods. In another study, Gong et al. [55] proposed a multiscale squeeze-and-excitation pyramid pooling network (MSPN), and used a hybrid 2D-3D-CNN MSPN framework (which can learn and fuse deeper hierarchical spatial-spectral features with fewer training samples). The results demonstrated that a 97.31% classification accuracy was obtained based on the proposed method using only 0.1% of the training samples in their work. These methods are lightweight and convenient,

which could also be applied to detect PWD and other forest diseases and pests. There are also some recent studies in the monitoring of PWD. For example, Zhang et al. [56] designed a spatiotemporal change detection method in a complex landscape, using deep learning algorithms to capture the spectral, temporal, and spatial characteristics of the target from the image, thereby reducing false detections in tree-scale PWD monitoring. In another study, in order to obtain the detailed shape and size of infected pines, high-performance deep learning models (e.g., fully convolutional networks for semantic segmentation) were applied to perform image segmentation to evaluate the disease's degree of damage, and achieved good results [57].

Additionally, although many of widely used deep learning-based HI classification methods have achieved good classification accuracy, these methods are often accompanied by a large number of parameters, a long training time, and a high-complexity algorithm. Therefore, it is often inconvenient to adjust the hyperparameters. These limitations lie in the theoretical research of algorithms and the high dimensionality of the HI data. Therefore, how to enhance the generalization ability of these methods and the robustness of the model needs to be further explored in future studies.

In this study, the classification task was performed based on a supervised classification method. With each sample labeled to its own corresponding category, this method constantly learns the corresponding features through deep neural networks, finally realizing the classification task. To estimate the accuracies of the classification model, we manually labeled each sample based on the field investigation results, which was time- and labor-consuming and resulted in a smaller sample size. To solve these problems, migration learning and data enhancement methods can be employed. For example, the generative adversarial network (GAN) [58] uses a generator and a discriminator, where the function of the generator is to produce the target output, and the function of the discriminator is to discriminate the true data in the output. During the training process, the generator that captures the data distribution and the discriminator that estimates the probability finally reach a dynamic balance through continuous confrontation: that is, the image generated by the generator is very close to the distribution of the real image. The GAN can also be used to enrich hyperspectral data: GAN learns a category in the hyperspectral image to generate new data that match the characteristics of this category, increasing the amount of data in this category and expanding the sample size [59]. In addition, the unsupervised classification method [60] can be used to construct the network using an end-to-end encoder-decoder approach. Unsupervised methods can solve the problem of deep learning models relying on a large number of learning samples. Therefore, in the future, unsupervised classification models can be considered in large-scale practical forestry applications, such as the control of diseases and pests, which will enable the forest managers to better grasp the distribution and spreading trend of pests and diseases in the forest.

Another potential tool to detect PWD is light detection and ranging (LiDAR). As an active remote sensing technology, LiDAR can penetrate the tree canopy and quickly obtain information about the vertical structure of the forest [61–65]. More importantly, LiDAR data have been widely used in forest health monitoring [21,24,61–65]. When we use HI data alone, we cannot accurately segment the canopy, and the shadows, understory, and overlapping canopies can easily cause spectral confusion. LiDAR can solve these problems by collecting the structural features of trees, and the metrics derived from LiDAR data can also be used for the detection of forest pests [48,66–68]. However, the combination and fusion of HI and LiDAR has not been well studied in PWD detection, which will be researched in our next study.

To sum up, in view of the above-mentioned advanced models (e.g., the split-transform-merge strategy, unsupervised classification methods, some novel lightweight model, and the use of Lidar data), we can also apply them to PWD and other forest damage monitoring. These methods can be employed in future research to achieve accurate, automatic, and early monitoring of diseases and pests in the forest.

## 5. Conclusions

In this paper, we applied 3D CNN and residual blocks to construct a 3D-Res CNN for early detection of PWD based on hyperspectral images. Additionally, we compared the classification accuracy of the 2D-CNN, 3D-CNN, 2D-Res CNN, and 3D-Res CNN models in identifying pine trees infected by PWD. The results demonstrated that the 3D-Res CNN model was most efficient in detecting the early infected pine trees by PWD. Using only 20% of the training samples, the OA and EIP accuracy of the 3D-Res CNN can still achieve 81.06% and 51.97%, which is superior to the state-of-the-art method in the early detection of PWD based on hyperspectral images. Further, 3D-Res CNN can simultaneously extract the spectral and spatial information from the hyperspectral images, and add a residual module to improve the recognition accuracy. Although its training time was longer, it can still meet the requirements of practical forestry applications in a large area.

More importantly, there are still many obstacles that need to be addressed in PWD detection (e.g., the lack of training data). Therefore, we expect to use more effective and lightweight technologies to carry out early monitoring of PWD on a larger scale in the future. This will enable forest managers to better determine the distribution and spreading trend of PWD, so as to take preventive measures as early as possible to reduce ecological and economic losses. When applying large-scale RS data in monitoring of large areas of PWD, it is crucial to develop such methods.

**Author Contributions:** Conceptualization, R.Y.; Formal analysis, R.Y.; Funding acquisition, Y.L.; Investigation, R.Y., H.L. and L.Y. (Liyuan Yang); Methodology, R.Y., Y.L., L.Y. (Linfeng Yu) and L.R.; Writing—original draft, R.Y.; Writing—review & editing, Y.L., H.H., L.R. and R.Y. performed the experiment and wrote the paper; H.H., L.Y. (Linfeng Yu), L.R. and Y.L. revised the papers. R.Y., H.L., L.Y. (Liyuan Yang) conducted the field investigation and collected data. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was funded by the “National Key Research & Development Program of China (2018YFD0600200)”, “Beijing’s Science and Technology Planning Project (Z191100008519004)”, “Major emergency science and technology projects of National Forestry and Grassland Administration (ZD202001-05)”, and “Chinese National Natural Science Foundation (31971581)”.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** The authors would like to thank H.N.L. and L.Y.Y. for the field investigation and Y.Q.L., H.G.H., L.F.Y. and L.L.R. for their suggestion and modification to this paper.

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

## References

- Mamiya, Y.; Kiyohara, T. Description of *Bursaphelenchus lignicolus* N. Sp. (Nematoda: Aphelenchoididae) from pine wood and histopathology of nematode-infested trees. *Nematologica* **1972**, *18*, 120–124. [[CrossRef](#)]
- Rodrigues, M. National Eradication Programme for the Pinewood Nematode. In *PineWilt Disease: A Worldwide Threat to Forest Ecosystems*; Mota, M.M., Vieira, P., Eds.; Springer: Kato Bunmeisha, Japan, 2008; pp. 5–14.
- Robertson, L.; Cobacho Arcos, S.; Escuer, M.; Santiago Merino, R.; Esparrago, G.; Abelleira, A.; Navas, A. Incidence of the pinewood nematode *Bursaphelenchus xylophilus* Steiner & Buhner, 1934 (Nickle, 1970) in Spain. *Nematology* **2011**, *13*, 755–757.
- Fonseca, L.; Santos, M.; Santos, M.; Curtis, R.; Abrantes, I. Morpho-biometrical characterization of Portuguese *Bursaphelenchus xylophilus* isolates with mucronate, digitate or round tailed females. *Phytopathol. Mediterr.* **2009**, *47*, 223–2334.
- Ye, J. Epidemic status of pine wilt disease in China and its prevention and control techniques and counter measures. *Sci. Silv. Sin.* **2019**, *55*, 1–10.
- Togashi, K.; Arakawa, Y. Horizontal Transmission of *Bursaphelenchus xylophilus* between Sexes of *Monochamus alternatus*. *J. Nematol.* **2003**, *35*, 7–16. [[PubMed](#)]
- Kobayashi, F.; Yamane, A.; Ikeda, T. The Japanese pine sawyer beetle as the vector of pine wilt disease. *Annu. Rev. Entomol.* **2003**, *29*, 115–135. [[CrossRef](#)]
- Mamiya, Y. Pathology of the pine wilt disease caused by *Bursaphelenchus xylophilus*. *Annu. Rev. Phytopathol.* **1983**, *21*, 201–220. [[CrossRef](#)] [[PubMed](#)]

9. Yu, H.; Wu, H.; Zhang, X.; Wang, L.; Zhang, X.; Song, Y. Preliminary study on *Larix* spp. infected by *Bursaphelenchus xylophilus* in natural environment. *For. Pest Dis.* **2019**, *38*, 7–10.
10. Pan, L.; Li, Y.; Liu, Z.; Meng, F.; Chen, J.; Zhang, X. Isolation and identification of pine wood nematode in *Pinus koraiensis* in Fengcheng, Liaoning Province. *For. Pest Dis.* **2019**, *38*, 1–4.
11. 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](#)]
12. Lin, X. Review on damage and control measures of pine wilt disease. *East China For. Manag.* **2015**, *29*, 28–30.
13. Allen, C.D.; Breshears, D.D.; McDowell, N.G. On underestimation of global vulnerability to tree mortality and forest die-off from hotter drought in the Anthropocene. *Ecosphere* **2015**, *6*, 1–55. [[CrossRef](#)]
14. Wong, C.M.; Daniels, L.D. Novel forest decline triggered by multiple interactions among climate, an introduced pathogen and bark beetles. *Glob. Chang. Biol.* **2017**, *23*, 1926–1941. [[CrossRef](#)] [[PubMed](#)]
15. Kärvelo, S.; Johansson, V.; Schroeder, M.; Ranius, T. Local colonization-extinction dynamics of a tree-killing bark beetle during a large-scale outbreak. *Ecosphere* **2016**, *7*, e01257. [[CrossRef](#)]
16. Umabayashi, T.; Yamada, T.; Fukuhara, K.; Endo, R.; Kusumoto, D.; Fukuda, K. In situ observation of pinewood nematode in wood. *Eur. J. Plant Pathol.* **2017**, *147*, 463–467. [[CrossRef](#)]
17. Kim, S.-R.; Lee, W.-K.; Lim, C.-H.; Kim, M.; Kafatos, M.; Lee, S.-H.; Lee, S.-S. Hyperspectral analysis of pine wilt disease to determine an optimal detection index. *Forests* **2018**, *9*, 115. [[CrossRef](#)]
18. Iordache, M.-D.; Mantas, V.; Baltazar, E.; Pauly, K.; Lewycky, N. A machine learning approach to detecting pine wilt disease using airborne spectral imagery. *Remote Sens.* **2020**, *12*, 2280. [[CrossRef](#)]
19. 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](#)]
20. Yu, R.; Luo, Y.; Zhou, Q.; Zhang, X.; Ren, L. Early detection of pine wilt disease using deep learning algorithms and UAV-based multispectral imagery. *For. Ecol. Manag.* **2021**, *497*, 119493. [[CrossRef](#)]
21. Mäyrä, J.; Keski-Saari, S.; Kivinen, S.; Tanhuanpää, T.; Hurskainen, P.; Kullberg, P.; Poikolainen, L.; Viinikka, A.; Tuominen, S.; Kumpula, T.; et al. Tree species classification from airborne hyperspectral and lidar data using 3D convolutional neural networks. *Remote Sens. Environ.* **2021**, *256*, 112322. [[CrossRef](#)]
22. Mou, L.; Ghamisi, P.; Zhu, X.X. Deep recurrent neural networks for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 3639–3655. [[CrossRef](#)]
23. LeCun, Y.; Bengio, Y.; Hinton, G. Deep Learning. *Nature* **2015**, *521*, 436–444. [[CrossRef](#)]
24. Zhang, B.; Zhao, L.; Zhang, X. Three-dimensional convolutional neural network model for tree species classification using airborne hyperspectral images. *Remote Sens. Environ.* **2020**, *247*, 111938. [[CrossRef](#)]
25. Hinton, G.; Salakhutdinov, R. Reducing the dimensionality of data with neural networks. *Science* **2006**, *313*, 504–507. [[CrossRef](#)] [[PubMed](#)]
26. Vincent, P.; Larochelle, H.; Lajoie, I.; Bengio, Y.; Manzagol, P.A.; Bottou, L. Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion. *J. Mach. Learn. Res.* **2010**, *11*, 3371–3408.
27. Krizhevsky, A.; Sutskever, I.; Hinton, G. Imagenet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.* **2012**, *25*, 1097–1105. [[CrossRef](#)]
28. 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, Columbus, OH, USA, 23–28 June 2014; pp. 580–587.
29. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv* **2014**, arXiv:1409.1556.
30. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 1–9.
31. 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](#)]
32. Tao, H.; Li, C.; Zhao, D.; Deng, S.; Jing, W. Deep learning-based dead pine trees detection from unmanned aerial vehicle images. *Int. J. Remote Sens.* **2020**, *41*, 8238–8255. [[CrossRef](#)]
33. Chen, Y.; Jiang, H.; Li, C.; Jia, X.; Ghamisi, P. Deep Feature Extraction and Classification of Hyperspectral Images Based on Convolutional Neural Networks. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 6232–6251. [[CrossRef](#)]
34. He, M.; Bo, L.; Chen, H. Multi-scale 3D deep convolutional neural network for hyperspectral image classification. In Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP), Beijing, China, 17–20 September 2017; pp. 3904–3908.
35. Li, Y.; Zhang, H.; Shen, Q. Spectral-spatial classification of hyperspectral imagery with 3D convolutional neural network. *Remote Sens.* **2017**, *9*, 67. [[CrossRef](#)]
36. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
37. He, K.; Zhang, X.; Ren, S.; Sun, J. Identity Mappings in Deep Residual Networks. *Lect. Notes Comput. Sci.* **2016**, 630–645. [[CrossRef](#)]

38. Zhong, Z.; Li, J.; Luo, Z.; Chapman, M. Spectral-spatial residual network for hyperspectral image classification: A 3-D deep learning framework. *IEEE Trans. Geosci. Remote Sens.* **2017**, *56*, 847–858. [[CrossRef](#)]
39. Lu, Z.; Xu, B.; Sun, L.; Zhan, T.; Tang, S. 3-D Channel and Spatial Attention Based Multiscale Spatial–Spectral Residual Network for Hyperspectral Image Classification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 4311–4324. [[CrossRef](#)]
40. Zhao, L.; Zhang, X.; Wu, Y.; Zhang, B. Subtropical Forest Tree Species Classification Based on 3D-CNN for Airborne Hyperspectral Data. *Sci. Silv. Sin.* **2020**, *11*, 97–107.
41. Gu, J.; Wang, J.; Braasch, H.; Burgermeister, W.; Schroder, T. Morphological and molecular characterization of mucronate isolates (M form) of *Bursaphelenchus xylophilus* (Nematoda: Aphelenchoididae). *Russ. J. Nematol.* **2011**, *19*, 103–120.
42. Iwahori, H.; Futai, K. Lipid peroxidation and ion exudation of pine callus tissues inoculated with pinewood nematodes. *Jpn. J. Nematol.* **1993**, *23*, 79–89. [[CrossRef](#)]
43. Ji, S.; Xu, W.; Yang, M.; Yu, K. 3D convolutional neural networks for human action recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2013**, *35*, 221–231. [[CrossRef](#)]
44. Tran, D.; Bourdev, L.; Fergus, R.; Torresani, L.; Paluri, M. Learning spatiotemporal features with 3D convolutional networks. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 4489–4497.
45. Hinton, G.E.; Srivastava, N.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R.R. Improving neural networks by preventing co-adaptation of feature detectors. *Comput. Sci.* **2012**, arXiv:1207.0580.
46. Congalton, R. A review of assessing the accuracy of classification of remotely sensed data. *Remote Sens. Environ.* **1991**, *37*, 35–46. [[CrossRef](#)]
47. Hao, Z.; Li, Y.; Jiang, Y. Deep learning for hyperspectral imagery classification: The state of the art and prospects. *Acta Autom. Sin.* **2018**, *44*, 961–977.
48. 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.* **2021**, *101*, 102363. [[CrossRef](#)]
49. Huo, L.; Persson, H.J.; Lindberg, E. Early detection of forest stress from European spruce bark beetle attack, and a new vegetation index: Normalized distance red & SWIR (NDRS). *Remote Sens. Environ.* **2021**, *255*, 112240.
50. Zhang, N.; Zhang, X.; Yang, G.; Zhu, C.; Huo, L.; Feng, H. Assessment of defoliation during the *Dendrolimus tabulaeformis* Tsai et Liu disaster outbreak using UAV-based hyperspectral images. *Remote Sens. Environ.* **2018**, *217*, 323–339. [[CrossRef](#)]
51. Li, Y.; Zhang, X. Research advance of pathogenic mechanism of pine wilt disease. *J. Environ. Entomol.* **2018**, *40*, 231–241.
52. Szegedy, C.; Ioffe, S.; Vanhoucke, V.; Alemi, A. Inception-v4, inception-Resnet and the impact of residual connections on learning. In Proceedings of the Thirty-first AAAI conference on artificial intelligence, San Francisco, CA, USA, 4–9 February 2017; Volume 31.
53. Xie, S.; Girshick, R.; Dollár, P.; Tu, Z.; He, K. Aggregated Residual Transformations for Deep Neural Networks. In Proceedings of the IEEE conference on computer vision and pattern recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 5987–5995.
54. Yin, J.; Qi, C.; Chen, Q.; Qu, J. Spatial-Spectral Network for Hyperspectral Image Classification: A 3-D CNN and Bi-LSTM Framework. *Remote Sens.* **2021**, *13*, 2353. [[CrossRef](#)]
55. Gong, H.; Li, Q.; Li, C.; Dai, H.; He, Z.; Wang, W.; Li, H.; Han, F.; Tuniyazi, A.; Mu, T. Multiscale Information Fusion for Hyperspectral Image Classification Based on Hybrid 2D-3D CNN. *Remote Sens.* **2021**, *13*, 2268. [[CrossRef](#)]
56. Zhang, B.; Ye, H.; Lu, W.; Huang, W.; Wu, B.; Hao, Z.; Sun, H. A Spatiotemporal Change Detection Method for Monitoring Pine Wilt Disease in a Complex Landscape Using High-Resolution Remote Sensing Imagery. *Remote Sens.* **2021**, *13*, 2083. [[CrossRef](#)]
57. 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](#)]
58. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial networks. *Adv. Neural Inf. Process. Syst.* **2014**, arXiv:1406.2661.
59. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial networks. *Commun. ACM* **2020**, *63*, 139–144. [[CrossRef](#)]
60. Zhang, M.; Gong, M.; He, H.; Zhu, S. Symmetric All Convolutional Neural-Network-Based Unsupervised Feature Extraction for Hyperspectral Images Classification. *IEEE Trans. Cybern.* **2020**, 1–13. [[CrossRef](#)] [[PubMed](#)]
61. Riaño, D.; Valladares, F.; Condés, S.; Chuvieco, E. Estimation of leaf area index and covered ground from airborne laser scanner (Lidar) in two contrasting forests. *Agric. For. Meteorol.* **2004**, *114*, 269–275. [[CrossRef](#)]
62. Donoghue, D.N.M.; Watt, P.J.; Cox, N.J.; Wilson, J. Remote sensing of species mixtures in conifer plantations using lidar height and intensity data. *Remote Sens. Environ.* **2007**, *110*, 509–522. [[CrossRef](#)]
63. Hovi, A.; Korhonen, L.; Vauhkonen, J.; Korpela, I. Lidar waveform features for tree species classification and their sensitivity to tree- and acquisition related parameters. *Remote Sens. Environ.* **2016**, *173*, 224–237. [[CrossRef](#)]
64. Liu, L.; Coops, N.C.; Aven, N.W.; Pang, Y. Mapping urban tree species using integrated airborne hyperspectral and LiDAR remote sensing data. *Remote Sens. Environ.* **2017**, *200*, 170–182. [[CrossRef](#)]
65. Li, Q.; Wong, F.; Fung, T. Mapping multi-layered mangroves from multispectral, hyperspectral, and lidar data. *Remote Sens. Environ.* **2021**, *258*, 112403. [[CrossRef](#)]
66. Meng, R.; Dennison, P.E.; Zhao, F.; Shendryk, I.; Rickert, A.; Hanavan, R.P.; Cook, B.D.; Serbin, S.P. Mapping canopy defoliation by herbivorous insects at the individual tree level using bi-temporal airborne imaging spectroscopy and LiDAR measurements. *Remote Sens. Environ.* **2018**, *215*, 170–183. [[CrossRef](#)]

- 
67. Lin, Q.; Huang, H.; Wang, J.; Huang, K.; Liu, Y. Detection of pine shoot beetle (PSB) stress on pine forests at individual tree level using UAV-based hyperspectral imagery and lidar. *Remote Sens.* **2019**, *11*, 2540. [[CrossRef](#)]
  68. Lin, Q.; Huang, H.; Chen, L.; Wang, J.; Huang, K.; Liu, Y. Using the 3D model RAPID to invert the shoot dieback ratio of vertically heterogeneous Yunnan pine forests to detect beetle damage. *Remote Sens. Environ.* **2021**, *260*, 112475. [[CrossRef](#)]