PINE WOOD NEMATODE DISEASE WOOD EXTRACTION USING TRIPLESAT SATELLITE IMAGES BASED ON RICHER CONVOLUTIONAL FEATURES

Dongyang Yan (1), Ting Chen (1), Pengfei Yan (1), Wenwen Qi (1), Yuexia Qiao (1), Jianjun He (1), Chong Chen(2)

[1] Twenty First Century Aerospace Technology Co. Ltd, Beijing, 100096, China

[2] Twenty First Century Aerospace Technology (Asia) Pte. Ltd, 61 Science Park Road, #05-17 The Galen, Singapore Science Park II, Singapore 117525

Email: yandy@21at.com.cn; chenting@21at.com.cn; yanpengfei@21at.com.cn; qiww@21at.com.cn; qiaoyx@21at.com.cn; hejj@21at.com.cn; chenchong@21at.sg;

KEYWORDS: Triplesat Satellite Images, Richer Convolutional Features, Pine Wood Nematode, Disease Wood Extraction.

ABSTRACT: Pine wood nematode disease is an important quarantine forestry pest in China, which threatens pine forests and forest ecosystems worldwide and causes serious economic losses. The traditional extraction method mainly analyzes spectral feature of the disease wood, with low accuracy and limited applicable regions. In this paper, we propose a deep learning method to extract the disease wood with pine wood nematode, which is based on TripleSat satellite images and Richer convolutional features (RCF) network. Due to the RCF's image-to-image network structure and deep supervision of side-output, the method can realizes multi-scale and multi-level disease wood feature fusion in order to make a full use of the information from a low level to a high level and guides the learning of the correct disease wood information. The study area is located in Taian city, Shandong Province, in China. First, we produce training sample dataset of the study region using TripleSat satellite images and ground truth data of the pine wood nematode disease wood. Then, this paper trains the RCF network based on Caffe using the training sample dataset. Lastly, the predicting results are processed using OTSU method to segment the probabilities map and get disease wood classification map. The Recovery rate and False alarm rate of our method are 97.9% and 34.3%. The compared experiment results show that our method has a better performance in classification and object position precision than machine learning algorithm, such as SVM. The method can quickly monitor and intelligently identify susceptible pine trees in a wide range in practical application and can provide accurate extraction results for decision maker in order to prevent and control pine wood nematode disease. So, the study of deep learning in disease wood extraction can be one of future development direction.

1 INTRODUCTION

China is one of the countries with the most serious forest pests in the world. According to statistics, the annual incidence of forest pests in China has exceeded 180 million **mu**, resulting in an annual economic loss of 110 billion **Yuan**. Since the first discovery of pine wood nematode disease in Nanjing in 1982, there have been 18 provincial-level administrative districts infected with epidemics in the country, involving more than 500 county-level administrative regions. Pine plants such as Pinus koraiensis and Larix chinensis have also become new hosts. This is the main afforestation tree species in China. Major forestry projects and ecological security pose enormous threats and greater challenges.

Extracting forest information based on remote sensing data has become one of the important technical approaches in modern forestry, and has been widely used in forestry resource survey, fire monitoring, pest monitoring and other fields. At the end of the last century, researchers began to study the use of remote sensing images to monitor and control pine wood nematode epidemic. Landsat data was first used to monitor forest disasters (Vogelmann, 1988; Nakane, 1992), then aerial photography technology (Qinwen, 1997) and GPS (Jin, 2006) were used for forest disease monitoring. Fractal theory was applied to the application of early hyperspectral detection of pine wood nematode disease (Ge, 2009). However, in the early stage of pine wood nematode disease, the external morphology of trees was difficult to identify, and the progress of initial monitoring using remote sensing images was slow.

In 2006, Hinton proposed the concept of deep learning (DL) (Hinton, 2006), which provides the basis for the revival of DL. After that, deep learning techniques are also widely used in the field of remote sensing, like image classification (Luus, 2015), object detection (Girshick, 2014) and etc. Liu firstly proposed the Richer Convolutional Features (RCF) network (Liu, 2017) which developed from the Holistically Nested Edge Detection (HED) network (Xie, 2015). The RCF model changes the locally handled way of the traditional network. Based on the training and prediction of the entire image, high-level semantic information can be obtained. Then, the RCF network used to extract road (Hong, 2018) from a high spatial resolution remote sensing image and detect the edge of buildings (Lu, 2018). The compared experiment results show that the RCF network has a better performance than other information extraction methods

Despite the huge advantages that deep learning shown in information extraction, few researchers have turned their attention to deep learning and using deep learning to extract pine nematode disease wood in recent years. Due to the special architecture of RCF and its excellent performance in the deep-learning-based information extraction, this paper proposed a new method to extract pine nematode disease wood by fine-tuning the pre-trained RCF network, and the generated RCF model can exclusively extract the disease wood. The rest of this paper is organized as follows. In Section 2, we briefly present study area and datasets. RCF-based Pine nematode disease wood detection model is described in Section 3. Section 4 presents the experiment and contrast results and analyzes the performance of the proposed methods. Finally, the discussion and conclusions are drawn in Sections 5.

2 STUDY AREA AND DATASETS

2.1 Study Area

The study area is divided into two parts: sample collection area (Fig. 1) and model test area (Fig. 2). The sample area is located in the northwest of Quzhou City, Zhejiang Province, close to 29° 7' N, 118° 24' E, and about 500 meters above sea level. The test area is located in the north of Taian City, Shandong Province, close to 36°30′N,117° 20′E, and about 1000 meters above sea level. Both areas are cover of vegetations, which are main of pine wood and are important ecological functional area of China. However, in recent years, with the expansion of the pine wood nematode, pine nematode diseased wood has also been found in both areas. Since year 2018, the two regions have caused nearly 10,000 dead pine trees due to pine wood nematode disease. If strong control measures are not taken, once a large-scale outbreak of pine wood nematode disease is caused, it will directly threaten the safety of pine forest resources and cause devastating ecological disasters. This also brings great challenges to the local forestry sector.

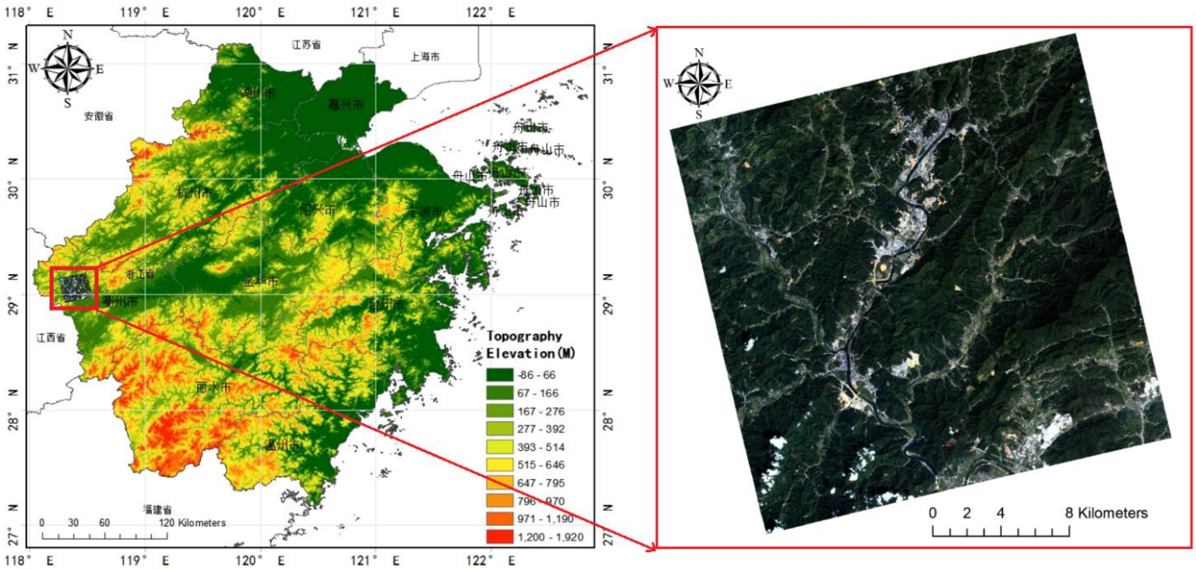


Fig 1. Location (left) and TripleSat satellite images scene (right) of the study area in the Quzhou City, Zhejiang Province

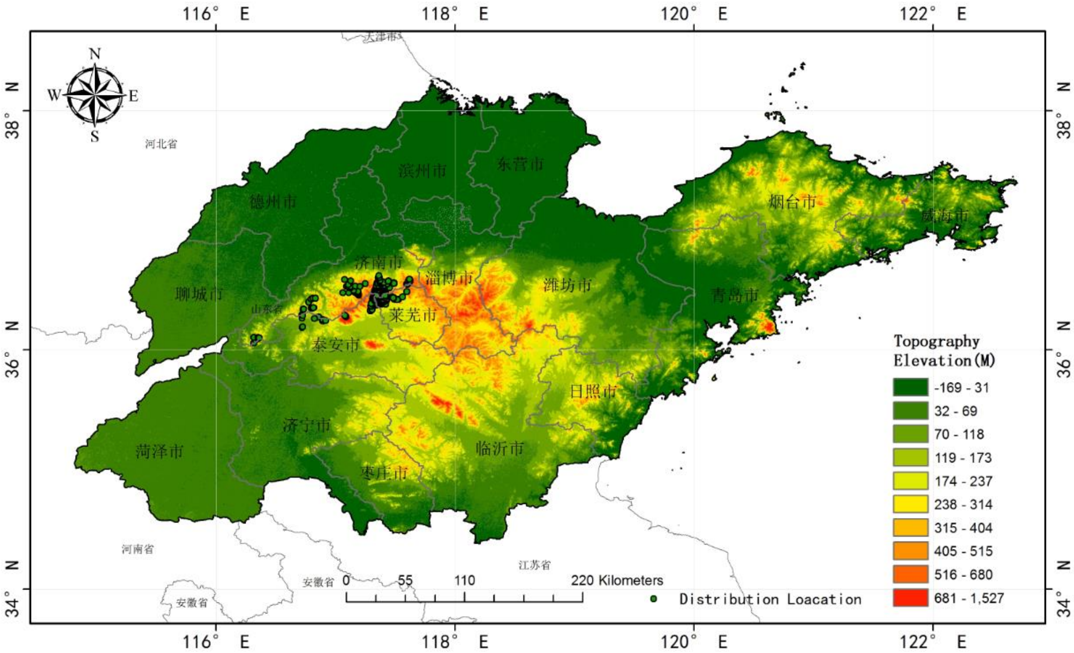
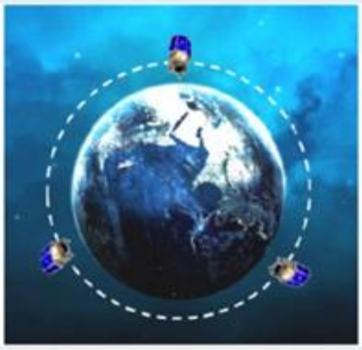


Fig 2. The geographic location of the test area in Taian City, Shandong Province

From July to October, the symptom of the pine wood nematode disease-susceptible strain was concentrated, and the normal forest trees had not changed naturally during this period (no tree interference changes). Therefore, we select several TripleSat satellite images acquired on September 2018 to study.

2.2 Triplesat Satellite Images and Preprocessing

The TripleSat satellite images used in this paper are come from the TripleSat Constellation (Wen, 2017) operated by Twenty First Century Aerospace Technology Co., Ltd., which is a high resolution (<1 m) satellite constellation, consists of three identical satellites, launched on 10 July 2015. The constellation can provide daily targeting capability anywhere on Earth. Both space and ground segments have been designed to efficiently deliver guaranteed timely information. The TripleSat Constellation is the enabler for customers' operational and sustainable geospatial applications. A diagram of TripleSat constellation satellites in the same orbit is shown in Figure 3.



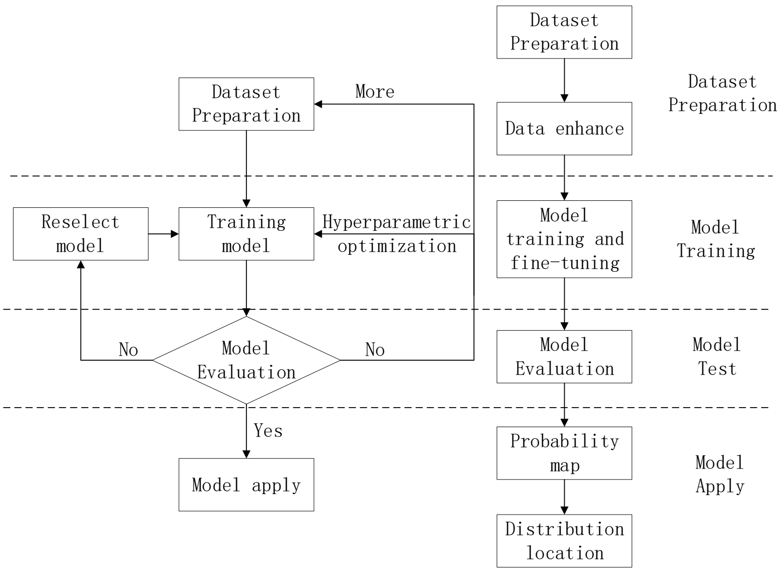
Orbit	SSO: 650km LTAN: 10:30am
Resolution	0.8m PAN; 3.2m MS
Swath Width	24km
Bands	PAN/B/G/R/NIR
Off-pointing	±45°
Revisit	1 Day
Launch Date	10 July 2015

Figure 3. Diagram of Triplesat constellation orbit and specification.

The images before used were subjected to strict geometric and radiation corrections and color normalization to make the difference between the images as small as possible. The original depth of TripleSat satellite images is 16 bits, and the bands are R, G, B, NIR. But limited by models and frameworks, the bit depth was reduced to 8 bits from 16 bits, and only R, G, B bands were used.

3 METHODOLOGY

This study use RCF network to extract pine nematode disease wood from TripleSat satellite images. Similar to the traditional deep learning model application, the workflow of proposed method is mainly divided into four stages. The workflow of traditional deep learning model and proposed method is shown in Figure 4.



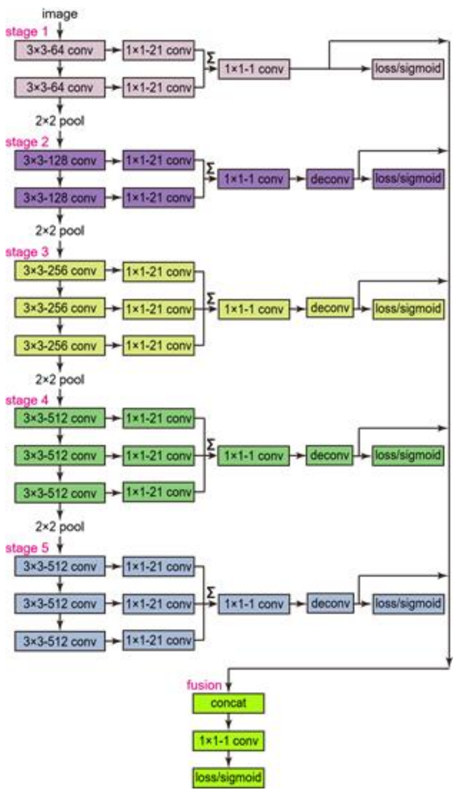


Fig 4. A typical application mode of traditional deep learning model (left) and the workflow of proposed method(right)

Fig 5. The RCF network architecture

3.1 Pine Nematode Disease Wood Extraction

As shown in Figure 4, the method proposed by this paper was divided into four stages:(1):sample dataset preparation. (2): network training and fine-tuning. (3): trained model test. (4): model apply.

Stage 1: The production and processing of datasets are essential for the training of the network and the final prediction. Sample production is a very important step because Sample production accuracy directly images to model convergence effects. Usually, sample production is very time consuming and laborious, and it is very difficult. This study is no exception. Detailed sample production is presented in the experimental section.

Stage 2: We implement our network using the publicly available Caffe (Jia, 2014b) frameworks which is well-known in deep learning community. The VGG16 (Simonyan, 2014) model that is pre-trained on ImageNet (Deng, 2009a) is used to initialize our network. In RCF training, the weights of 1 × 1 conv layer in stage 1-5 are initialized from zero-mean Gaussian distributions with standard deviation 0.01 and the biases are initialized to 0. The weights of 1×1 conv layer in fusion stage are initialized to 0.2 and the biases are initialized to 0. For other SGD hyper-parameters, the global learning rate is set to 1e-6 and will be divided by 10 after every 10k iterations. The momentum and weight decay are set to 0.9 and 0.0002 respectively. We run SGD for 24k iterations totally. All experiments in this paper are finished using a NVIDIA Tesla M40 GPU.

Stage 3: Through the verification data provided by the forestry department, we can evaluate the training effect of the model. Since only the point data is provided, we need to take certain measures to determine whether the point is extracted or not. And use the relevant index for accuracy evaluation. Detailed results were evaluated in the experimental section.

Stage 4: At the last, if the model passes the test, that is, the model's effect is as expected, it can be applied, and we will apply the results for subsequent processing. We use GDAL related techniques to splicing the entire extraction results and vectorizing them, and then calculating the centroid of the pine wood nematode wood objects. The centroid provides a relatively precise position, suggesting that pine wood nematode disease may occur nearby. Through this geographical location, ground forest maintenance personnel can quickly verify and process.

3.2 RCF Network

The RCF network was originally proposed by Liu in 2017. It was optimized based on VGG16 network. Since receptive field sizes of conv layers in VGG16 are different from each other, the RCF network can learn multiscale, including low-level and object level, information that is helpful to information extraction. The input of the RCF network is an RGB image with unlimited size, and the output is a probability map with the same size. An advantage of the RCF network is that it uses the side-output image to image and deep supervision network architecture. In this architecture, the side-output layer is added to incorporate the feature responses from different levels of the primary network stream. Also, deep supervision uses the label (ground truth) to guide the correct side-output of the object information. The RCF network architecture is illustrated in Fig 5.

The CNN autonomously learns features of multiple levels through the convolutional layer and the pooling layer. The hidden network with smaller receptive field can learn some local information of objects in the image. However, as the number of layers increases, the receptive field also becomes larger, which leads to a higher level of information output. Unlike the CNN, the RCF network takes full advantage of the complementary information between different convolutional layers to obtain more accurate results for information extraction. The difference between the RCF network and the traditional neural network lies in: the previous neural networks only use the last layer as the output, and lose many feature details, while the RCF network fuses the convoluted element_wise layers of each stage (convoluted element_wise layers of 2, 3, 4, and 5 stages need to be restored to its original image size by deconvolution) with the same weights to get a fusion output. This special network architecture allows the RCF network to make full use of semantic information and extract detailed information.

4 EXPERIMENT

4.1 Samples Preparation

The importance of the sample in deep learning is self-evident. How to make high-precision samples is especially important. We produced sample data based on the ground monitoring data provided by the staff of the forestry department and the TripleSat satellite images of the corresponding period. The monitoring data provides a geo-referenced coordinate for the outbreak of pine wilt disease, through which we can find and mark the location of the diseased wood. Usually there is one or more diseased trees around a monitoring point. The diseased trees are marked by monitoring points as shown in Figure 6. Figure 7 shows a ground-based verification photograph based on a ground-based monitoring data and a comparison of characterizations on the TripleSat satellite images.

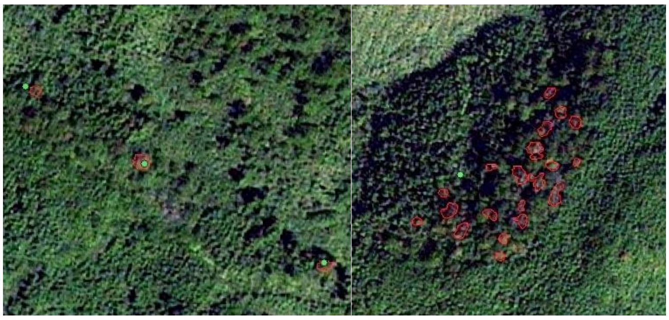


Fig 6. Monitor data points and manually plot ground truth. The green points are the ground monitoring data, and the red regions are samples drawn by hands.

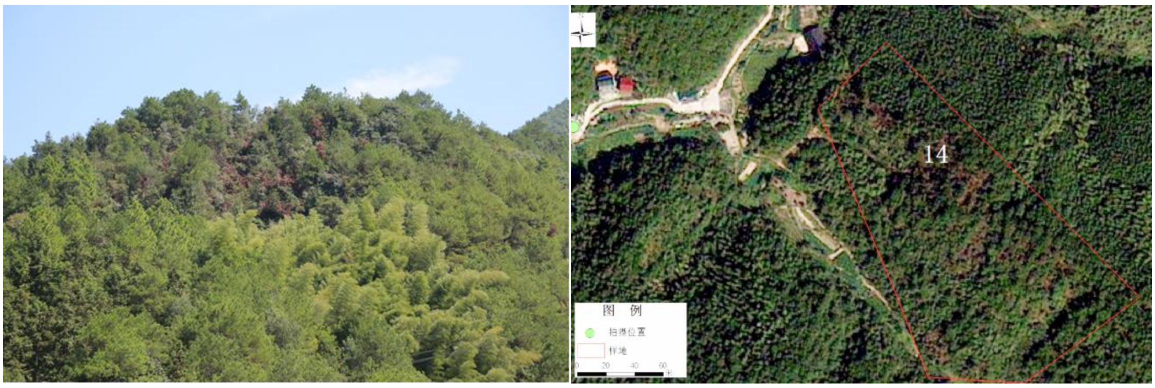


Fig 7. Ground verification site photo (left) and TripleSat satellite image of sample 14 (right)

We have a total of 443 monitoring data in the sample area, and a total of 765 diseased wood objects were identified by these monitoring data. The finished sample vector is then exported to a grayscale image by Arcgis. The gray value of 255 represents the diseased tree, the rest is the background, the gray value is 0. Then the gray image is cropped, a total of 78 samples of size 400x400 are obtained, and some samples are shown in Fig 8.

In deep network training, the model needs to ensure that enough data is entered to avoid overfitting. Therefore, data augmentation has been proved to be a key technology in deep networks. In this study, data augmentation was achieved through vertical flip the training images and labels first and then rotating them in three different angles. Samples of this dataset and data augmentation are shown in Fig 9.

Fig 8. The diagram of samples

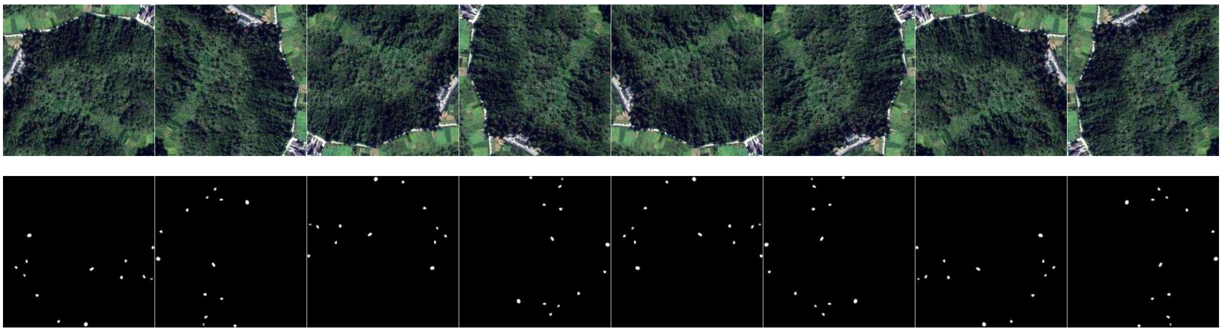


Fig 9. Samples of the dataset and data augmentation through vertical flip and clockwise rotation. The first line to the second line is the cropped image and ground truth augmented through vertical flip and rotating respectively.

4.2 Model Test

After the sample is completed, we use the training parameters described in Section 3 for RCF network training, and the training ends after 34 hours. Then use the trained model to test and get the extracted probability map. The extracted probability maps can be binarized with different thresholds to obtain the final extraction results. Given a probability map, a threshold is needed to produce the binary image. There are two choices to set this threshold. The first one is referred as optimal dataset scale which employs a fixed threshold for all images in the test set. And the second is called optimal image scale which use a method called OTSU to selects an optimal threshold for each image.

Contrast experiments were performed using the SVM algorithm and RF method that came with eCognition. Since the RF method may have something wrong with the parameter setting or other problems, the extraction result was particularly poor. Thus, only the SVM test results were used for comparison. The results of the two methods were compared with the ground monitoring data of Taian area in Shandong Province for accuracy evaluation. The comparisons results obtained by SVM algorithm and RCF netwotk with different binarization threshold with the ground truth values as shown in Fig 10.

4.3 Accuracy Evaluation

we usually use precision, recall, and F-measure as the criteria to evaluate a model. However, this method needs four types according to the combination of the real category and the predicted category: true positive, false positive, true negative, and false negative. Pine nematode disease wood extraction is a binary classification of many-to-many relationships based on objects, that is disease wood objects only true positive or false positive, which is different with pixels classification. So, we use another two indexes, Recovery rate and False alarm rate, an assessment method often

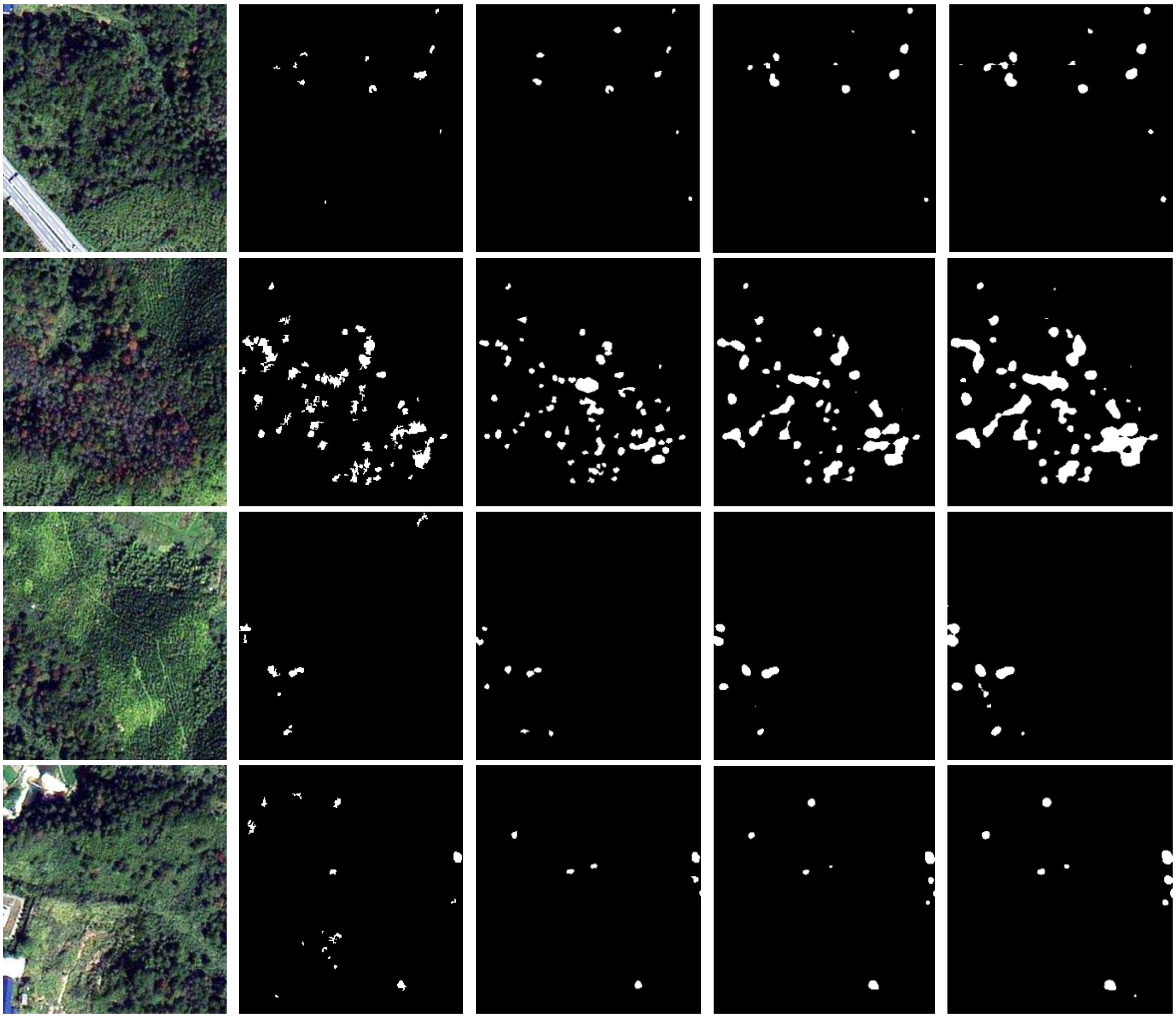


Fig 10. the results comparison of SVM and RCF network with different threshold. From left column to the right column: test images, the results of SVM, the results of RCF network with threshold 100, 150, 180. Through comparison of the results, it can be found that as the threshold increases, the range of results of RCF extraction gradually decreases.

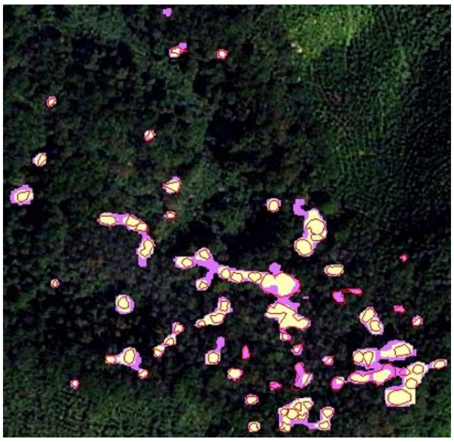


Table 1. The comparison of Recovery rate and False alarm rate

	Recovery rate (%)	False alarm rate (%)
SVM	66.7	33.3
TH100	99.1	56.9
TH150	98.2	45.3
TH180	97.9	34.3

Fig 11. The comparison of the results extracted by different thresholds of binary. The red boundaries are the ground truth drawn by hands, the pink areas are the range of the diseased wood extracted from the threshold value 100, and the yellow areas are the range of the diseased wood segmented by the threshold value of 180. By contrast, raising the threshold can reduce the scope of the extraction, but it can cause missed.

used by forestry workers. Recovery rate is used to evaluate the percentage of disease wood objects that are correctly predicted as the actual disease wood objects. The False alarm rate is the percentage of misclassified objects among all predicted objects. We provided the quantitative comparisons of SVM and RCF network with thresholds 100 (TH100), 150 (TH 150), 180 (TH180).

Table 1 shows the evaluation results between SVM and RCF network with different Thresholds. The method proposed by this paper achieve substantially higher results than SVM. When the segmentation threshold is set to 100, we get a highest recovery rate (99.1%), but also have the highest False alarm rate (56.9%). When we increase the threshold to 180, we get a false alarm rate (34.3%) similar to the SVM (33.3%) but have a much higher recall (for RCF is 97.9%, for SVM is 66.7%). By analyzing Recovery rate and False alarm rate values from Table 1, it is clear that the proposed RCF network model yields more reliable and acceptable results and improves the general classification performance, which is more suitable for the extraction of pine nematode disease wood.

5 DISCUSSION AND CONCLUSIONS

Traditional methods, such as SVM, RF can only use texture, spectral or exponential features, which are not enough to extract pine nematode disease wood information from complex remote sensing images. Different with those methods, the RCF network fully combines the low-level features (e.g., spectral and texture) from the bottom level and the high-level features from the top level to perform precise extraction of pine nematode disease wood. The optimal fits in different scales are obtained via constant learning of pine nematode disease wood features under the guidance of deep supervision. Many shallow and hidden information are learned by multi-scale information fusion and under the guidance of manual labels.

This paper proposes a method for pine wood nematode disease wood extraction using Triplesat satellite images based on the RCF network. The highlights of our work are lists as follows:

(1) For the first time, the RCF network is used to extract pine nematode disease wood with TripleSat satellite images. Compared to the traditional pine nematode disease wood extraction method, the method used in this paper can make full use of high-level semantic information and get a higher accuracy evaluation value.

(2) In this paper, a typical application of TripleSat satellite images was introduced in the case of pine nematode wood extraction, including image processing, sample library construction and training and testing using TripleSat satellite images. As the leading remote sensing satellite platform in China, TripleSat satellite constellation can meet the needs of various tasks and be applied in different information extraction tasks.

(3) In the experimental section, the extraction of pine wood nematode was taken as the research object. Based on the TripleSat satellite images, the key points of rapid processing analysis, sample production, data enhance and intelligent interpretation of the remote sensing image of the pine wood nematode lesion were analyzed. Technology, to achieve intelligent, high-time, high-precision, full coverage monitoring and forecasting of pine wood nematode disease epidemic, providing a theoretical basis for actual production needs. It has an important practical significance in forestry protection.

However, in actually, it is difficult to balance Recall and false alarm rate. A high Recall is usually meaning a high False alarm rate. Reduced accuracy will result in the inability to detect all diseased Pine wood and excessive false alarm rate will increase the workload of the forest ranger. Therefore, how to ensure a lower False alarm rate with a high recall is still a big challenge.

REFERENCES

Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. 2009a. Imagenet: A large-scale hierarchical image database. In 2009 IEEE conference on computer vision and pattern recognition, pp. 248-255.

Hinton, G. E. , Osindero, S. , & Teh, Y. W. . 2006. A fast learning algorithm for deep belief nets. Neural Computation, 18(7), 1527-1554.

Ge, H., Fan, W., Jin, W., & Li, J., 2009. Application of fractal theory in hyperspectral detecting the early stage of pine wood nematode disease (Bursaphelenchus xylophilus) of Pinus massoniana with hyperspectrum. Scientia Silvae Sinicae, 45(6), pp. 68-76.

Girshick, R., Donahue, J., Darrell, T., & Malik, J., 2014. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 580-587.

Hong, Z., Ming, D., Zhou, K., Guo, Y., & Lu, T., 2018. Road extraction from a high spatial resolution remote sensing image based on richer convolutional features. IEEE Access, 6, pp. 46988-47000.

Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., ... & Darrell, T., 2014b. Caffe: Convolutional architecture for fast feature embedding. In Proceedings of the 22nd ACM international conference on Multimedia, pp. 675-678.

Jin, S. H. I. 2006. Application of aerial remote sensing technique in monitoring pine wood nomatode disease. Forest Pest and Disease.

Liu, Y., Cheng, M. M., Hu, X., Wang, K., & Bai, X., 2017. Richer convolutional features for edge detection. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 3000-3009.

Lu, T., Ming, D., Lin, X., Hong, Z., Bai, X., & Fang, J., 2018. Detecting building edges from high spatial resolution remote sensing imagery using richer convolution features network. Remote Sensing, 10(9), 1496.

Luus, F. P., Salmon, B. P., Van den Bergh, F., & Maharaj, B. T. J., 2015. Multiview deep learning for land-use classification. IEEE Geoscience and Remote Sensing Letters, 12(12), pp. 2448-2452.

Nakane, K., & Kimura, Y. 1992. Assessment of pine forest damage by blight based on Landsat TM data and correlation with environmental factors. Ecological Research, 7(1), pp. 9-18.

Qinwen, Y. E. 1997. Application of aerial video techniques to the monitoring of pinewood nematodes. FOREST PEST AND DISEASE, 3.

Simonyan, K., & Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

Vogelmann, J. E., & Rock, B. N. 1988. Assessing forest damage in high-elevation coniferous forests in Vermont and New Hampshire using Thematic Mapper data. Remote Sensing of Environment, 24(2), pp. 227-246.

Wen, Q., He, J., Guan, S., Chen, T., Hu, Y., Wu, W., ... & Yeong, S. 2017. The TripleSat constellation: a new geospatial data service model. Geo-spatial Information Science, 20(2), pp. 163-173.

Xie, S., & Tu, Z., 2015. Holistically-nested edge detection. In Proceedings of the IEEE international conference on computer vision, pp. 1395-1403.