

Article

Accurate Identification of Pine Wood Nematode Disease with a Deep Convolution Neural Network

Jixia Huang ^{1,2,†}, Xiao Lu ¹ , Liyuan Chen ^{1,†} , Hong Sun ³, Shaohua Wang ⁴ and Guofei Fang ^{3,*}

¹ Key Laboratory for Silviculture and Conservation of Ministry of Education, Beijing Forestry University, Beijing 100083, China; huangjx@bjfu.edu.cn (J.H.); luxiao@bjfu.edu.cn (X.L.); chenliyuan@mail.bnu.edu.cn (L.C.)

² Key Laboratory of Land Surface Pattern and Simulation, Chinese Academy of Sciences, Beijing 100101, China

³ Center for Biological Disaster Prevention and Control, National Forestry and Grassland Administration, Shenyang 110034, China; sunhongcaf@163.com

⁴ Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China; wangshaohua@aircas.ac.cn

* Correspondence: fgfly@163.com

† These authors contributed equally to this work.

Abstract: Pine wood nematode disease is a devastating pine disease that poses a great threat to forest ecosystems. The use of remote sensing methods can achieve macroscopic and dynamic detection of this disease; however, the efficiency and accuracy of traditional remote sensing image recognition methods are not always sufficient for disease detection. Deep convolutional neural networks (D-CNNs), a technology that has emerged in recent years, have an excellent ability to learn massive, high-dimensional image features and have been widely studied and applied in classification, recognition, and detection tasks involving remote sensing images. This paper uses Gaofen-1 (GF-1) and Gaofen-2 (GF-2) remote sensing images of areas with pine wood nematode disease to construct a D-CNN sample dataset, and we train five popular models (AlexNet, GoogLeNet, SqueezeNet, ResNet-18, and VGG16) through transfer learning. Finally, we use the “macroarchitecture combined with micromodules for joint tuning and improvement” strategy to improve the model structure. The results show that the transfer learning effect of SqueezeNet on the sample dataset is better than that of other popular models and that a batch size of 64 and a learning rate of 1×10^{-4} are suitable for SqueezeNet’s transfer learning on the sample dataset. The improvement of SqueezeNet’s fire module structure by referring to the Slim module structure can effectively improve the recognition efficiency of the model, and the accuracy can reach 94.90%. The final improved model can help users accurately and efficiently conduct remote sensing monitoring of pine wood nematode disease.

Keywords: deep convolutional neural networks; pine wood nematode disease; SqueezeNet; transfer learning



Citation: 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. <https://doi.org/10.3390/rs14040913>

Academic Editor: Siyuan Wang

Received: 30 December 2021

Accepted: 7 February 2022

Published: 14 February 2022

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



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

1. Introduction

Pine wood nematode disease is one of the most dangerous forest biological infections worldwide, and it is a disease that is devastating to pine tree species. Because of its high infectivity and its high fatality rate, pine wood nematode disease is also called the “cancer” of pine trees. Pine wood nematodes originated in North America and have spread to other areas through the timber trade. The disease is currently prevalent in America, Canada, and Mexico in North America, China, Japan, South Korea, and North Korea in East Asia, and Portugal in Europe as well as in other countries. Japan has experienced the worst losses due to pine wood nematode disease. Pine wood nematode disease was reported in Nagasaki Prefecture in Japan in 1905. In the following decades, pine wood nematode disease spread to most parts of the country and caused serious economic losses [1]. In 1982, pine wood nematode disease was first reported in China in Nanjing, Jiangsu Province, and in the

following decades the disease spread to the surrounding areas. At present, the disease has spread to high-altitude areas, such as the Qinling Mountains, where it seriously threatens more than 300,000 km² of pine forests [2]. Pine wood nematode disease has caused great losses to China's forestry ecology and economy. In the 35 years between 1982 and 2017, the disease killed more than 50 million pine trees, and clear and selective pine forest cutting for epidemic control was conducted over an area of more than 4667 km². The related economic losses amount to tens of billions of dollars, and the epidemic has caused massive damage to China's forest resources and ecological environment [3].

The use of remote sensing technology to monitor forest pests offers the advantages of real-time dynamic monitoring, coverage of a large area, limited susceptibility to environmental interference, and short periods. The use of remote sensing technology to monitor pine wood nematode disease has recently made some progress. In the past, remote sensing to monitor pine wood nematode disease was mainly based on spectral histogram analysis of multispectral images. Kim et al. combined the normalized difference vegetation index (NDVI) with spectral histogram analysis of IKONOS images to identify areas affected by pine wood nematode disease [4]. Then, based on on-site spectral observations, the researchers identified the characteristics of typical bands (green, red, and near-infrared bands) and constructed spectral characteristic indicators (red-edge parameters, vegetation indices, and time series characteristics). Determining the relationship between observed spectral characteristics and plant physiological characteristics, such as chlorophyll content, transpiration rate, and water content can help detect pine wood nematode disease [5–7]. Huang et al. analyzed the hyperspectral time series characteristics and sensitive characteristics of healthy and susceptible plants, and they reported that the time series of plants infected with pine wood nematode disease showed large spectral differences, including a decrease in red-edge spectral reflectance and red-edge blueshift [6]. Multiple spectral feature values in the near-infrared, red and blue edges are significant hyperspectral features that indicate the presence of pine wood nematode disease [6]. Xu et al. collected the spectral characteristics of lodgepole and Masson pines at different susceptibility levels, and they found that the reflection spectrum curve and spectral characteristic parameters of different bands in the hyperspectral image can be used to analyze pathogenic mechanism at different stages, and the relationship model between spectral characteristics and chlorophyll can provide a reference for remote sensing monitoring of pine wood nematode disease [7]. The existing studies generally rely only on the spectral characteristics of images as the basis for the identification of pine wood nematode disease, and few studies have attempted to use new technological means of analyzing high-resolution satellite remote sensing images to recognize pine wood nematode disease.

Deep convolutional neural networks (D-CNNs), which have efficient and accurate image recognition capabilities, have been widely used in computer vision and other fields. In recent years, D-CNNs have been introduced into the field of remote sensing and used in remote sensing big data analysis [8,9]. However, the current remote sensing processing methods based on D-CNNs have most often been applied to land use classification and feature target recognition [10–12], and only a few studies have addressed forest pest monitoring and control. Ha et al. used deep learning to process images captured by unmanned aerial vehicles (UAVs) at low altitudes and to identify infected radish plants. The CNN obtained an accuracy of 93.3% [13]. Rançon et al. obtained and labelled pictures of diseased and healthy vine plants, and 91% overall accuracy was obtained using deep features extracted from the MobileNet network trained on the ImageNet database [14]. The above studies show that the use of deep learning can achieve higher accuracy than traditional machine learning methods provide. At the same time, the use of deep learning for pest detection can use pretrained networks without the need to redesign the network structure. In the monitoring of forest diseases and insect pests based on aerial images, Sylvain et al. used a D-CNN to identify the health status of trees, and the accuracy reached 94% [15]. Safonova et al. used D-CNNs to detect Siberian fir trees at different susceptibility stages based on UAV images and achieved an accuracy of 98.77% in detecting

susceptible fir trees at different stages [16]. Qiao, Deng et al. used deep learning methods to classify and detect pine wood nematode disease based on UAV images and achieved high accuracy [17,18]. Most current studies of forest diseases and insect pests that use remote sensing technology are based on images obtained by UAVs or aerial imagery. UAV images have higher image resolution and richer detailed information than satellite images, making it easy to classify and detect ground objects accurately. Satellite remote sensing images, which offer large coverage, low cost, and relatively rough spatial resolution, are not fully utilized.

At present, research on deep learning for satellite remote sensing recognition of pine wood nematode disease is lacking, especially research that focuses on suitable deep learning networks and manually labelled samples. This study explores a deep learning model that is suitable for remote sensing image classification of pine wood nematode disease and uses China's Gaofen-1 (GF-1) and Gaofen-2 (GF-2) images of pine wood nematode disease occurrence areas to construct a D-CNN sample dataset. Based on these samples, five excellent D-CNN models (AlexNet, GoogLeNet, SqueezeNet, ResNet-18, and VGG16) are selected for transfer learning, and the model with the best transfer learning effect is chosen for hyperparametric and structural optimization. The resulting model can accurately identify pine wood nematode disease. This study constructs a D-CNN model that is suitable for identifying satellite remote sensing images of pine wood nematode disease occurrence areas and provides technical support for the monitoring, prevention, and control of pine wood nematode disease.

2. Materials and Methods

2.1. Study Area

The spatial range of the remote sensing images used in this study is 41.6–42.2° N, 123.5–124.8° E (Figure 1), covering Shenyang City District, Tieling City District, Fushun City District, Kaiyuan City, Tieling County, Fushun County, Xinbin Manchu Autonomous County, Qingyuan Manchu Autonomous County, Liaoning Province, China. The research area is rich in vegetation resources and is dominated by mountain forests, such as those of the Daxi, Tiebei, and Nantianmen Mountains. The genus *Pinus* has a wide range of distribution and is present in large numbers. Among the species in the area, *Pinus densiflora* Sieb. et Zucc., *Pinus tabulaeformis* Carr., *Pinus thunbergii* Parl., and other pine tree species are hosts to pine wood nematodes. The study area has a northern temperate seasonal continental climate with cold, dry winters and warm, rainy summers, and its altitude ranges from 5.3 m to 1346.7 m. The average annual temperature is approximately 6–10 °C. The maximum temperature in August can reach 38 °C, and the minimum temperature in January can be below −35 °C. The average daily minimum temperature is above 0 °C beginning in April, and the average daily minimum temperature is below 0 °C beginning in November. An average of 600 to 850 mm of rainfall occurs yearly, and there are approximately 2500 h of sunshine annually, with longer sunshine hours in May and June and shorter sunshine hours in November and December. The annual average wind speed is 4.5 m per second. Relevant studies have shown that pine wood nematodes have a strong ability to adapt to temperatures above 0 °C [2]. The range of latitude within the study area is also suitable for pine wood nematodes, and the study area provides suitable breeding conditions for pine wood nematodes [2]. The studied area is a key area for the detection, prevention, and control of pine wood nematode disease. In recent years, the incidence of pine wood nematode disease in Fushun, Dandong, and Liaoning Province has been expanding, the degree of damage caused by these pests has been increasing, and massive economic losses have occurred [19].

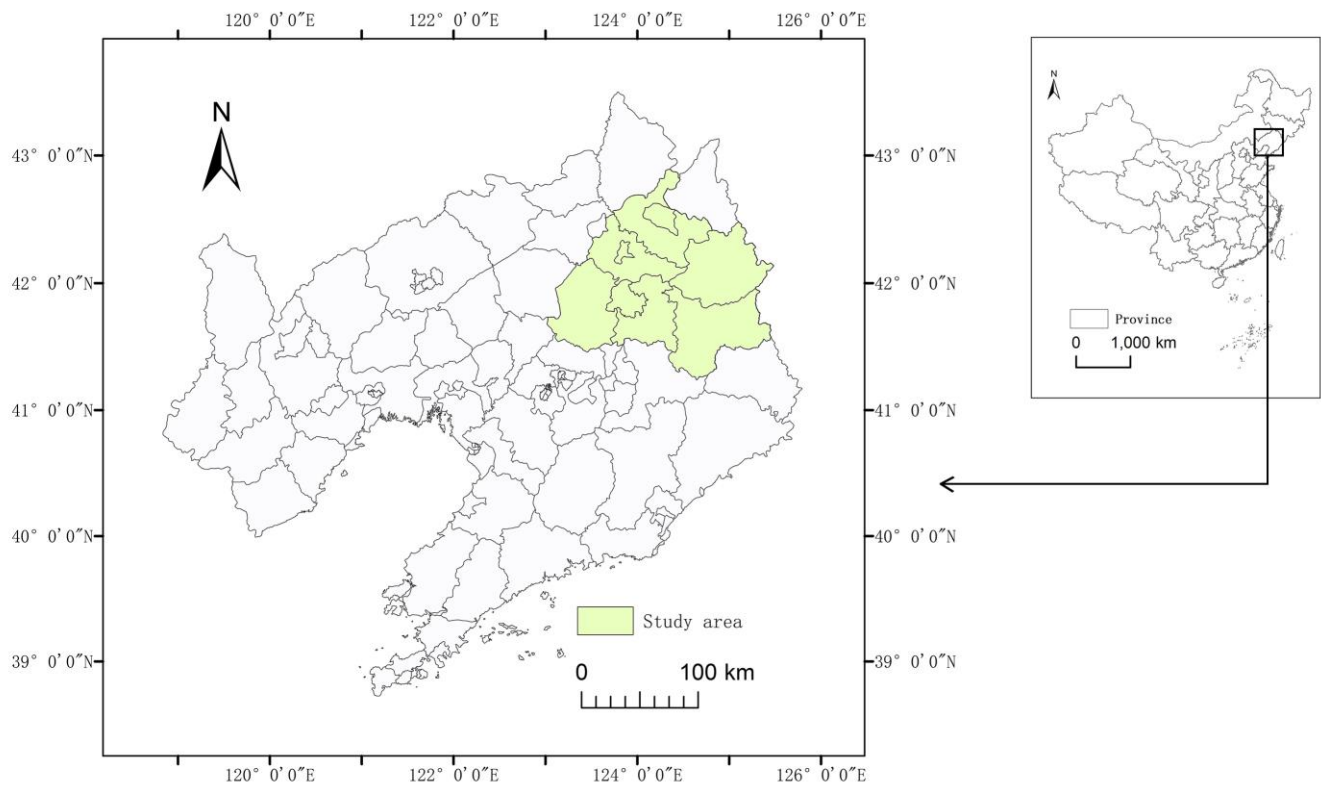


Figure 1. Liaoning Province, China.

2.2. High-Resolution Remote Sensing Image Data

This study used 99 GF-1 and 50 GF-2 images with 1A product grades. The images were obtained from May to October of each year from 2013 to 2017. Pine wood nematode disease has been reported to occur in many cities and counties in Liaoning. The images from 2015–2017 provided information on the susceptible area, and the images from 2013–2014 provided reference information for normal forestland. High-spatial-resolution satellite images with a wide imaging range and short revisit period have advantages in forestry remote sensing applications. Such images have been widely used in forest resource monitoring and forest information extraction research, and can be used effectively to detect dynamic changes in forestland and vegetation cover [20,21]. The GF-1 images were obtained using 2 panchromatic/multispectral (PMS) cameras with spatial resolutions of 2 m for the panchromatic bands and 8 m for the multispectral bands and a width greater than 69 km. The multispectrum contains 4 bands (blue, green, red, and near-infrared bands), and the revisit period is only 4 days; the system thus integrates the advantages of high spatial resolution and high temporal resolution and can accurately reflect the spatial texture characteristics of the target. The GF-2 images were obtained using two PMS cameras with a spatial resolution of 1 m panchromatic bands/4 m multispectrum at a width of 45 km. The multispectral bands are the same as those of the GF-1 system and the revisit period is 5 days, further expanding the spatial resolution to the submeter level while maintaining excellent time resolution. The width reaches the highest level among international satellites with submeter resolution.

To build a deep learning model that is suitable for remote sensing image classification of pine wood nematode disease, we mainly go through three steps: dataset construction, transfer learning, and model optimization (Figure 2).

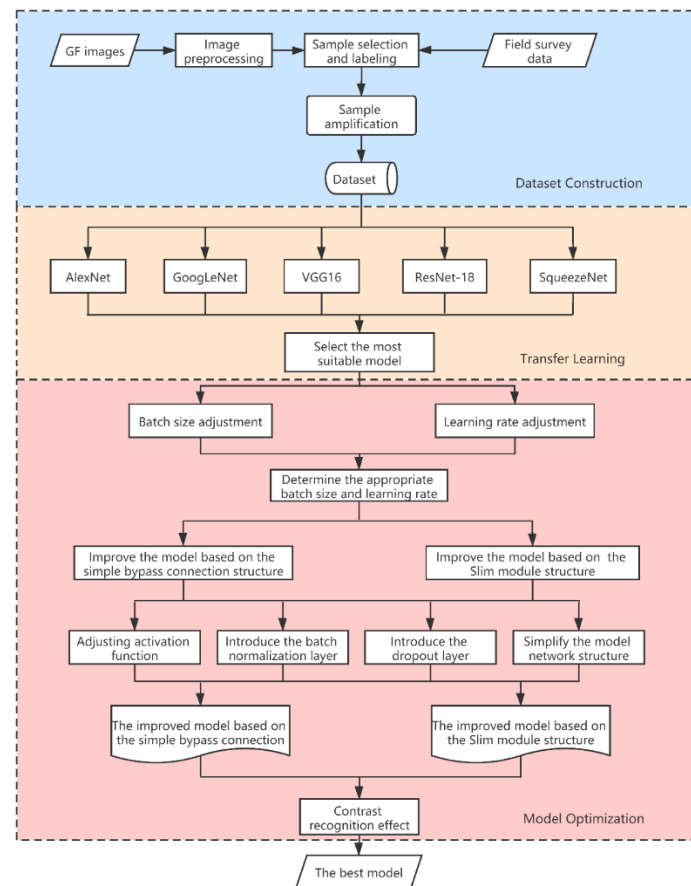


Figure 2. Flow diagram of experiment.

2.3. Construction of a Manually Annotated Sample Dataset

The sample dataset is the basis for building the D-CNN. The D-CNN iteratively learns a large number of samples and uses that information to adjust the weight parameters of each neuron to achieve the extraction and recognition of multidimensional image features. The sample dataset acts directly on the parameters of the D-CNN, which has a profound impact on the recognition performed by the model.

We constructed the sample dataset in seven steps, including image selection, image fusion, band combinations, visual interpretation, sample cutting, Jeffries–Matusita distance separability calculation, sample balance, and augmentation. First, 76 remote sensing images with low cloud cover were selected from among 149 remote sensing images. To obtain remote sensing images with high spatial resolution and containing multispectral information as the basis for constructing the samples, the multispectral bands and the panchromatic bands of the remote sensing images were merged using the NNDiffuse pan sharpening method. The use of a combination of bands highlights the spectral characteristics of the vegetation disturbed by pine wood nematode disease. A large number of studies have confirmed that the red and green bands are very sensitive to color and to physiological changes caused by pine wood nematode disease. The red–green ratio index ($RGRI = R/G$) was calculated as one of the discriminant spectra. The near-infrared band is the most sensitive to changes caused by pine wood nematode disease, and the spectral difference between diseased and healthy plants is the largest in this range [22]. The blue band can be used to increase the dimensionality of the spectral features. Therefore, RGRI, the near-infrared band, and the blue band are used as the input bands of the R, G, and B channels to synthesize the base image used to label the sample.

In the visual interpretation of the preprocessed remote sensing images based on the field survey data provided by the Forest and Grassland Pest Control Station of the State

Forestry and Grassland Administration of China (<http://www.forestpest.org/> (accessed on 20 December 2018)), the spectral, textural, and other characteristics of the images obtained from each area in the corresponding periods are compared and evaluated for the presence of features characteristic of pine wood nematode disease, and the visual interpretation characteristics of the pine wood nematode disease-affected area are determined (Figure 3). Plants in areas affected by pine wood nematode disease often present dark blue-green/blue-violet/dark green discolored areas, needles that show wilting and a clear granular texture without a large amount of shedding, and a clustered spatial distribution. Healthy forests are mostly characterized by green/light green areas with dense canopies and no obvious texture.

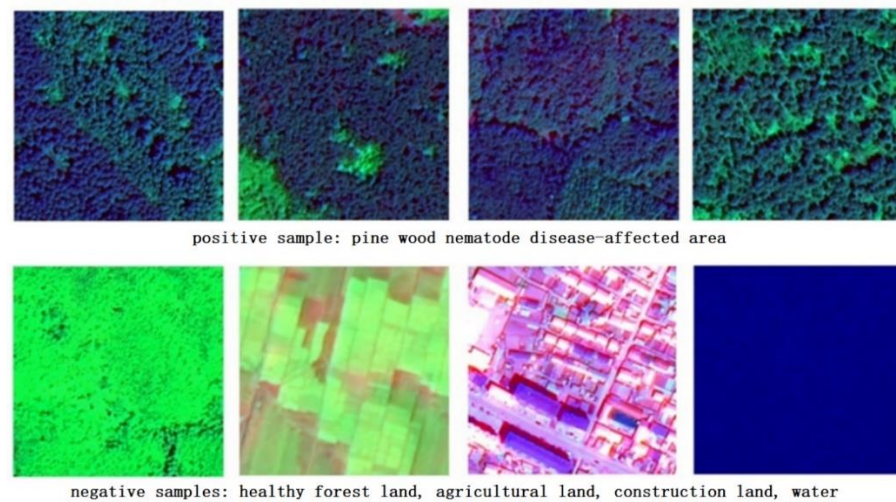


Figure 3. Examples of the samples. The first row shows images of a pine wood nematode disease-affected area. In the second row, from left to right, are images of healthy forestland, agricultural land, construction land, and water.

The samples are labelled according to their visual interpretation features. ENVI is used for precise positioning and cutting of the samples. The uniform sample size is $200 \times 200 \times 3$ pixels, and all of the sample images are in TIFF/GeoTIFF format. After classification and labelling of each image, the labelled samples are divided into positive samples and negative samples. Positive samples represent woodlands infected with pine wood nematode disease and provide a direct reference for the identification and classification of research targets. They were generated from images obtained during the period of onset of pine wood nematode disease (2015–2017) in the study area. The negative samples present a collection of noninfectious surface features, including healthy forestland, agricultural land, construction land, and water (Figure 3).

To avoid training errors caused by the phenomenon of “different bodies with the same spectrum” and “the same bodies with different spectra” between high-resolution images in susceptible forest land, healthy forest land, and agricultural land, this study uses the above three types of samples from the same image data source to calculate the separability of the Jeffries–Matusita (JM) distance for each image, and the result is used for separability testing and sample optimization [23]. To minimize the feature learning bias that may result from the use of imbalanced data, this research uses the label shuffle algorithm proposed by Hikvision as the sample category balancing strategy [24]. The labelled samples are flipped horizontally and vertically to increase the sample size 3-fold and thereby obtain the final sample size. When training and verifying samples, random image flipping is performed to improve the generalizability of the model.

According to the network training requirements, the samples are further divided into a training/validation dataset and a test dataset. A total of 3570 samples were used for training and testing. Of these, 3030 samples were used for model training and validation,

and 540 samples were used for model testing. Of the 3030 training/validation samples, 2424 samples were used for training, and 606 samples were used for validation.

2.4. Transfer Learning of D-CNN

This study implements the transfer learning of CNNs. Five commonly used models pretrained on ImageNet, including AlexNet [25], GoogLeNet [26], SqueezeNet [27], ResNet-18 [28], and VGG16 [29], are selected for training. These five models exhibit high accuracy in many tasks. We carried out the experiment with MATLAB software, and the training environment used in this study is: Windows10 64 bit operating system, 8 GB RAM, i7-5500U quad-core processor, and NVIDIA GeForce920M was used to accelerate the training of the models. The commonly used hyperparameters for transfer learning are set as follows: the batch size is 64, the learning rate is 0.001, and a total of 20 epochs of training are performed. To reduce memory usage during VGG16 model training, training is performed under the conditions of small sample size, small batch size (16), and a high learning rate (0.1). According to (i) the training time of the model, (ii) the classification accuracy result of the validation data, (iii) the convergence speed (runtime) of the model, and (iv) the stability of the accuracy and loss after convergence, we compare the effects of these pretrained network models on the transfer learning of the sample dataset and determine the best model for subsequent research.

2.5. Training Parameter Optimization of the Deep Convolution Neural Network

Not only the structure of the D-CNN but also the hyperparameters set in the training model have a direct impact on the transfer learning effect of the network. The initial training parameters that have a decisive effect on the feature learning performance are the batch size and the learning rate.

Increasing the batch size within a reasonable range can improve the efficiency of hardware memory usage, reduce the number of parameter updates (iterations) during each epoch, speed up processing of the same amount of data, improve the accuracy of the stochastic gradient descent direction, and stabilize the model training process. If the batch size exceeds a reasonable range, the full batch learning strategy may, in extreme cases, lead to insufficient hardware memory capacity and slow changes in the direction of the stochastic gradient descent, resulting in slow model training.

A learning rate that is too high will cause the model to fall too rapidly and thereby fail to arrive at the solution needed to minimize the loss function, and this will limit or even reduce the model's classification accuracy. In contrast, a learning rate that is too low will cause the correction of the weight parameter to be slow, and this may make the model fall into the local optimal solution of the loss function instead of the global optimal solution; this not only reduces the network training speed but also fails to achieve appropriate model accuracy.

In this study, the batch size is set to 32, 64, 128, and 256, values that are commonly used in previous research and applications, and the transfer training effects of suitable models under these 4 batch size conditions are compared. To compare the transfer training effects of suitable models under different learning rates, the learning rate parameter is set to a constant learning rate series and to a variable learning rate. The constant learning rate series includes $1e-4$, $5e-4$, $1e-3$, $3e-3$, and $5e-3$, the initial value of the variable learning rate is $1e-3$, the drop coefficient is 0.5, and the variable learning rate changes every 5 training epochs.

2.6. Structure Optimization of the Deep Convolution Neural Network

The diversity of the structural design of D-CNNs and the complex interactions between the various layers of the network provide room for improving the network model, making it possible for the existing model to achieve optimal training accuracy and efficiency through structural adjustments. This study adopts the strategy of "macroarchitecture combined with a micromodule for joint tuning and improvement" to improve the best model obtained

by transfer learning. The macroarchitecture of the model is improved using two model structure optimization methods: one method is based on a simple bypass connection structure [27], and the other is based on a Slim module structure [30]. The micromodules are adjusted by replacing the activation functions, introducing a batch normalization (BN) layer and a dropout layer and reducing the network structure. For model optimization, we compare the learning effect of each adjustment strategy using the sample dataset.

The improvement based on the simple bypass connection structure involves adding shortcut connections to the D-CNN that skip one or more layers and one or more modules [27,28]. When the network deepens, the use of shortcut connections can partially solve the network degradation problem and alleviate the disappearing gradient problem during back propagation.

Based on the improvement of the Slim module structure, one or more modules in the model are replaced with a Slim module. The Slim module introduces the idea of group convolutions and singular bottlenecks. The group convolution is for the channel: the input channel is divided into multiple groups so that the convolutions reduce the number of parameters and the number of calculations. The singular bottleneck is a nonlinear transformation that is preserved only once in the structure of the bottleneck, thereby improving the classification accuracy [30].

The input layer of the proposed network follows the squeezenet, with the input size of $227 \times 227 \times 3$ and zero-center normalization. The output layer is the softmax layer. It performs the classification by respectively calculating the probability of five categories (pine nematode disease-affected area, healthy forest, agricultural land, construction land, and water) of each feature map.

2.7. Evaluation of the Recognition Effect

This study uses evaluation indicators that are widely used in existing studies [31], namely, overall accuracy (OA), recall (true positive rate, TPR), and false alarm rate (false positive rate, FPR), to evaluate the recognition effect of the improved D-CNN model on the test samples. In addition, considering that pine wood nematode disease-affected areas and healthy forestlands are easily confused, the inter-forestland TPR (TPRF) and the inter-forestland FPR (FPRF) indicators for the two classifications are also calculated. The formulas used to calculate the evaluation indices are as follows:

$$OA = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$TPR = \frac{TP}{TP + FN} \quad (2)$$

$$FPR = \frac{FP}{FP + TN} \quad (3)$$

$$TPRF = \frac{TP}{TP + FN_{forest}} \quad (4)$$

$$FPRF = \frac{FP_{forest}}{FP_{forest} + TN_{forest}} \quad (5)$$

In the formulas, OA represents the ratio of correctly identified samples to total samples, TP represents the number of correctly identified samples from susceptible areas, FN represents the number of samples from susceptible areas that were incorrectly identified as nonsusceptible areas, FP represents the number of samples from nonsusceptible areas that were incorrectly identified as susceptible areas, and TN represents the number of correctly identified samples from nonsusceptible areas. FN_{forest} represents the number of samples from susceptible areas that were incorrectly identified as healthy forestlands, FP_{forest} represents the number of samples from healthy forestlands that were incorrectly identified as susceptible areas, and TN_{forest} represents the number of correctly identified samples from healthy forestlands. The TPR represents the accuracy of samples from pine wood

nematode disease-affected areas identified by the network model; the FPR represents the false positive rate of samples from nonsusceptible areas identified by the network model. The TPRF and FPRF are similar to the TPR and FPR, respectively, mainly reflecting the correct identification of positive samples and the misclassification of negative samples between forestlands. In general, the larger the OA, TPR, and TPRF are, the smaller the FPR and FPRF are and the better is the model's ability to recognize pine wood nematode disease-affected areas.

3. Results

3.1. Transfer Learning of D-CNN

Research has shown that SqueezeNet is the most suitable model for automatically identifying pine wood nematode disease-affected areas. The model achieved a good balance between training time and classification accuracy (Table 1). SqueezeNet's training time and verification sample accuracy are better than those of VGG16. Compared with GoogLeNet, SqueezeNet needs less than 1/2 of the training time to exceed the former's verification sample accuracy. Compared with AlexNet, SqueezeNet increases the classification accuracy by 3.25% at the cost of only 3 min of training time. In contrast, compared to ResNet-18, SqueezeNet's classification accuracy is only 0.32% lower, but the training time required by SqueezeNet is nearly 42 min less than that required by ResNet-18. In addition, the convergence speed of the SqueezeNet training process is better than that of the other networks, and its recognition accuracy of the verification samples is closest to that of the training samples.

Table 1. Comparison of the training effects of the different network models.

	AlexNet	VGG16	GoogLeNet	ResNet-18	SqueezeNet
Training time	27 min, 54 s	282 min, 29 s	67 min, 21 s	72 min, 29 s	30 min, 46 s
Validation sample accuracy	95.45%	47.57%	98.05%	99.02%	98.70%
Convergence rate (time)	3 epochs	4 epochs	2 epochs	1 epoch	1 epoch
Stability after convergence	Stable	Fluctuating	Stable	Very stable	Very stable

3.2. Suitable Training Parameters

Comparing the transfer training effects of SqueezeNet under different batch sizes and learning rates, it is found that a batch size of 64 is optimal for SqueezeNet's transfer learning on the sample dataset. A learning rate of 1×10^{-4} is appropriate for SqueezeNet to perform transfer learning on the sample dataset.

From the comparison results (Table 2), it can be inferred that a batch size of 64 is the appropriate batch size for SqueezeNet to perform transfer learning on the sample dataset. First, a batch size of 64 will be more representative of the samples during each parameter update than a batch size of 32, making the direction of the stochastic gradient descent more accurate; the convergence speed of the model can thereby be accelerated or maintained at a high-speed level, with less fluctuation occurring during training. Second, increasing the batch size to 64 reduces the number of iterations in each epoch and enables the model to make full use of the amount of sample information to achieve a balance between the parameter update frequency and the accuracy of the gradient descent direction; this improves the classification accuracy and achieves a balance between the number of iterations and the amount of learning each time. The time required for training is shortened by nearly 1.5 min. Therefore, it can be concluded that a batch size of 64 is reasonable. Comparing the effect of increasing the batch size to 128 or 256, it can be found that these excessive increases reduce the number of parameter updates. In addition, the memory occupancy and computation time increase, resulting in a degradation of convergence speed and training duration.

Table 2. SqueezeNet training effects of different batch sizes.

	32	64	128	256
Training time	32 min, 11 s	30 min, 46 s	32 min, 55 s	58 min, 9 s
Validation sample accuracy	98.05%	98.70%	98.86%	99.19%
Convergence rate (time)	1 epoch	1 epoch	2 epochs	3 epochs
Stability after convergence	Stable	Very stable	Very stable	Stable

From the comparison results (Table 3), it can be inferred that a learning rate of 1×10^{-4} is the appropriate learning rate for SqueezeNet to perform transfer learning on the sample dataset. First, when a small learning rate is used as the initial parameter, as is the case in this study, differences in the learning rates did not lead to significant changes in the training time; all of the training times are close to 30 min and 40 s, indicating that the parameters set for each training process do not fall into the local optimum solution that causes the gradient to decrease slowly. Second, comparison of the learning rates over the commonly used range of 5×10^{-3} to 1×10^{-3} shows that a reasonable decrease in the learning rate increases the model classification accuracy, accelerates the convergence speed, and enhances the stability. This phenomenon may arise from the fact that the features of the sample dataset are complex, and the network needs to realize fine learning for the features through a small learning rate within a reasonable range. Furthermore, reducing the learning rate again by a constant or variable amount can still improve the training effect of the model. The overall effect of the model is best when the learning rate is 1×10^{-4} , the model accuracy converges quickly, the validation sample accuracy reaches nearly 100%, and the model remains stable during the short training process.

Table 3. Effect of different learning rates on SqueezeNet training.

	5×10^{-3}	3×10^{-3}	1×10^{-3}	5×10^{-4}	1×10^{-4}	Piecewise
Training time	30 min, 48 s	30 min, 39 s	30 min, 46 s	30 min, 46 s	30 min, 40 s	30 min, 39 s
Validation sample accuracy	57.07%	94.80%	98.70%	99.67%	99.84%	98.86%
Convergence rate (time)	9 epochs	7 epochs	1 epoch	1 epoch	1 epoch	1 epoch
Stability after convergence	Greatly fluctuating	Fluctuating	Stable	Stable	Stable	Stable

3.3. Structure Optimization of SqueezeNet

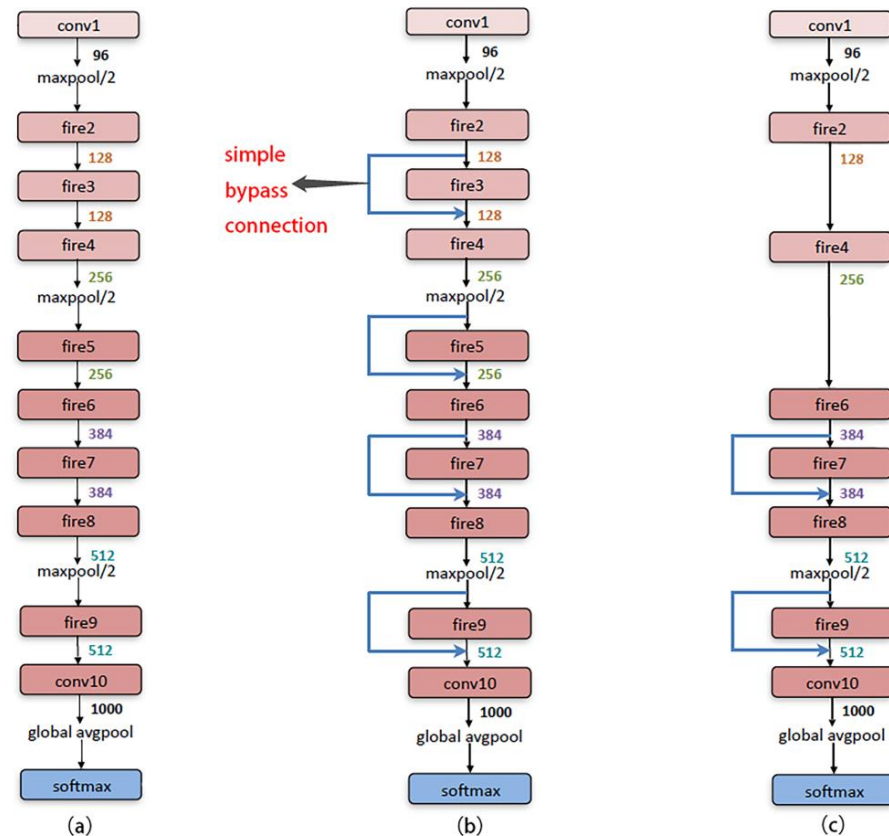
Simple bypass connections are added to the network, and microscopic modules, such as the replacement activation function are adjusted to train the model using the appropriate initial training parameters obtained in the previous section.

Comparing the training results for each tuning strategy (Table 4), based on the simple bypass connection structure condition, the best improved model can be obtained by the following adjustments (Figure 4): the BN layer is introduced into Fire8 and Fire9, an additional dropout layer is introduced, Fire3 and Fire5 are reduced, and a rectified linear unit (ReLU) function is selected as the activation function. Through this adjustment strategy, the accuracy of the model with respect to the validation samples reaches 97.89%. Although the accuracy of the adjusted model is not the highest, its training time is 12 min and 29 s less than the adjustment strategy with the highest accuracy. Therefore, the best improved model based on the simple bypass connection architecture is chosen as a suitable improved model for subsequent comparative analysis.

Table 4. Comparison of the effects of individual tuning strategies based on a simple bypass connection.

Strategies	Training Time	Validation Sample Accuracy	Convergence Rate (Time)	Stability after Convergence
ReLU	42 min, 28 s	71.38%	2 epochs	Greatly fluctuating
Leaky ReLU	43 min, 24 s	93.50%	14 epochs	Fluctuating
ELU	59 min, 6 s	92.20%	10 epochs	Fluctuating
tanh	38 min, 44 s	92.36%	7 epochs	Greatly fluctuating
A + ReLU	42 min, 15 s	96.91%	3 epochs	Fluctuating
A + Leaky ReLU	43 min, 41 s	98.54%	3 epochs	Stable
A + ELU	59 min, 19 s	96.59%	2 epochs	Fluctuating
A + tanh	39 min, 19 s	96.91%	1 epoch	Stable
A + B + ReLU	31 min, 9 s	97.24%	3 epochs	Fluctuating
A + B + Leaky ReLU	32 min, 7 s	96.26%	3 epochs	Fluctuating
A + B + ELU	43 min, 16 s	95.12%	2 epochs	Fluctuating
A + B + tanh	30 min, 47 s	96.91%	1 epoch	Fluctuating
A + B + C + ReLU	31 min, 12 s	97.89%	4 epochs	Stable
A + B + C + tanh	30 min, 25 s	94.15%	1 epoch	Stable

Note: A, B, and C in the table represent different adjustment strategies. A represents the introduction of a BN layer into Fire8 and Fire9, B represents the removal of Fire3 and Fire5, and C represents the introduction of a dropout layer to the model.

**Figure 4.** (a) SqueezeNet [27]; (b) SqueezeNet with simple bypass; (c) the improved model based on simple bypass connections.

Unlike simple bypass connections, which improve the cascading mode between modules, the Slim module implements another macrostructure improvement mode of SqueezeNet and improves the network module organization. Based on the experience gained through the previous model adjustment process, this study uses an improved structure of SqueezeNet based on the Slim module structure as a suitable improvement model. The specific adjustments are as follows (Figure 5): (1) introduce the idea of the bottleneck convolution structure in the Slim module and reduce the activation layer, reduce the convolution operation parameters, reduce the nonlinear calculation amount, and improve the network training speed; (2) reduce redundant convolutional layers, retain the two complete Slim module structures (Slim8 and Slim9) at the end of the model, and reduce Slim3, Slim5, and Slim7; (3) add the BN layer to the last two Slim modules to reduce the model's dependence on the initial parameters and improve the model training speed; and (4) introduce an additional dropout layer in Slim8 to improve the generalizability of the model.

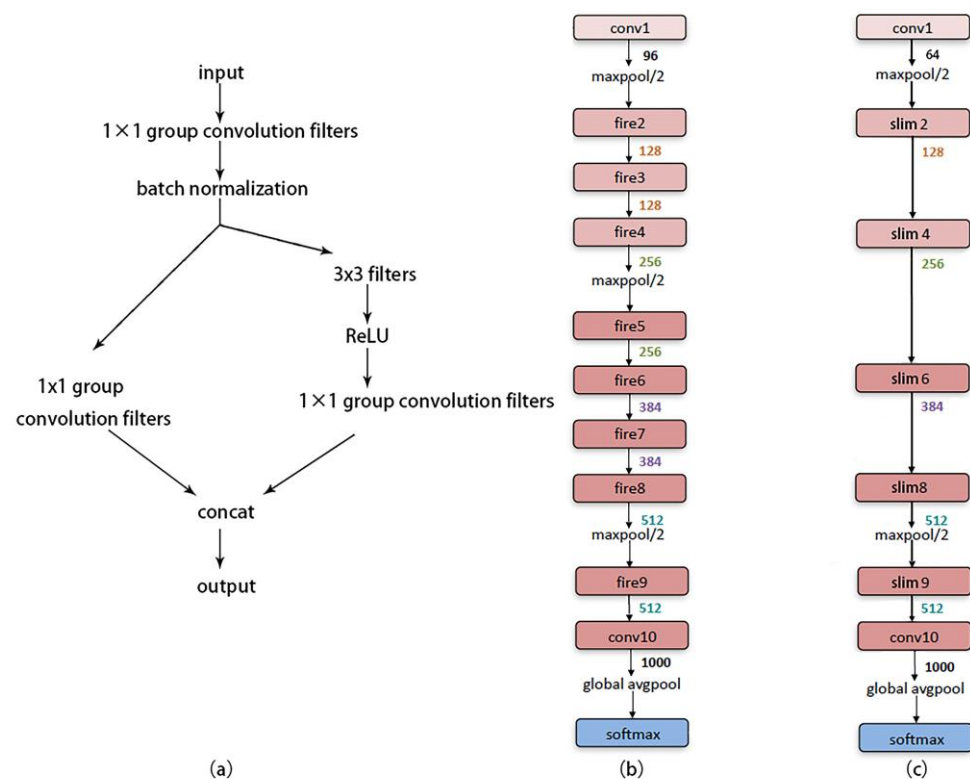


Figure 5. (a) Slim module [30]; (b) SqueezeNet [27]; (c) the improved model based on the Slim module structure.

Comparing the training effects for the original SqueezeNet model, the improved model based on the simple bypass connection structure, and the improved model based on the Slim module structure, it is found that compared with the training results obtained for the original SqueezeNet model (Table 5), the improvement in the module cascading mode through the simple bypass connection did not obviously improve the training effect of the network, but the convergence speed and the stability of the model decreased significantly. The improvement in the model based on the Slim module structure has the advantage of shortening the training time, and the accuracy, convergence speed, and stability of this model are basically equivalent to that of the original SqueezeNet model. This structure has a training time that is faster than that of SqueezeNet by nearly 8 min, equivalent to an approximately 25% increase in efficiency.

Table 5. Comparison of effects before and after improvement.

	Original SqueezeNet	Improved Model Based on the Simple Bypass Connections	Improved Model Based on the Slim Module Structure
Training time	30 min, 40 s	31 min, 12 s	22 min, 59 s
Validation sample accuracy	99.84%	97.89%	96.59%
Convergence rate (time)	1 epoch	4 epochs	2 epochs
Stability after convergence	Very stable	Fluctuating	Stable
Test data processing time	11.75 s	11.21 s	9.76 s
OA	96.68%	91.93%	94.90%
TPR	99.42%	94.74%	95.91%
FPR	4.39%	6.59%	1.80%
TPRF	99.42%	95.58%	96.76%
FPRF	10.71%	18.03%	4.69%

Comparison of the classification accuracy of the test samples between each improved model and the original model (Table 5) showed that both improved methods shorten the classification processing time of the same number of samples; this is more obvious for the improved network based on the Slim module structure, in which the classification time is shortened by 3.2 s. After the improvement, the overall classification accuracy of the model is slightly reduced, but compared with the original model, the TPR and FPR of the improved model based on the Slim module structure are reduced; in particular, the FPR is significantly reduced to 1.80%. When considering only the classification accuracy between forestland types, the improved model based on the Slim module structure achieves a significant decrease in the FPRF (6.02%) at a smaller accuracy cost (1.78%). The performance of the improved model regarding avoiding misclassifications of healthy forestland is even better.

Compared with the original SqueezeNet model, the improved model based on the Slim module structure has higher training and classification efficiency, good classification accuracy, and a more balanced distribution of the accuracy and FPR of the test sample. It can sensitively identify diseased areas and reduce false alarms and is consistent with actual application requirements. Therefore, in this study, this model was selected for use in the image classification of areas in which pine wood nematode disease occurs.

4. Discussion

This study mainly explored the method and effect of using D-CNN technology to identify pine wood nematode disease-affected areas using high-spatial-resolution satellite remote sensing images. A sample dataset is constructed based on GF remote sensing images of areas in which pine wood nematode disease is present. Using five commonly used CNN models for transfer learning, SqueezeNet is found to be the best model for transfer learning of the sample dataset. The training parameters of SqueezeNet are then optimized, and it is found that a batch size of 64 and a learning rate of 1×10^{-4} are suitable. Then, using the strategy of “macroarchitecture combined with micromodule for joint tuning and improvement” to optimize the SqueezeNet structure, it is found that an improved model based on the Slim module structure has the best accuracy for identifying pine wood nematode disease occurrence areas. The improved model can be used to identify areas susceptible to pine wood nematode disease and provides an important technical method for the monitoring and control of pine wood nematode disease.

Although some studies have shown that conventional image processing techniques can accurately identify trees infected with pine nematode disease [32,33], such identification has two basic requirements. The first is the need for a large amount of data, including ground survey data, forest cover data, satellite remote sensing data, airborne aerial photography data, and other types of data, from multiple sources. Second, airborne images with a

resolution of 20 cm or higher are needed, and even satellite images with a resolution of 0.5 m cannot meet the requirements for identification of areas affected by the disease. The cost of traditional technology is very high and this limits its scope of application. A few previous studies [16–18,34,35] have applied deep learning techniques to remote sensing images to detect and identify forest pests. In these studies, deep learning technology is basically applied to UAV remote sensing images to enable the identification, classification, and detection of damaged trees; for this purpose, airborne images of 20 cm or higher resolution are needed, and no studies based on high-resolution satellite remote sensing imagery have been performed. Based on UAV images, Deng et al. used the improved faster region convolutional neural network method to detect trees killed by pine wood nematode disease, and the detection accuracy reached approximately 90% [18]. Safonova et al. used a D-CNN based on UAV images to detect fir trees in different susceptible stages, and the detection accuracy of fir in some stages reached 98.77% [16]. UAV images have higher resolution and more richly detailed information than high-resolution satellite remote sensing images and can be used to classify and detect objects more accurately. However, high-resolution satellite remote sensing images have the advantages of large coverage, wide monitoring area, richness of time-series information, and low cost, and they can therefore be applied over large areas. This study, which is based on high-resolution satellite remote sensing images, uses the improved SqueezeNet model based on the Slim module structure to classify the test samples with an accuracy of 94.90%; thus, it can better identify and classify images of areas in which pine wood nematode disease is present than the comparison methods.

In many cases, when we use deep learning technology for classification and recognition tasks, we do not create a new D-CNN model but select existing network models that have strong feature extraction ability, high classification accuracy, and pretraining for transfer learning. Based on the powerful ability of the existing weighting parameters in the pretrained network to extract rich features from natural images and the basic features common among samples from different datasets, the network model can be adapted to new visual tasks with minimal weight readjustment. However, different network models employ different design philosophies, model structures, and weighting parameters, and these differences have different effects on the classification and recognition of new datasets. In this study, we use five popular pretrained models with strong feature extraction capabilities, namely, AlexNet, GoogLeNet, SqueezeNet, ResNet-18, and VGG16, for transfer learning on sample datasets to find the most appropriate network model for our task.

D-CNN models have much room for design optimization. A number of scholars have proposed excellent model optimization strategies, such as the use of a 1×1 convolution kernel in GoogLeNet to reduce the number of parameters and the use of a residual structure to solve the network degradation problem in ResNet. SqueezeNet uses fire modules and global average pooling to replace the fully connected layers and thus to compress the parameters significantly. At the same time, it retains large feature maps before global average pooling, thereby preserving more information and improving the classification accuracy of the model. This study optimizes SqueezeNet using improvements based on a simple bypass connection structure and improvements based on the Slim module structure. The training speed of the improved method based on the Slim module structure is faster than that of the model based on a simple bypass connection; this is directly related to the former's use of the group convolution strategy to reduce the number of weighting parameters and operations, and the reduction in the number of weighting parameters does not have a significant impact on the accuracy of identification of areas in which pine wood nematode disease is present. In addition, the optimization strategies of replacing the activation function, introducing a BN layer, reducing the number of modules, and introducing a dropout layer were conducted in a step-by-step manner. The results show that these methods are very helpful in improving the performance of the network model in terms of both speed and accuracy. Introducing a BN layer can improve gradient dispersion and thus improve the training accuracy; reducing the number of modules can remove

redundant layers and thus speed up network training, and introducing a dropout layer can prevent network overfitting in some structures, thereby improving generalizability.

In this study, deep learning technology was applied to the classification of pine wood nematode disease satellite remote sensing images, and good results were achieved with an accuracy of 94.90%. However, some aspects of this work need to be improved and expanded. In this study, only GF-1 and GF-2 images were used in the construction of the datasets. The use of only a few data types limits the scope of application of the trained network model. In addition, this study used D-CNN to identify and classify satellite remote sensing images of pine wood nematode disease occurrence areas, but it did not attempt to detect dead wood or to study how factors, such as the age of the trees and the characteristics of the terrain affect the results. Other causes, such as drought, can also kill pine trees, combined with ground investigation, the error can be limited to an acceptable range. These areas of research are key areas in which research will be conducted next.

5. Conclusions

Satellite remote sensing image processing methods based on deep learning have seldom been applied for the detection of forest pests and diseases, especially pine wood nematode disease. In this paper, through transfer learning of five commonly used D-CNN models and structural parameter adjustment, an improved model based on the Slim module structure is obtained, and the classification accuracy of the test samples reaches 94.90%. The experimental results show that the improved model based on the Slim module structure can obtain good results in identifying and classifying remote sensing images of pine wood nematode disease-affected areas and that it can achieve dynamic macroscopic, accurate, and efficient monitoring of pine wood nematode disease-affected areas. Information and decision support should be provided for the monitoring and control of pine wood nematode disease, and ecological and economic losses should be reduced. This study also provides a reference for the application of D-CNN technology to forest disturbance and forest resource monitoring.

Author Contributions: Conceptualization, J.H. and G.F.; methodology, L.C. and J.H.; validation, L.C. and J.H.; formal analysis, L.C.; resources, J.H. and G.F.; data curation, H.S. and L.C.; writing—original draft, X.L., S.W. and L.C.; writing—review and editing, X.L., L.C. and J.H.; project administration, J.H. and G.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by National Science and Technology Major Project [No.21-Y30B02-9001-19/22], National Key Research and Development Program of China (No.2019YFA0606600), Major Emergency Science and Technology Projects of State Forestry and Grassland Administration [ZD202001-06].

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

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

References

1. Rodrigues, J.M. *Pine Wilt Disease: A Worldwide Threat to Forest Ecosystems*; Springer: Dordrecht, The Netherlands, 2008.
2. Li, Y.X.; Zhang, X.Y. High risk of invasion and expansion of pine wood nematode in middle temperate zone of china. *J. Temp. For. Res.* **2018**, *1*, 3–6.
3. 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.
4. Kim, J.B.; Kim, Y.K.; Jo, M.H.; Kim, I.H. A study on the extraction of damaged area by pine wood nematode using high resolution IKONOS satellite images and GPS. *J. Korean For. Soc.* **2003**, *92*, 362–366.
5. Wang, Z. Dynamic Monitoring of Forest Ecosystem with Remote Sensing Technology after *Bursaphelenchus Xylophilus* Invasion. Master's Thesis, Beijing Forestry University, Beijing, China, 2007.
6. Mingxiang, H.; Jianhua, G.; Shun, L.; Bo, Z.; Qianting, H. Study on pine wilt disease hyper-spectral time series and sensitive features. *Remote Sens. Technol. Appl.* **2013**, *27*, 954–960.

7. Xu, H.; Luo, Y.; Zhang, T.; Shi, Y. Changes of reflectance spectra of pine needles in different stage after being infected by pine wood nematode. *Spectrosc. Spect. Anal.* **2011**, *31*, 1352–1356.
8. Zhang, L.; Zhang, L.; Du, B. Deep learning for remote sensing data: A technical tutorial on the state of the art. *IEEE Geosci. Remote Sens. Mag.* **2016**, *4*, 22–40. [[CrossRef](#)]
9. Li, W.; Hsu, C.Y. Automated terrain feature identification from remote sensing imagery: A deep learning approach. *Int. J. Geogr. Inf. Sci.* **2020**, *34*, 637–660. [[CrossRef](#)]
10. Zhang, C.; Sargent, I.; Pan, X.; Li, H.; Gardiner, A.; Hare, J.; Atkinson, P.M. Joint Deep Learning for land cover and land use classification. *Remote Sens. Environ.* **2019**, *221*, 173–187. [[CrossRef](#)]
11. Carranza-García, M.; García-Gutiérrez, J.; Riquelme, J.C. A framework for evaluating land use and land cover classification using convolutional neural networks. *Remote Sens.* **2019**, *11*, 274. [[CrossRef](#)]
12. Lin, Z.; Ji, K.; Leng, X.; Kuang, G. Squeeze and excitation rank faster R-CNN for ship detection in SAR images. *IEEE Geosci. Remote Sens. Lett.* **2018**, *16*, 751–755. [[CrossRef](#)]
13. Ha, J.G.; Moon, H.; Kwak, J.T.; Hassan, S.I.; Dang, M.; Lee, O.N.; Park, H.Y. Deep convolutional neural network for classifying Fusarium wilt of radish from unmanned aerial vehicles. *J. Appl. Remote Sens.* **2017**, *11*, 42621. [[CrossRef](#)]
14. Rañon, F.; Bombrun, L.; Keresztes, B.; Germain, C. Comparison of SIFT encoded and deep learning features for the classification and detection of Esca disease in Bordeaux vineyards. *Remote Sens.* **2019**, *11*, 1. [[CrossRef](#)]
15. Sylvain, J.; Drolet, G.; Brown, N. Mapping dead forest cover using a deep convolutional neural network and digital aerial photography. *ISPRS J. Photogramm.* **2019**, *156*, 14–26. [[CrossRef](#)]
16. Safonova, A.; Tabik, S.; Alcaraz-Segura, D.; Rubtsov, A.; Maglinets, Y.; Herrera, F. Detection of fir trees (*Abies sibirica*) damaged by the bark beetle in unmanned aerial vehicle images with deep learning. *Remote Sens.* **2019**, *11*, 643. [[CrossRef](#)]
17. Qiao, R.; Ghodsi, A.; Wu, H.; Chang, Y.; Wang, C. Simple weakly supervised deep learning pipeline for detecting individual red-attacked trees in VHR remote sensing images. *Remote Sens. Lett.* **2020**, *11*, 650–658. [[CrossRef](#)]
18. 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](#)]
19. Yu, H.Y.; Wu, H. Discovery of new host plants and new vector insects of the pine beetle nematode in Liaoning. *For. Pest. Dis.* **2020**, *37*, 61.
20. Yin, L.; Qin, X.; Sun, G.; Zu, X.; Deng, G.; Cassanova, J. Method for forest vegetation change monitoring using GF-1 images. *ESA-SP* **2016**, *739*, 98.
21. Hao, R.; Chen, E.; Li, Z. Forest cover change detection method using bi-temporal GF-1 multi-spectral data. In Proceedings of the 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Beijing, China, 10 July 2016.
22. 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.
23. Li, M.; Li, C.; Huai, H.; Hu, H.; Shi, S.; Quan, B. The best phase selection for monitoring winter wheat sowing time using remote sensing data based on simulated spectral data. *J. Trit. Crop.* **2015**, *35*, 1148–1154.
24. Library, U. Towards Good Practices for Recognition & Detection. Available online: http://imagenet.org/challenges/talks/2016/Hikvision_at_ImageNet_2016.pdf (accessed on 25 December 2018).
25. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. *Adv. Neural Inf. Proc. Syst.* **2012**, *25*, 1097–1105. [[CrossRef](#)]
26. 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 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015.
27. Iandola, F.N.; Han, S.; Moskewicz, M.W.; Ashraf, K.; Dally, W.J.; Keutzer, K. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB model size. *arXiv* **2016**, arXiv:1602.07360.
28. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. Available online: https://www.cv-foundation.org/openaccess/content_cvpr_2016/papers/He_Deep_Residual_Learning_CVPR_2016_paper.pdf (accessed on 26 December 2018).
29. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv* **2014**, arXiv:1409.1556.
30. Dong, Y.W.; Yu, J. Light-weight convolutional neural network slimnet based on squeezeNet. *Comput. Appl. Softw.* **2018**, *35*, 226–232.
31. Tian, T.; Li, C.; Xu, J.; Ma, J. Urban area detection in very high resolution remote sensing images using deep convolutional neural networks. *Sensors* **2018**, *18*, 904. [[CrossRef](#)]
32. JRC. The Feasibility of Detecting Trees Affected by the Pine Wood Nematode Using Remote Sensing. Available online: <https://publications.jrc.ec.europa.eu/repository/bitstream/JRC95972/lb-na-27290-en-n%20.pdf> (accessed on 13 December 2018).
33. Lee, S.; Cho, H.; Lee, W. Detection of the pine trees damaged by pine wilt disease using high resolution satellite and airborne optical imagery. *Korean J. Remote Sens.* **2007**, *23*, 409–420.
34. Hu, G.; Yin, C.; Wan, M.; Zhang, Y.; Fang, Y. Recognition of diseased Pinus trees in UAV images using deep learning and AdaBoost classifier. *Biosyst. Eng.* **2020**, *194*, 138–151. [[CrossRef](#)]
35. Sun, Y.; Zhou, Y.; Yuan, M.; Liu, W.; Luo, Y.; Zong, S. UAV real-time monitoring for forest pest based on deep learning. *Trans. Chin. Soc. Agric. Eng.* **2018**, *34*, 74–81.