

Article

Detecting Pest-Infested Forest Damage through Multispectral Satellite Imagery and Improved UNet++

Jingzong Zhang, Shijie Cong, Gen Zhang, Yongjun Ma, Yi Zhang and Jianping Huang *

School of Mechanical and Electrical Engineering, Northeast Forestry University, Harbin 150040, China

* Correspondence: jphu@nefu.edu.cn

Abstract: Plant pests are the primary biological threats to agricultural and forestry production as well as forest ecosystem. Monitoring forest-pest damage via satellite images is crucial for the development of prevention and control strategies. Previous studies utilizing deep learning to monitor pest-infested damage in satellite imagery adopted RGB images, while multispectral imagery and vegetation indices were not used. Multispectral images and vegetation indices contain a wealth of useful information for detecting plant health, which can improve the precision of pest damage detection. The aim of the study is to further improve forest-pest infestation area segmentation by combining multispectral, vegetation indices and RGB information into deep learning. We also propose a new image segmentation method based on UNet++ with attention mechanism module for detecting forest damage induced by bark beetle and aspen leaf miner in Sentinel-2 images. The ResNeSt101 is used as the feature extraction backbone, and the attention mechanism scSE module is introduced in the decoding phase for improving the image segmentation results. We used Sentinel-2 imagery to produce a dataset based on forest health damage data gathered by the Ministry of Forests, Lands, Natural Resource Operations and Rural Development (FLNRORD) in British Columbia (BC), Canada, during aerial overview surveys (AOS) in 2020. The dataset contains the 11 original Sentinel-2 bands and 13 vegetation indices. The experimental results confirmed that the significance of vegetation indices and multispectral data in enhancing the segmentation effect. The results demonstrated that the proposed method exhibits better segmentation quality and more accurate quantitative indices with overall accuracy of 85.11%, in comparison with the state-of-the-art pest area segmentation methods.

Keywords: deep learning; attention mechanism; Sentinel-2; semantic segmentation; pest area detecting; vegetation indices



Citation: Zhang, J.; Cong, S.; Zhang, G.; Ma, Y.; Zhang, Y.; Huang, J. Detecting Pest-Infested Forest Damage through Multispectral Satellite Imagery and Improved UNet++. *Sensors* **2022**, *22*, 7440. <https://doi.org/10.3390/s22197440>

Academic Editor: Thomas Udelhoven

Received: 26 August 2022

Accepted: 28 September 2022

Published: 30 September 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

Forests play a crucial part in preserving the natural environment's biodiversity as well as the cyclical flow of energy and mineral elements in the biosphere, which are essential to the ecosystem. The effect of outbreaks of forest pests on healthy forests is severe. Consequently, research on forest pests and their impacts has received considerable attention over the years [1]. It is essential to boost the management of forest pests to minimize the growth of diverse forest pests, improve forest quality, protect the genetic integrity of forest species, maintain ecological balance, and promote coordinated ecological development. The bark beetle, which is a species of the Coleoptera order, *Scolytidae* family, is among the most devastating pests in western North America [2]. It has a significant symptom lag and is highly insidious. In the last few decades, billions of trees have been destroyed by it in the US and Canada [3], causing severe damage to North American forests. The aspen leaf miner is a transcontinental pest of trembling or small tooth aspen, feeding on the poplar leaves' epidermal cells [4]. It has wreaked havoc in the northwestern Canadian and Alaskan forests since the late 1990s, leading to a wide array of consequences [5,6]. Its epidemics can last for a decade or longer [7]. Both pests have a significant impact on timber

production in North America and destabilize the ecological equilibrium of local forests. For the protection of forest ecosystems, their timely and accurate monitoring is crucial.

Traditional monitoring techniques rely on field surveys conducted by personnel. However, traditional manual monitoring methods are time-consuming, making it difficult to meet the application requirements for rapid pest monitoring during production. Field surveys are constrained by the natural environment, and many areas of the field are inaccessible to humans. Large-scale, objective, rapid, and non-destructive are all characteristics of remote sensing technology. For this reason, a growing number of researchers are employing remote sensing technology to monitor forest pests.

Unmanned aerial vehicles (UAVs) have emerged as valuable tools in the monitoring of forest health over large areas. Safonova, et al. [8] extracted areas of UAV imagery that might contain tree crowns and then assessed the degree to which they had been damaged by the Blandford bark beetle. Yu, et al. [9] utilized UAV to monitor pine wilt disease (PWD) in pine forests at an early stage. Many forest areas, however, are unsuitable for the use of UAVs for monitoring due to limitations imposed by sovereignty and natural conditions (e.g., uninhabited areas and uninhabited islands). Although UAVs have the capability to conduct detailed surveys of regional pests, it is difficult for them to conduct a large-scale pest census due to factors such as endurance and relatively limited detection scope.

Satellite remote sensing has a vast observation range, and the cost for the same monitoring area is less than that of UAVs. In addition, it can observe continuously for an extended period of time and generate time-series data. Unlike traditional monitoring techniques and UAV monitoring, satellite remote sensing makes it easier to obtain time-series data without requiring repeated surveys by traveling to the field at specified intervals. In recent years, numerous studies have demonstrated that satellite remote sensing can effectively evaluate the health of insect-infested forests. Dennison, et al. [10] classified mountain pine beetle-damaged canopy cover areas using panorama-sharpened GeoEye-1 images. Meddens and Hicke [3] used multi-temporal Landsat data to analyze and predict tree death due to mountain pine beetle. Using a support vector machine algorithm, GF-2 and Sentinel-2 imagery in combination for detecting individual and stand-scale tree mortality, respectively, that was attributed to red turpentine beetle, by Zhan, et al. [11]. With its multi-temporal, high-resolution, and large-scale range characteristics, satellite remote sensing technology has gradually become an essential instrument for forest-pest surveillance. Its application provides new opportunities for enhancing the location, nature, area range, and frequency assessments of forest pests.

The identification of forest-pest areas in satellite images is a classification task for remote sensing images. With the continuous development of deep learning technology in recent years, numerous researchers have applied it to the field of remote sensing imagery classification [12]. Hong, et al. [13] developed a general multimodal deep learning framework to fuse multimodal data for the classification of surface objects in remote sensing images. Wu, et al. [14] proposed a plug-and-play cross-channel reconstruction module, which enables more efficient information exchange between different remote sensing data and provides new insight into the task of land cover classification. A novel version of graph convolutional networks (GCNs) was proposed by Hong, et al. [15] that outperformed traditional GCNs for the task of hyperspectral remote sensing image classification. Shi, et al. [16] integrated spatial and spectral information in order to reduce information loss during feature extraction and increase the accuracy of land cover classification. The classification of remote sensing images can be solved by image semantic segmentation methods [17]. It is one of the core elements of computer vision research, which analyzes the content of the involved images using various algorithms, allowing the artificial intelligence system to recognize the images' semantics at the pixel level. Image semantic segmentation can be used to extract the areas damaged by bark beetle and aspen leaf miner in satellite images, thereby classifying pest-affected areas in remote sensing images.

However, in this field, multispectral data and deep learning are not combined. RGB bands were utilized in prior studies that employed deep learning to identify pest areas in the

forest. For instance, Kislov, Korznikov, Altman, Vozmishcheva, Krestov, Disney and Cord [18] segmented areas of damaged forest in RGB bands of Worldview-2 and Worldview-3 images using deep convolutional neural network. Using a deep learning approach, Zhou, et al. [19] identified trees damaged by PWD in BJ-2 images (RGB). Multispectral data contain more spectral information than RGB data. Insect damage, unlike physical damage such as storms and floods, can significantly alter the internal chemical composition of trees [20]. For analyses involving the chemical composition of plant tissues, extensive spectral information is essential [21–23]. Using vegetation indices generated from multispectral data, some researchers have obtained favorable results in pest monitoring [24–29]. Bárta, et al. [30] employed a random forest algorithm to analyze seasonal changes in vegetation indices of Norway spruce forests in Sentinel-2 imagery in order to monitor bark beetle infestations, achieving an overall accuracy of 78% in separating healthy and green attack categories. Choi, et al. [31] used Landsat image-based computation of NDMI to detect defoliation caused by the Asian gypsy moth. Kern, et al. [32] used NDVI calculated from MODIS data to detect oak lace bug hazard in oak forests and compared it to field data with greater than 61.1% agreement. Using Fisher discriminant analysis, Xu, et al. [33] analyzed vegetation indices derived from Landsat-7 ETM+ imagery for forest areas affected by *Dendrolimus Punctatus* Walker and obtained a kappa coefficient of 74.77% for samples with severe hazard classes. Consequently, it can be hypothesized that multispectral data and vegetation indices can achieve better results than RGB data when extracting pest-damaged forest areas from satellite images using a deep learning-based method.

Among the most obvious indicators for identifying pests and diseases of plants is the red-edge (680–780 nm). Variations in reflectance on the left side of the red-edge primarily reflect changes in the chlorophyll content of vegetation, whereas variations on the right side of the red-edge reflect changes in the leaf's tissue structure and water content. Therefore, when the chlorophyll content of vegetation or the tissue structure within the leaves and the water content of vegetation change, the red-edge will shift, revealing the health status of the vegetation [34–36]. Existing studies have demonstrated that forest health monitoring models based on red-edge locations perform better than other models [37], with monitoring of vegetation stress in the red-edge and near-infrared bands is earlier compared to other bands [9,38–42]. Among optical satellites, Sentinel-2 is the only one with three bands in the red-edge area, providing information that is highly useful for monitoring the health of plants, and is favored by many forest-related researchers [43–46].

The attention mechanism has been extensively implemented in a variety of deep learning tasks, including natural language processing, image recognition, and speech recognition. Its essence is to selectively select input data so that the model focuses more on the important information in the data and ignores the irrelevant information. Multispectral remote sensing images have rich contextual semantic and spectral information, so it can be speculated that the addition of attention mechanism can enhance the segmentation of pest regions.

The complex surroundings as well as spectral and textural characteristics of pest regions in remotely sensed images pose a formidable challenge to their accurate extraction. This study proposes a new image segmentation method (called RSPR-UNet++) based on UNet++ [47] with an attention mechanism module for detecting forest damage induced by bark beetle and aspen leaf miner in Sentinel-2 images. There are blank spaces in research related to the problem of segmenting pest-infested forest areas, especially those connected to employing deep learning and multispectral satellite imagery. The main objective of this study are: (i) to exploit deep learning for improving forest-pest infestation area segmentation in multispectral satellite imagery; (ii) to illustrate the potential of multispectral data and vegetation indices for pest area segmentation by comparing with RGB images.

2. Study Area and Data

2.1. Study Area

The study area is in the southeast portion of the Skeena region of BC, Canada (53°7'19" N~55°39'12" N, 125°52'3" W~128°22'8" W), covering an area of 48,224.16 km².

Since 2019, the Skeena region has experienced the greatest increase in mountain pine beetle and spruce beetle attacks in BC. The area damaged by aspen leaf miner in the region has nearly doubled since 2019, and the host species most affected by aspen leaf miner is trembling aspen, with a small percentage of other poplars also being damaged [48]. In the study area, the area affected by bark beetle is approximately 7200 km² and the area affected by aspen leaf miner is approximately 4700 km². Figure 1 depicts the study area's geo-location together with an example map of the four Sentinel-2 scenes used in this study.

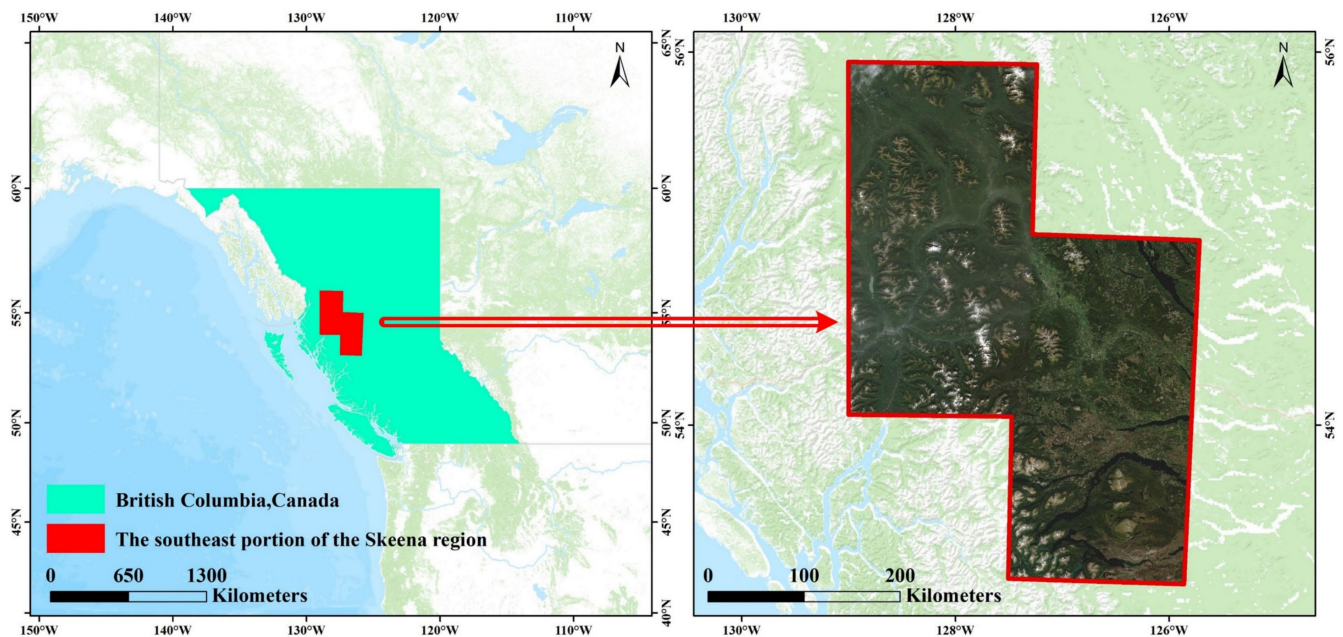


Figure 1. The study area's location and an illustration of the Sentinel-2 images utilized for the experiment.

2.2. Sentinel-2 Data and Preprocessing

2.2.1. Sentinel-2 Data

Sentinel-2 comprises two satellites, 2A and 2B, with one satellite having a 10 day revisit period and two complementary satellites having a 5 day revisit period. The Sentinel-2 data were downloaded from the U.S. Geological Survey's data download site. We used Sentinel-2 data from 8 and 16 September 2020, to be as close as possible to the dates the Ministry of Forests, Lands, Natural Resource Operations and Rural Development (FLNRORD) conducted the 2020 aerial overview surveys (AOS) in the study area (late August to mid-September 2020) and to minimize cloud cover. Table 1 displays the Sentinel-2 band divisions and resolution for each band. The European Space Agency's website provides additional information about the data characteristics of Sentinel-2 imagery.

Table 1. The features of Sentinel-2 imagery.

Band	Band Name	Resolution (m)
B1	Coastal aerosol	60
B2	Blue	10
B3	Green	10
B4	Red	10
B5	Vegetation Red Edge 1	20
B6	Vegetation Red Edge 2	20
B7	Vegetation Red Edge 3	20
B8	NIR	10
B8A	Narrow NIR	20
B9	Water vapor	60
B10	SWIR-Cirrus	60
B11	SWIR 1	20
B12	SWIR 2	20

2.2.2. Data Preprocessing

The Sentinel-2 images obtained are multispectral Level 1C data, which are orthorectified and geometrically fine-corrected atmospheric-apparent reflectance products with no atmospheric correction. We utilized Sen2Cor to execute atmospheric correction on the downloaded Sentinel-2 1C level images, which resulted in 2A level images, with the B1 and B10 bands disappearing after correction.

As shown in Table 1, Sentinel-2 has three different resolutions for its bands: 10 m, 20 m, and 60 m. Many subsequent operations will not be possible if they are not unified to the same resolution. Consequently, following atmospheric correction, we resampled the images using the software Snap, resampling the bands with 20 m and 60 m resolutions to 10 m.

After resampling, we calculated NDWI, DWSI, NGRDI, RDI, GLI, NDRE2, PBI, NDVI, GNDVI, CIG, CVI, NDRE3, and DRS. Based on the common degree of vegetation indices in remote sensing image analysis and existing research results in pest research, 13 vegetation indices were calculated [49]. These vegetation indices were calculated using the bands of Sentinel-2 and the corresponding formulas; the calculation procedure is detailed in Table 2.

Table 2. Method for calculating vegetation indices.

Vegetation Indices	Calculation Method	Calculation Details in Sentinel-2
NDWI	$\frac{(\text{Narrow NIR}) - (\text{SWIR 1})}{(\text{Narrow NIR}) + (\text{SWIR 1})}$	$\frac{B8A - B11}{B8A + B11}$
DWSI	$\frac{\text{NIR} + \text{Green}}{\text{Red} + (\text{SWIR 1})}$	$\frac{B8 + B3}{B4 + B11}$
NGRDI	$\frac{\text{Green} - \text{Red}}{\text{Green} + \text{Red}}$	$\frac{B3 - B4}{B3 + B4}$
RDI	$\frac{\text{SWIR 2}}{\text{Narrow NIR}}$	$\frac{B12}{B8A}$
GLI	$\frac{2 * \text{Green} - \text{Blue} - \text{Red}}{2 * \text{Green} + \text{Blue} + \text{Red}}$	$\frac{2 * B3 - B2 - B4}{2 * B3 + B2 + B4}$
NDRE2	$\frac{(\text{Vegetation Red Edge 3}) - (\text{Vegetation Red Edge 1})}{(\text{Vegetation Red Edge 3}) + (\text{Vegetation Red Edge 1})}$	$\frac{B7 - B5}{B7 + B5}$
PBI	$\frac{\text{NIR}}{\text{Green}}$	$\frac{B8}{B3}$
NDVI	$\frac{(\text{Narrow NIR}) - \text{Red}}{(\text{Narrow NIR}) + \text{Red}}$	$\frac{B8A - B4}{B8A + B4}$
GNDVI	$\frac{(\text{Narrow NIR}) - \text{Green}}{(\text{Narrow NIR}) + \text{Green}}$	$\frac{B8A - B3}{B8A + B3}$
CIG	$\frac{\text{Narrow NIR} - 1}{\text{Green}}$	$\frac{B8A}{B3} - 1$
CVI	$\frac{(\text{Narrow NIR}) * (\text{Vegetation Red Edge 1})}{(\text{Green})^2}$	$\frac{B8A * B5}{(B3)^2}$
NDRE3	$\frac{(\text{Narrow NIR}) - (\text{Vegetation Red Edge 3})}{(\text{Narrow NIR}) + (\text{Vegetation Red Edge 3})}$	$\frac{B8A - B7}{B8A + B7}$
DRS	$\sqrt{(\text{Red})^2 + (\text{SWIR 2})^2}$	$\sqrt{(B4)^2 + (B12)^2}$

To avoid the effect of any extreme values that could be outliers, each vegetation index was ordered from smallest to largest and then linearly stretched to 0–255 (8-bit), with the value at 2.5% as the minimum and the value at 97.5% as the maximum. The original Sentinel-2 (12-bit) bands are directly scaled to 0–255. Each of these 13 vegetation index images was then added to the original image as a band.

2.3. Dataset

British Columbia, Canada, is largely forested, and forestry is a vital economic pillar for the province. As a result, the B.C. Ministry of FLNRORD conducts an annual AOS of the province's forests in order to monitor their current state of health. The pest area labels for this study were determined using provincial fieldwork data from 2020.

Using Arcmap, the shapefile-formatted pest area labels are converted to raster form, which is the same size as the corresponding Sentinel-2 image. Using 1 of the 4 scenes of the Sentinel-2 image as an example, Figure 2 depicts the corresponding raster label of the image.

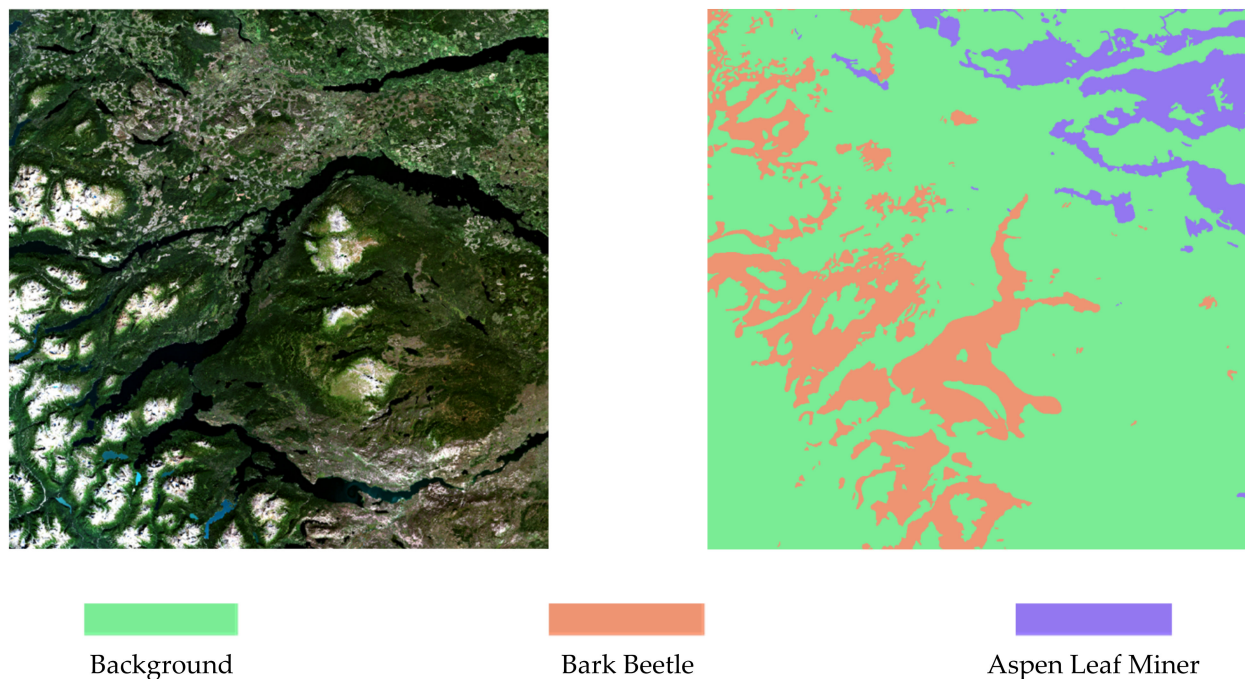


Figure 2. The raster label corresponding to the Sentinel-2 image.

The remote sensing image and the corresponding label is cropped into several image blocks and then is inputted into the deep learning network to avoid memory overflow. In this study, the Sentinel-2 images and labels of the study area were cropped in a regular grid pattern to produce 4984 3D cubes and 4984 labels corresponding to 3D cubes. The length and width of the 3D cube are 256 pixels. The number of channels, i.e., the total number of original bands and vegetation indices, is 24. A total of 20% of the dataset was allocated to the test set, 64% to the training set, with 16% allocated to the validation set.

3. Methodology

3.1. Model Architecture

Figure 3 shows the proposed model's overall structure (called RSPR-UNet++). It consists of an encoder sub-network and a decoder sub-network. $X^{i,0}$ ($i = 0, 1, 2, 3, 4$) stands for ResNeSt Layer and $X^{i,j}$ ($j \neq 0$) stands for convolutional layer of 3×3 size. A 1×1 sized convolutional layer is added after each $X^{0,j}$ ($j \neq 0$). Following the 1×1 convolutional layer is the sigmoid activation function. This design improves the gradient propagation by directly connecting $X^{0,j}$ ($j = 1, 2, 3$) to the final output while supervising the output of the dense convolutional block of each branch. After experimental adjustment, the channels' number of the feature-maps output from $X^{0,j}$, $X^{1,j}$, $X^{2,j}$, $X^{3,j}$ and $X^{4,j}$ ($j = 0, 1, 2, 3, 4$) are set to 16, 32, 64, 128 and 256 (the original UNet++ is 32, 64, 128, 256 and 512) in turn.

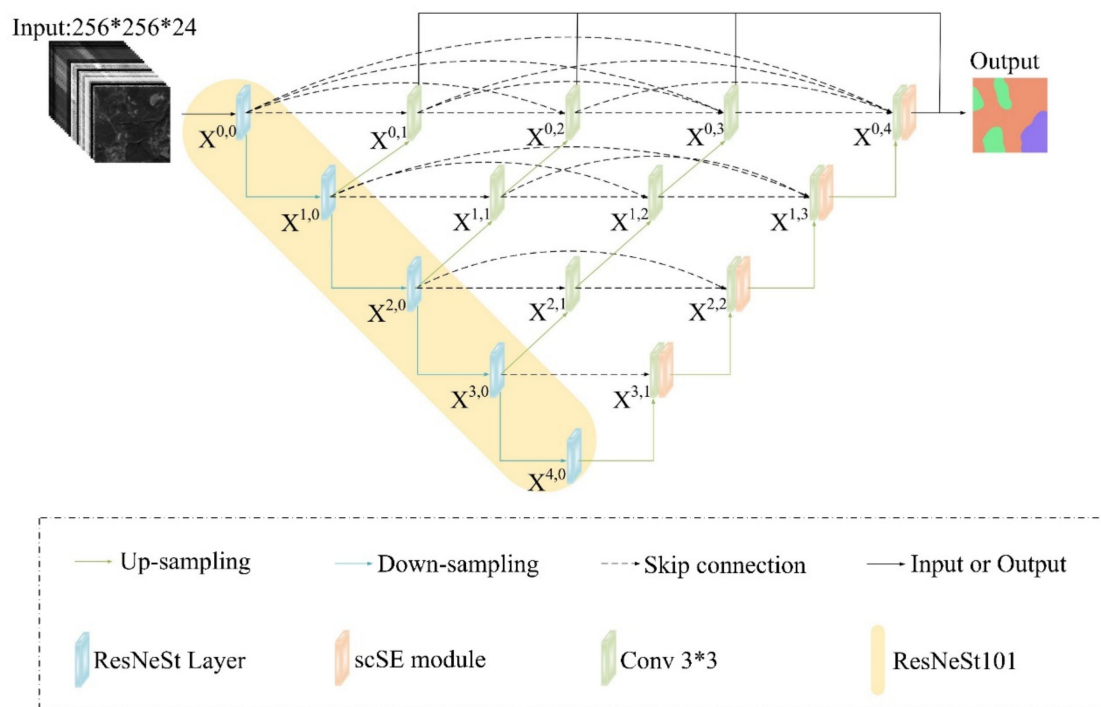


Figure 3. The structure of RSPR-UNet++.

Based on the original UNet++ network, we use ResNeSt101 [50] as the feature extraction backbone network. ResNeSt introduces Split-Attention module while retaining the Residual Network's structure (ResNet [51]), and stacks the Split-Attention module. The structure of the Split-Attention module in the model is shown in Figure 4. H , W , and C represent the height, width, and number of channels, respectively, of the input feature-map X . The module is a computational unit that consists of two parts: feature-map splits and split-attention operations. It divides the extracted feature-map into 2 feature-map splits, performs a series of transformations on each split, then fuses the weighted feature-maps after the transformations. Finally, as with the standard residual block, the final output feature-map Y is obtained by connecting with feature-map X and feature-map V using a shortcut connection. It realizes the information interaction between the feature-map splits, thus improving the model's feature extraction capability. Using ResNeSt as the feature extraction backbone network, features with different weights can be obtained from different splits of feature-map, and richer feature information of the infested area can be extracted, making the segmentation results more accurate.

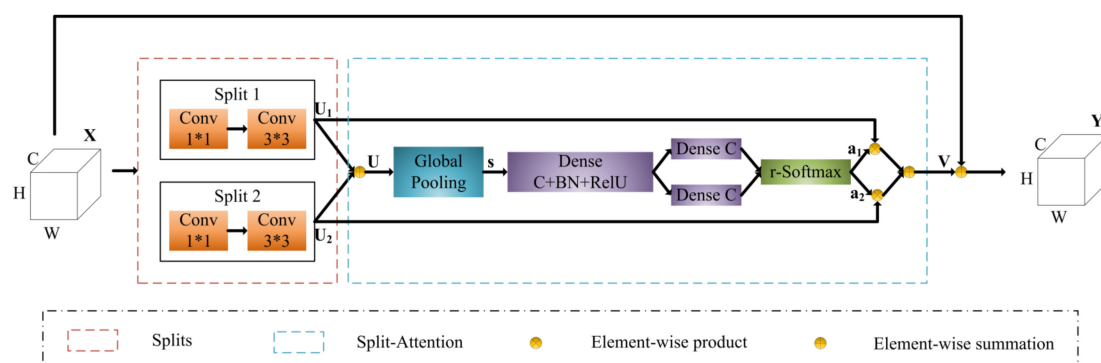


Figure 4. Split-Attention Block.

To further focus the model's attention on the infested region and thus extract its more detailed features, we add the scSE [52] attention mechanism module at the end of each

nested skip pathway in the decoding stage. Figure 5 depicts the specific structure of scSE, which consists of a combination of sSE module and cSE module. The scSE module calibrates and excites the spatial and channel features of the image in the application of pest region segmentation, reducing the influence of redundant features and effectively improving the model's ability to automatically learn the image's effective features, further enhancing the model's segmentation accuracy.

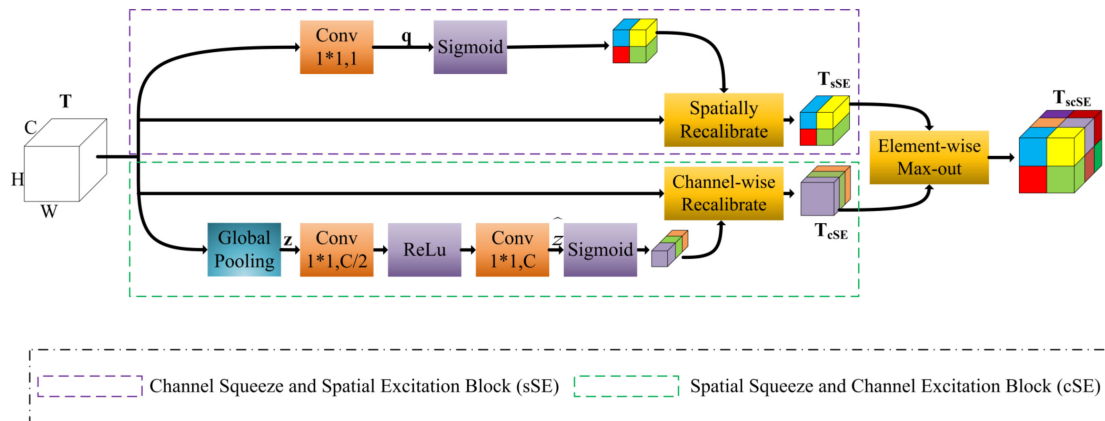


Figure 5. Spatial and Channel Squeeze and Excitation Block(scSE).

Table 3 illustrates the characteristics of other common semantic segmentation models. RSPR-UNet++ combines attention mechanism and encoder–decoder structure, which is more advanced than them.

Table 3. Other common semantic segmentation models and their characteristics.

Model	Characteristics	Reference
UNet	The architecture contains 2 paths (contraction path and symmetric expanding path). It is an end-to-end fully convolutional network (FCN).	[53]
DeeplabV3+	The spatial pyramid pooling module and the encoder–decoder structure were combined. The depthwise separable convolution was applied to both the Atrous Spatial Pyramid Pooling and decoder modules.	[54]
Feature Pyramid Networks (FPN)	Developed a top-down architecture with lateral connections for building high-level semantic feature maps at all scales.	[55]
Pyramid Attention Network (PAN)	Exploited the impact of global contextual information in semantic segmentation.	[56]
UNet++	The architecture is an encoder–decoder network where the encoder and decoder sub-networks are connected through a series of nested, dense skip pathways. It optimizes the topology of UNet and is an improved version of the UNet network structure.	[47]

The flow of making the dataset as well as the process of training, validating, and testing the model are shown in Figure 6.

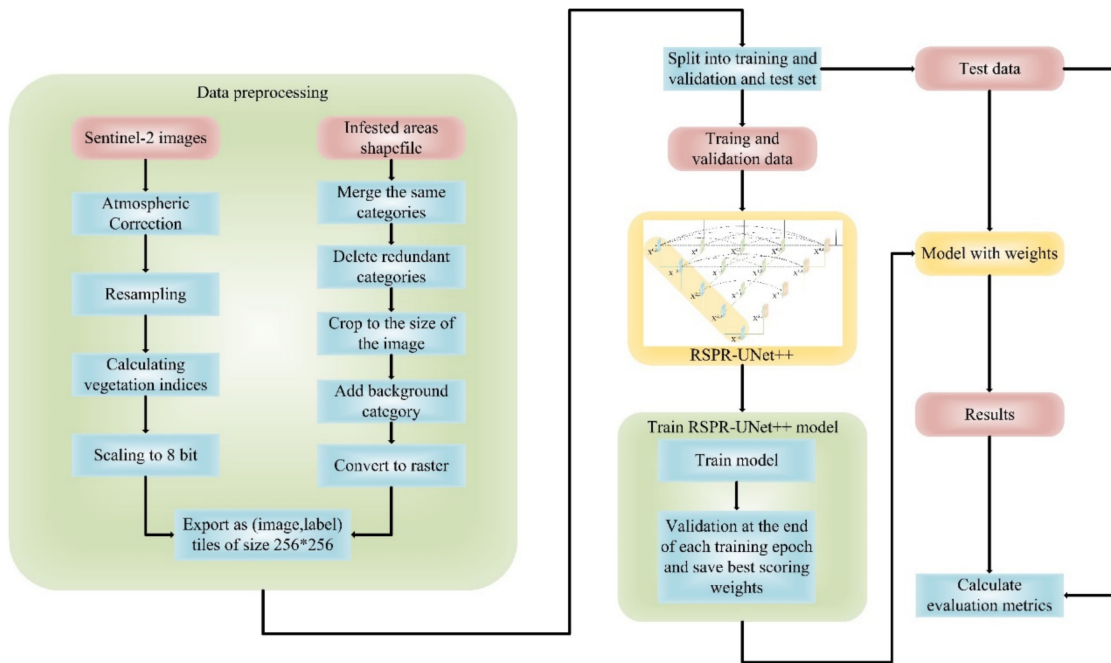


Figure 6. A flow chart of the process of creating the dataset together with the process of training, validating, and testing the model.

3.2. Loss Function

It is common in remote sensing images' segmentation that the number of foreground pixels and background pixels differ significantly, i.e., there is a sample imbalance between the segmented object and the background. The Dice Loss function is more suitable for the case of sample imbalance, the specific formula is as follows:

$$L_{\text{dice}} = 1 - \frac{2|X \cap Y|}{|X| + |Y|} \quad (1)$$

where: X denotes the tensor of the true segmentation label of the image, and Y denotes the tensor of the image segmentation result predicted by the model. $|X \cap Y|$ denotes the sum obtained by element-by-element summation for the result of the dot product of X and Y . $|X|$ denotes the sum obtained by adding X element by element, and $|Y|$ denotes the sum obtained by adding Y element by element.

In extreme cases, however, the Dice Loss function may cause the gradient values to be very high, which negatively impacts back propagation and renders model training unstable. Soft Cross Entropy Loss function does cross entropy calculation with the predicted values after label smoothing of the labeled values, which can improve the model's generalization to some extent, the specific formula is as follows:

$$L_{\text{sce}} = - \sum_{i=1}^n p(x_i) \log(q(x_i)) \quad (2)$$

$$x_{i(\text{one-hot})} = \begin{cases} 1, & x_i = \text{target} \\ 0, & x_i \neq \text{target} \end{cases} \quad (3)$$

$$p(x_i) = x_{i(\text{one-hot})} * (1 - \alpha) + \frac{\alpha}{K} \quad (4)$$

where: n denotes the number of pixel points, x_i denotes a pixel point, and $q(x_i)$ denotes the probability that the model predicts x_i to be the category in the labeling. K denotes the total number of categories in the segmentation, and $K = 3$ in this study. α denotes the smooth factor, and $\alpha = 0.1$ in this study.

Consequently, the following equation combines the Dice Loss function with the Soft Cross Entropy Loss function as the experiment's loss function:

$$L = 0.5 * L_{\text{dice}} + 0.5 * L_{\text{sce}} \quad (5)$$

3.3. Model Training

The software environment used for the experiments is: Windows 10 Professional 64-bit system, Pytorch (1.8.0) framework, and Python (3.6). Hardware environment: 32 GB RAM; Intel(R) Core i7-10700, 2.90 GHz; NVIDIA GeForce RTX 2080Ti, 11 GB graphics memory.

The model training process is optimized using the AdamW optimizer, with the batch-size set to 6 and the weight decay set to 10^{-3} , for a total of 300 training epochs. To improve the model's convergence speed, we pretrain the ResNeSt101 network for migration learning using the ImageNet dataset. The gradient descent method may encounter local minima during training, and then the learning rate can be suddenly increased to "jump out" of the local minima and find the path to the global minima. Therefore, we use a cosine annealing strategy to adjust the learning rate. We let T_0 be the epoch of restart, T_{mult} be the factor after restart, and $T_0 = T_0 * T_{\text{mult}}$ after each restart. In this study, the initial $T_0 = 2$, $T_{\text{mult}} = 2$, the initial learning rate is 10^{-4} , and the minimum learning rate is 10^{-5} . Throughout the training process, the model with the highest mean Intersection-over-Union (mIoU) on the validation set is chosen as the final output model.

The loss values during the training of the model are shown in Figure 7. It can be seen that the loss curve decreases rapidly at the beginning, then gradually converges to about 0.56 after about 225 iterations.

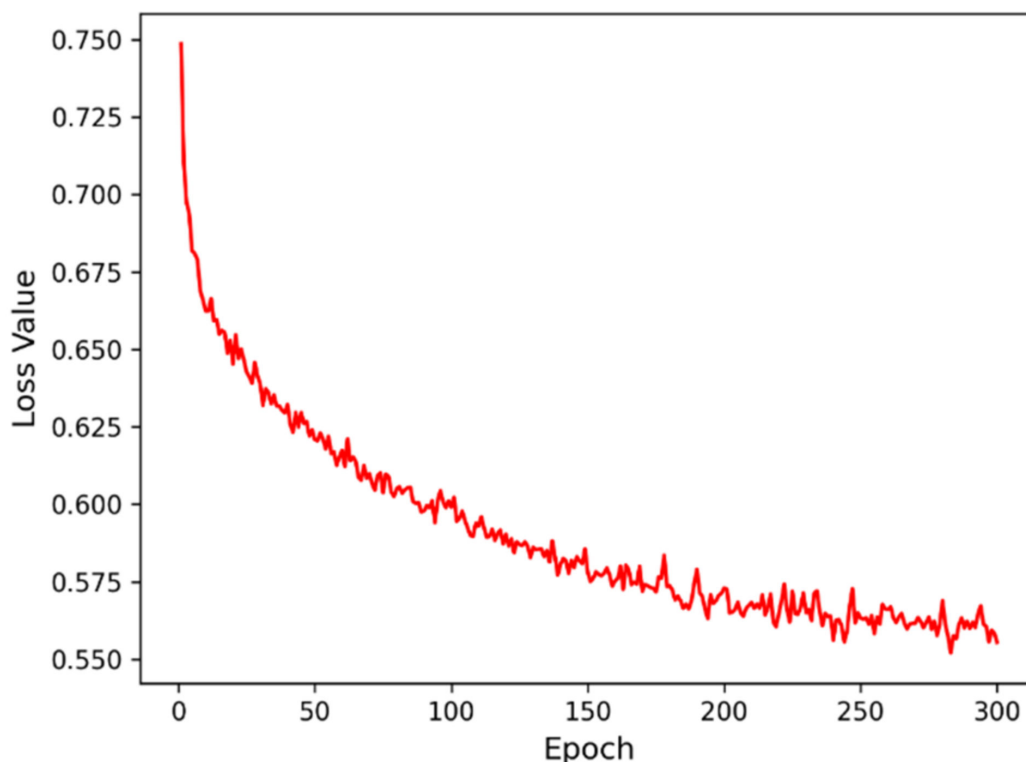


Figure 7. Loss value corresponding to different iterations.

3.4. Evaluation Metrics

To demonstrate the efficacy of RSPR-UNet++ and compare each model's performance on the test set, we quantitatively evaluated each model's performance using accuracy, precision, recall, F1-score, Intersection-over-Union (IoU), mIoU, and frequency weighted Intersection-over-Union (FWIoU). These evaluation metrics are calculated as follows:

The accuracy is calculated as follows:

$$\text{accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (6)$$

where, TP (True Positive) denotes the number of true positive samples; FP (False Positive) denotes false positive samples; TN (True Negative) denotes true negative samples; and FN (False Negative) denotes false negative samples.

The precision and recall are calculated as follows:

$$\text{precision} = \frac{TP}{TP + FP} \quad (7)$$

$$\text{recall} = \frac{TP}{TP + FN} \quad (8)$$

The two metrics, accuracy and recall, sometimes appear contradictory and cannot be high at the same time, so they need to be considered together. The most common composite evaluation metric is the F1-score, which is the harmonic mean of the precision and recall, defined as:

$$F1 = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (9)$$

The F1-score is determined by both precision and recall, when it is higher it indicates a better recognition result.

IoU is the ratio of intersection and union of the actual and predicted category samples, which is calculated as follows:

$$\text{IoU} = \frac{TP}{TP + FN + FP} \quad (10)$$

mIoU is the result of summing the IoU for each category and then averaging:

$$\text{mIoU} = \frac{\sum_N^i \text{IoU}}{N} \quad (11)$$

where N is the number of categories.

FWIoU sets weights according to the frequency of occurrence of each category, and the weights are multiplied by the IoU of each category and summed:

$$\text{FWIoU} = \frac{TP + FN}{TP + FP + TN + FN} * \frac{TP}{TP + FP + FN} \quad (12)$$

3.5. Analyze the Impact of Different Data

To analyze the effects of RGB images, multispectral images, and vegetation indices, we compared and analyzed the results of training and testing of the proposed model using RGB, Sentinel-2's original 11 bands, RGB plus 13 vegetation indices, and Sentinel-2's original 11 bands plus 13 vegetation indices.

In addition, for the purpose of analyzing the influence of the red-edge and the vegetation indices derived from the red-edge, we added 8 vegetation indices related to the red-edge to the original vegetation indices. The new vegetation indices added are shown in Table 4. We compared and analyzed the results of training and testing of the proposed model using all 32 bands and the 18 bands of the 32 bands that are not related to the red edge (The 8 new vegetation indices, Vegetation Red Edge 1, Vegetation Red Edge 2, Vegetation Red Edge 3, NDRE2, NDRE3, and CVI were excluded.).

Table 4. Method for calculating the 8 new vegetation indices related to red edge.

Vegetation Indices	Calculation Method	Calculation Details in Sentinel-2
ND790/670	$\frac{(\text{Vegetation Red Edge 3}) - \text{Red}}{(\text{Vegetation Red Edge 3}) + \text{Red}}$	$\frac{B7 - B4}{B7 + B4}$
NDVI690-710	$\frac{\text{Water vapor} - (\text{Vegetation Red Edge 1})}{\text{Water vapor} + (\text{Vegetation Red Edge 1})}$	$\frac{B9 - B5}{B9 + B5}$
NDRE	$\frac{\text{NIR} - (\text{Vegetation Red Edge 1})}{\text{NIR} + (\text{Vegetation Red Edge 1})}$	$\frac{B8 - B5}{B8 + B5}$
NDVI65	$\frac{(\text{Vegetation Red Edge 2}) - (\text{Vegetation Red Edge 1})}{(\text{Vegetation Red Edge 2}) + (\text{Vegetation Red Edge 1})}$	$\frac{B6 - B5}{B6 + B5}$
GNDVIhyper	$\frac{(\text{Vegetation Red Edge 3}) - \text{Green}}{(\text{Vegetation Red Edge 3}) + \text{Green}}$	$\frac{B7 - B3}{B7 + B3}$
RENDVI1	$\frac{(\text{Vegetation Red Edge 1}) - \text{Red}}{(\text{Vegetation Red Edge 1}) + \text{Red}}$	$\frac{B5 - B4}{B5 + B4}$
RENDVI2	$\frac{(\text{Vegetation Red Edge 2}) - \text{Red}}{(\text{Vegetation Red Edge 2}) + \text{Red}}$	$\frac{B6 - B4}{B6 + B4}$
RI	$\frac{(\text{Vegetation Red Edge 1}) - \text{Green}}{(\text{Vegetation Red Edge 1}) + \text{Green}}$	$\frac{B5 - B3}{B5 + B3}$

4. Results

4.1. Comparison of Predicted Results

4.1.1. Comparison between Different Models

To verify RSPR-UNet++'s performance, in Table 5 we compare the commonly used semantic segmentation models listed in Table 3. All models were trained and tested using all 24 bands. It can be seen that RSPR-UNet++ has 5.98%, 4.41%, and 2.9% higher mIoU, FWIoU, and accuracy, respectively, compared with UNet; 4.73%, 3.75%, and 2.6% higher mIoU, FWIoU, and accuracy, respectively, compared with DeeplabV3+; mIoU, FWIoU, and accuracy, respectively, compared with FPN was 5.31%, 3.9%, and 2.59% higher; mIoU, FWIoU, and accuracy was 4.98%, 3.63%, and 2.4% higher, respectively, compared with PAN; mIoU, FWIoU, and accuracy was 4.64%, 3.44%, and 2.29% higher, respectively, compared with UNet++. A more accurate segmentation result was achieved by RSPR-UNet++ than other models.

Table 5. The experimental outcomes of RSPR-UNet++ and other methods. BG represents the background, BB the bark beetle, and ALM the aspen leaf miner.

Model	Category	Precision (%)	Recall (%)	F1 (%)	IoU (%)	mIoU (%)	FWIoU (%)	Accuracy (%)
UNet	BG	85.44	88.63	87.00	77.00	63.84	70.09	82.21
	BB	76.64	69.74	73.03	57.52			
	ALM	74.03	71.26	72.62	57.01			
FPN	BG	86.33	87.93	87.12	77.18	64.51	70.60	82.52
	BB	75.09	74.12	74.60	59.49			
	ALM	75.08	70.11	72.51	56.87			
PAN	BG	86.39	88.34	87.36	77.55	64.84	70.87	82.71
	BB	75.00	73.05	74.01	58.75			
	ALM	76.15	71.21	73.59	58.22			
DeeplabV3+	BG	87.63	86.22	86.92	76.86	65.09	70.75	82.51
	BB	74.85	75.81	75.33	60.42			
	ALM	71.49	75.44	73.41	57.99			
UNet++	BG	86.81	87.81	87.31	77.47	65.18	71.06	82.82
	BB	75.34	74.71	75.03	60.03			
	ALM	75.00	71.93	73.44	58.02			
RSPR-UNet++ without scSE	BG	89.61	87.06	88.32	79.08	68.83	73.76	84.61
	BB	75.60	82.15	78.74	64.94			
	ALM	76.83	76.98	76.90	62.47			
RSPR-UNet++	BG	89.92	87.53	88.71	79.70	69.82	74.50	85.11
	BB	78.10	79.52	78.81	65.02			
	ALM	75.16	82.33	78.58	64.72			

To further investigate the impact of the ResNeSt and scSE modules on the proposed model. Table 5 compares the segmentation effect of RSPR-UNet++ with and without the scSE module. It can be seen that when the scSE module is removed, the mIoU falls by 0.99%, the FWIoU falls by 0.74%, and the accuracy falls by 0.5%. Compared to UNet++, RSPR-UNet++ with the scSE module removed improves mIoU by 3.65%, FWIoU by 2.2%, and accuracy by 1.79%. Both the ResNeSt and scSE modules have a positive effect on the model, as is evident.

To compare the segmentation effects of RSPR-UNet++ and other models on the full-size Sentinel-2 image, we used a scene of Sentinel-2 image from the Skeena region with the imaging date of 28 August 2019. After preprocessing, it is cropped in regular grid fashion into a number of 3D cubes of size $256 \times 256 \times 24$. Predictions are made for each cube, and then the predictions are stitched together sequentially to obtain the predicted result for the entire Sentinel-2 image. Figure 8 depicts the complete procedure. The segmentation effect of RSPR-UNet++ and other models on the full-size Sentinel-2 image are shown in Figure 9. Table 6 compares these models' overall accuracy for the segmentation results of this Sentinel-2 image. In comparison to other models, RSPR-UNet++ has a superior segmentation effect.

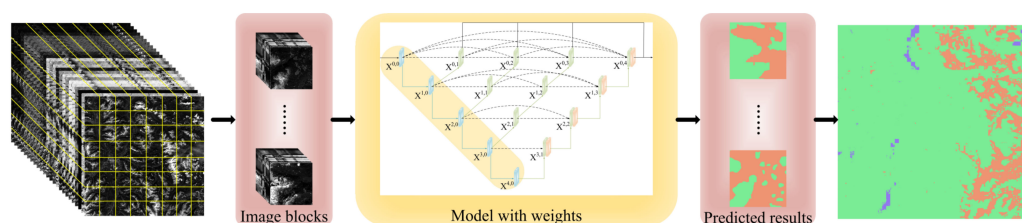


Figure 8. The process of the prediction of the entire Sentinel-2 image.

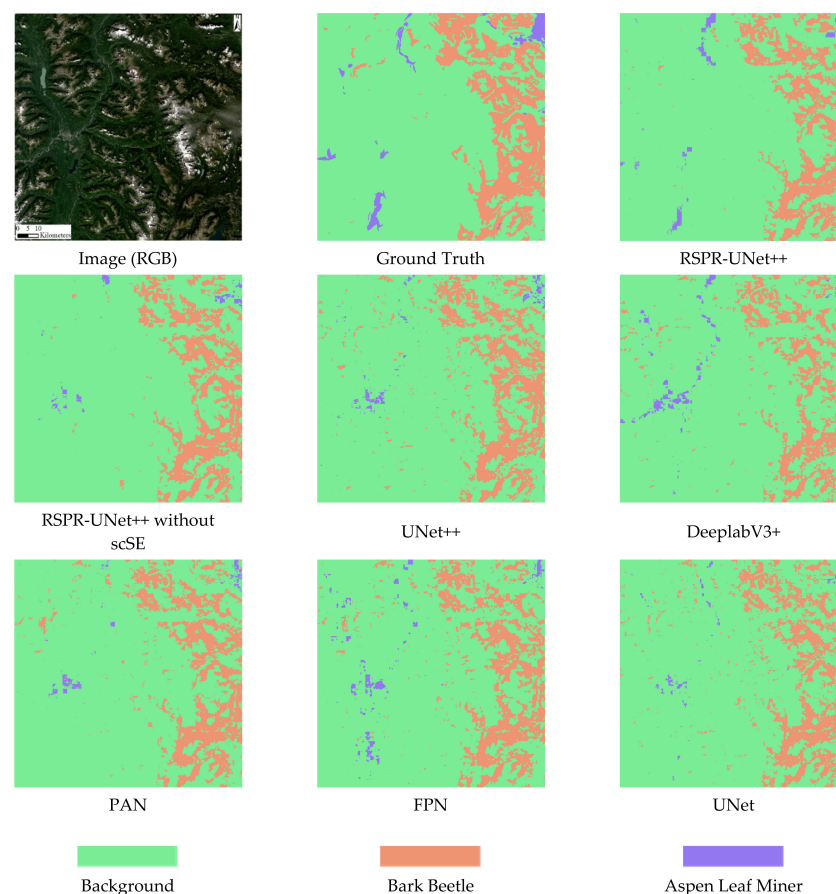


Figure 9. The segmentation effect of the whole Sentinel-2 image.

Table 6. The overall accuracy for the segmentation results of the Sentinel-2 image in Figure 9.

Model	Accuracy (%)
UNet	86.79
FPN	86.31
PAN	87.40
DeeplabV3+	85.87
UNet++	87.49
RSPR-UNet++ without scSE	88.29
RSPR-UNet++	89.10

4.1.2. Comparison between RGB Data, Multispectral Data, and Vegetation Indices

Figure 10 and Table 7 shows the experimental results of using RGB, Sentinel-2's original 11 bands, and Sentinel-2's original 11 bands plus 13 vegetation indices to train and test the proposed model. Compared with using RGB, when training with the original 11 bands of Sentinel-2, mIoU improved by 1.56%, FWIoU improved by 1.27%, and accuracy improved by 0.82%. It is clearly seen that, when training with all bands, mIoU improves by 4.06%, FWIoU improves by 3.47%, and accuracy improves by 2.48% compared to the original 11 bands of Sentinel-2.

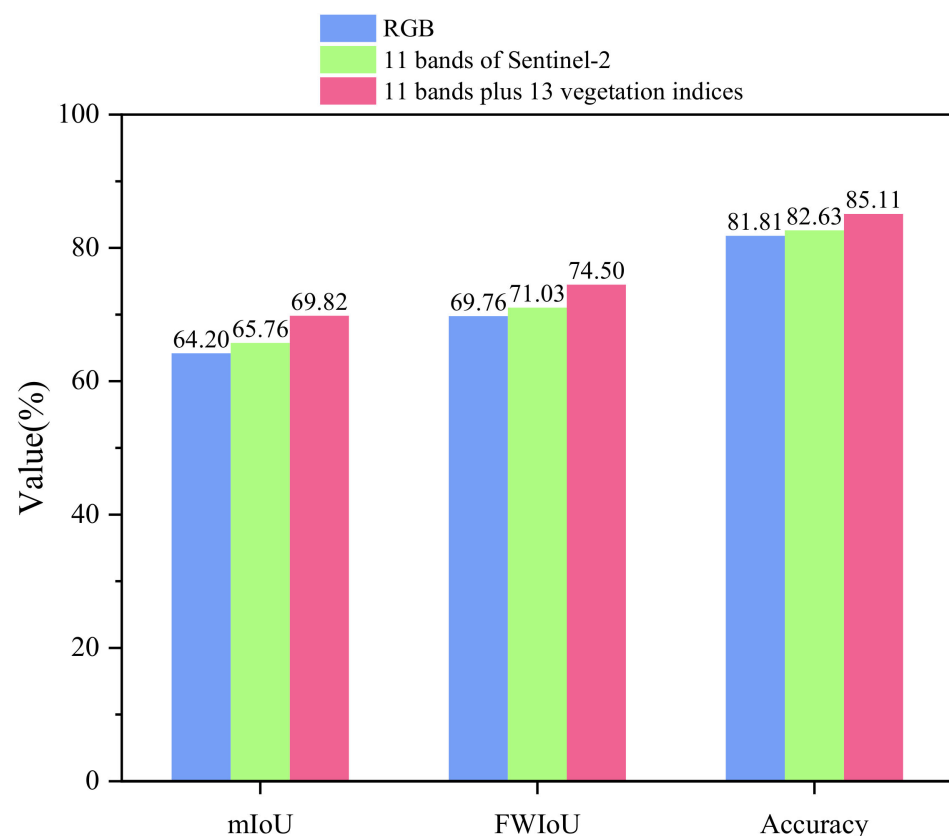
**Figure 10.** Comparison of the evaluation metrics after the two additions of the bands.

Table 7. Effect of different data on the segmentation result of infested areas using RSPR-UNet++.

Data	Category	Precision (%)	Recall (%)	F1 (%)	IoU (%)	mIoU (%)	FWIoU (%)	Accuracy (%)
RGB	BG	87.06	85.67	86.36	75.99	64.20	69.76	81.81
	BB	73.45	73.32	73.38	57.96			
	ALM	71.32	76.76	73.94	58.66			
11 bands	BG	88.96	84.98	86.93	76.87	65.76	71.03	82.63
	BB	70.04	81.40	75.29	60.38			
	ALM	76.43	73.65	75.02	60.02			
RGB plus 13 vegetation indices	BG	89.79	84.79	87.22	77.34	67.06	72.01	83.31
	BB	76.23	78.65	77.42	63.16			
	ALM	68.99	83.43	75.52	60.67			
11 bands plus 13 vegetation indices	BG	89.92	87.53	88.71	79.70	69.82	74.50	85.11
	BB	78.10	79.52	78.81	65.02			
	ALM	75.16	82.33	78.58	64.72			
8 bands plus 10 vegetation indices	BG	86.43	89.63	88.00	78.57	68.17	72.76	84.13
	BB	80.66	74.73	77.58	63.37			
	ALM	78.76	75.28	76.98	62.57			
11 bands plus 21 vegetation indices	BG	84.68	90.62	87.55	77.85	68.12	72.17	83.85
	BB	81.21	75.63	78.32	64.37			
	ALM	84.00	70.48	76.65	62.14			

To analyze the effects of the 13 vegetation indices and the remaining bands of Sentinel-2 except RGB, the results of training and testing by using RGB plus 13 vegetation indices are also compared in Table 7. Compared with the original 11 bands of Sentinel-2, the mIoU improved by 1.3%, the FWIoU improved by 0.98%, and the accuracy improved by 0.68% when using RGB plus 13 vegetation indices for training.

Table 7 also displays the experimental results of training and testing the proposed model using Sentinel-2's original 11 bands plus 13 previous vegetation indices plus eight newly added vegetation indices related to red edge and the 18 bands of the 32 bands that are not related to red edge. After deleting the bands and vegetation indices associated with red edge from the 11 bands and the previous 13 vegetation indices, the mIoU decreased by 1.65%, the FWIoU decreased by 1.74%, and the accuracy decreased by 0.98%. In contrast, the mIoU decreased by 1.7%, the FWIoU decreased by 2.33%, and the accuracy decreased by 1.26% following the addition of 8 vegetation indices related to the red edge to the 11 bands and 13 vegetation indices.

4.2. Comparison of the Spectral Characteristics of Different Infestation Types

We extracted the pixel values of each band for various types of pest-infested areas in the dataset and then averaged the pixel values of each band. As illustrated in Figure 11, the spectral figures of bark beetle and aspen leaf miner were acquired. The differences between bark beetle and aspen leaf miner on the original bands of Sentinel-2 are relatively small, whereas the differences between the vegetation indices, with the exception of CVI, are greater.

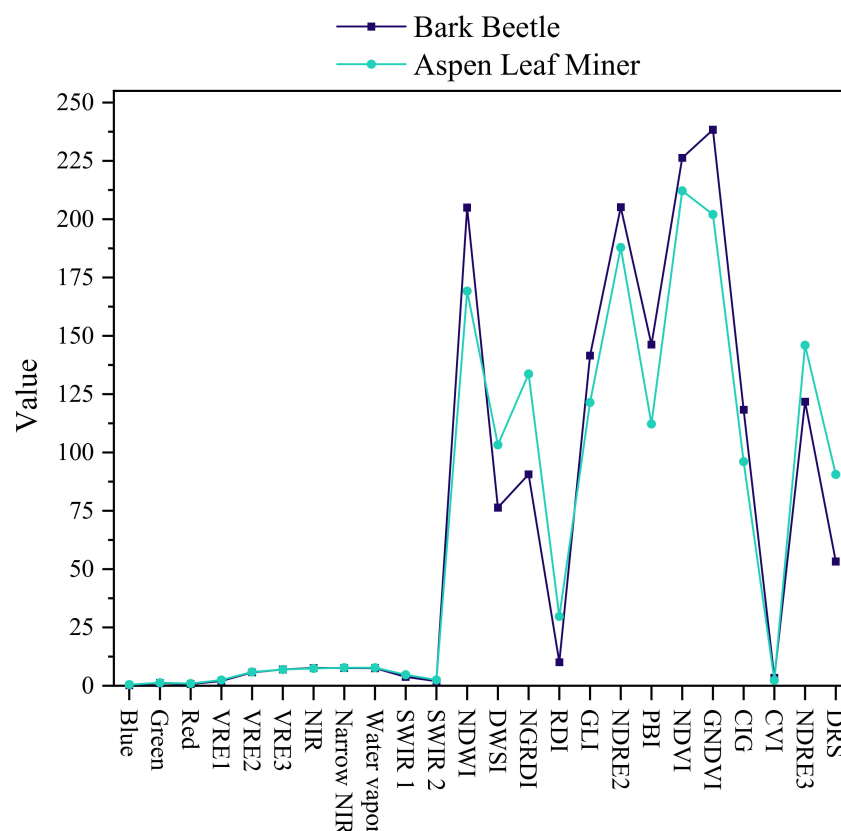


Figure 11. The spectral figure of bark beetle and aspen leaf miner. VRE represents Vegetation Red Edge.

5. Discussion

We proposed a segmentation model based on Unet++ and named RSPR-UNet++ for the extraction from optimized Sentinel-2 images of bark beetle and aspen leaf miner infested regions. Compared with other common semantic segmentation models, RSPR-UNet++ introduces attention mechanism [57–59], and the model is more advanced in structure. For the optimization of Sentinel-2 images, we added 13 bands based on the formulae for 13 vegetation indices to the original bands, drawing on the existing research results in pest research and the common degree of vegetation indices in remote sensing image analysis. Previous study demonstrated the feasibility of employing satellite remote sensing technology to monitor forest pests [60–64]. The current research utilizing deep learning to detect forest damage employ RGB images. Multispectral satellite images and vegetation indices were not utilized in this field. We demonstrated the importance of vegetation indices and multispectral data to improve the segmentation effect with experiments. As far as we know, our research is the first exploitation of deep learning for forest-pest infestation area segmentation on multispectral satellite images, particularly those containing numerous vegetation indices.

As shown in Table 5, RSPR-UNet++ outperforms other models in terms of evaluation metrics such as accuracy, mIoU, and FWIoU for pest region segmentation. We speculate that the main reason is that RSPR-UNet++ employs ResNeSt for feature extraction and incorporates the scSE attention mechanism module. However, we were uncertain that both ResNeSt and scSE played a positive role in the segmentation effect, so we added ablation experiments to determine the effect of the scSE module and ResNeSt on RSPR-UNet++ independently. After RSPR-UNet++ eliminates the scSE module, virtually all evaluation metrics are reduced. We also added contrast experiments to explore the influence of cSE and sSE on the segmentation effect. According to Table 8, both cSE and sSE contribute positively to the segmentation effect. It suggests that the mechanism of recalibration and excitation of features from both spatial and channel dimensions by the scSE module enables the

model to concentrate more on the features that are useful for detecting infested regions and suppress the useless features, thereby enhancing the segmentation effect. Comparing the evaluation metrics of RSPR-UNet++ without the scSE module to those of UNet++ reveals that nearly all of the former are superior to the latter, indicating that ResNeSt's capacity to extract features is superior than the original UNet++'s encoder portion. The Split-Attention block within ResNeSt divides the input feature-map into feature-map splits, calculates weights for each feature-map split, and then combines the feature-map splits to produce a new feature-map. This procedure extracts additional characteristics of the infested area. ResNeSt and scSE both have the attention operation of assigning weights to features and combining new feature-maps, and they each positively contribute to RSPR-UNet++. It demonstrates that redundant features can have a negative impact on the model's ability to extract pest regions from multispectral remote sensing images, which contain both rich contextual semantic and spectral information. The attention mechanism allows the model to concentrate more on useful characteristics, thereby improving the final results. Additionally, we adjusted the channels' number of the feature-maps output from X^{0j} , X^{1j} , X^{2j} , X^{3j} and X^{4j} ($j = 0, 1, 2, 3, 4$) experimentally. As shown in Table 9, we attempted these three sets of parameters due to our hardware limitations.

Table 8. Effects of different attention modules.

Attention Module	Accuracy (%)
scSE	85.11
cSE	84.93
sSE	84.85
None	84.61

Table 9. The channels' number of the feature-maps output.

Parameters	Accuracy (%)
16, 32, 64, 128 and 256	85.11
32, 64, 128, 256 and 512	85.03
64, 128, 256, 512 and 1024	83.24

As can be seen in Table 7, when training with the original 11 bands of Sentinel-2, nearly all evaluation metrics were enhanced compared to when RGB was used. It indicates that the other Sentinel-2 bands provide favorable characteristics for pest region segmentation. In addition, after adding 13 vegetation indices for training, nearly all evaluation metrics were further enhanced compared to Sentinel-2's initial bands. Compared to the original 11 bands of Sentinel-2, it is indicated that the bands of Sentinel-2 plus 13 vegetation indices provide more useful features for pest area segmentation. Comparing the results of these three experiments, as depicted in Figure 10, reveals that the proposed model's segmentation effect is improved more by adding 13 vegetation indices based on 11 Sentinel-2 bands than by adding the remaining Sentinel-2 bands based on RGB bands. We speculate that this is due to the fact that the 13 vegetation indices contain more characteristics than the native bands of Sentinel-2, excluding RGB. To further test this hypothesis, we employed a control variables approach, i.e., an experiment with RGB plus 13 vegetation indices. The results indicate that training with RGB plus 13 vegetation indices is superior to training with 11 Sentinel-2 bands. It indicates that adding 13 vegetation indices based on RGB improves the segmentation effect of the model more than adding the remaining bands of Sentinel-2 based on RGB. Our hypothesis was confirmed that vegetation indices can provide additional features such as spectral information about insect-infested and non-infested areas, thereby enhancing the model's ability to differentiate between insect-infested and non-infested areas and enhancing segmentation performance. These four experiments also demonstrate that multispectral data and vegetation indices provide more effective features than RGB data for the task of segmenting insect-infested areas, where the internal chemical

composition of trees is significantly altered, thereby enabling the model to segment the infested areas more precisely.

In addition, Table 7 shows that training the model with the data after removing the red-edge related bands and vegetation indices from Sentinel-2's original 11 bands and previous 13 vegetation indices resulted in a decrease in the majority of evaluation metrics. This suggests that the red-edge bands and the vegetation indices derived from the red edges contain useful information for assessing forest health. Therefore, we hypothesize that increasing the red-edge related vegetation indices may improve the model's precision. Consequently, we added eight red-edge-based vegetation indices to the Sentinel-2's original 11 bands and previous 13 vegetation indices for training. Contrary to expectations, nearly all evaluation indices decreased, falling even lower than the results of training the model with the data after removing the red-edge-related bands and vegetation indices from Sentinel-2's original 11 bands and previous 13 vegetation indices. We argue that this is due to the redundancy of information at this point, which has a negative impact on the model's segmentation effect.

The above experiments illustrate that the nonlinear combination of the original bands is important. The information in the vegetation indices comes from the original bands. To some extent, all the information is contained in the original bands, and the vegetation indices make the information more obvious.

Although the model has similar results in segmenting bark beetle and aspen leaf miner in the test set, the segmentation results shown in Figure 9 show that the model detects bark beetle better than aspen leaf miner. We speculate that this is because the coverage area of bark beetle in the dataset is much larger than that of aspen leaf miner, which causes the model to be less generalizable to aspen leaf miner.

6. Conclusions

Accurate extraction of bark beetle and aspen leaf miner infected areas from remote sensing images is important for monitoring forest health and protecting forest ecosystems. In this study, a UNet++-based semantic segmentation model is proposed for segmenting bark beetle and aspen leaf miner infected regions in remote sensing images. In the encoder, we employ ResNeSt to improve the quality of the extracted features. To enhance the most important features for segmentation, the scSE module is introduced in the decoder. The experimental results demonstrated that the proposed model outperforms the state-of-the-art methods, such as UNet, DeeplabV3+, UNet++, etc. Moreover, to illustrate the potential of multispectral data and vegetation indices for pest area segmentation, we also implement the segmentation with RGB images, the results indicate that multispectral data and vegetation indices are more advantageous for pest area extraction since the vegetation indices can provide a wealth of information regarding the pest areas' characteristics. Although this work focuses only on the forest pests of bark beetle and aspen leaf miner, our proposed method can also be extended for segmenting the other forest pest areas. In the future, investigating the optimal band and vegetation indices to further improve the segmentation performance will be interesting. The remote sensing data usually tend to suffer from various degradation, noise effects, or variabilities in the process of imaging [65]. Solving this problem and applying the solution to pest area extraction can also be a direction in future work.

Author Contributions: Conceptualization, J.Z. and J.H.; Data curation, J.Z.; Formal analysis, J.Z., S.C., G.Z., Y.M., Y.Z. and J.H.; Investigation, J.Z.; Methodology, J.Z.; Project administration, J.H.; Software, J.Z.; Supervision, J.H.; Validation, J.Z.; Visualization, J.Z.; Writing—original draft, J.Z.; Writing—review & editing, J.Z. and J.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (No. 61701105), the Natural Science Foundation of Heilongjiang Province of China (Grant No. TD2020C001), and the Fundamental Research Funds for the Central Universities (Grant No. 2572019CP19).

Data Availability Statement: Data sharing is not applicable to this article.

Acknowledgments: We are thankful to Jiuqing Liu for providing funding support and supervision in this study. We are also grateful for anonymous reviewers' hard work and review.

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

References

- Guo, Z.; Wu, C.; Wang, X. Forest insect-disease monitoring and estimation based on satellite remote sensing data. *Geogr. Res.* **2019**, *38*, 831–843.
- Hicke, J.A.; Xu, B.; Meddens, A.J.H.; Egan, J.M. Characterizing recent bark beetle-caused tree mortality in the western United States from aerial surveys. *For. Ecol. Manag.* **2020**, *475*, 118402. [\[CrossRef\]](#)
- Meddens, A.J.H.; Hicke, J.A. Spatial and temporal patterns of Landsat-based detection of tree mortality caused by a mountain pine beetle outbreak in Colorado, USA. *For. Ecol. Manag.* **2014**, *322*, 78–88. [\[CrossRef\]](#)
- Wagner, D.; DeFoliart, L.; Doak, P.; Schneiderheinze, J. Impact of epidermal leaf mining by the aspen leaf miner (*Phyllocnistis populiella*) on the growth, physiology, and leaf longevity of quaking aspen. *Oecologia* **2008**, *157*, 259–267. [\[CrossRef\]](#) [\[PubMed\]](#)
- Wagner, D.; Doak, P.; Sformo, T.; Steiner, P.M.; Carlson, B. Overwintering physiology and microhabitat use of *Phyllocnistis populiella* (Lepidoptera: Gracillariidae) in interior Alaska. *Environ. Entomol.* **2012**, *41*, 180–187. [\[CrossRef\]](#) [\[PubMed\]](#)
- Young, B.; Wagner, D.; Doak, P.; Clausen, T. Induction of phenolic glycosides by quaking aspen (*Populus tremuloides*) leaves in relation to extrafloral nectaries and epidermal leaf mining. *J. Chem. Ecol.* **2010**, *36*, 369–377. [\[CrossRef\]](#)
- Doak, P.; Wagner, D. The role of interference competition in a sustained population outbreak of the aspen leaf miner in Alaska. *Basic Appl. Ecol.* **2015**, *16*, 434–442. [\[CrossRef\]](#)
- 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. [\[CrossRef\]](#)
- Yu, R.; Ren, L.L.; Luo, Y.Q. Early detection of pine wilt disease in *Pinus tabulaeformis* in North China using a field portable spectrometer and UAV-based hyperspectral imagery. *For. Ecosyst.* **2021**, *8*, 44. [\[CrossRef\]](#)
- Dennison, P.E.; Brunelle, A.R.; Carter, V.A. Assessing canopy mortality during a mountain pine beetle outbreak using GeoEye-1 high spatial resolution satellite data. *Remote Sens. Environ.* **2010**, *114*, 2431–2435. [\[CrossRef\]](#)
- Zhan, Z.; Yu, L.; Li, Z.; Ren, L.; Gao, B.; Wang, L.; Luo, Y. Combining GF-2 and Sentinel-2 Images to Detect Tree Mortality Caused by Red Turpentine Beetle during the Early Outbreak Stage in North China. *Forests* **2020**, *11*, 172. [\[CrossRef\]](#)
- Ma, L.; Liu, Y.; Zhang, X.; Ye, Y.; Yin, G.; Johnson, B.A. Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS J. Photogramm. Remote Sens.* **2019**, *152*, 166–177. [\[CrossRef\]](#)
- Hong, D.; Gao, L.; Yokoya, N.; Yao, J.; Chanussot, J.; Du, Q.; Zhang, B. More diverse means better: Multimodal deep learning meets remote-sensing imagery classification. *IEEE Trans. Geosci. Remote Sens.* **2020**, *59*, 4340–4354. [\[CrossRef\]](#)
- Wu, X.; Hong, D.; Chanussot, J. Convolutional neural networks for multimodal remote sensing data classification. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 1–10. [\[CrossRef\]](#)
- Hong, D.; Gao, L.; Yao, J.; Zhang, B.; Plaza, A.; Chanussot, J. Graph convolutional networks for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **2020**, *59*, 5966–5978. [\[CrossRef\]](#)
- Shi, L.; Li, C.; Li, T.; Peng, Y. A Complementary Spectral–Spatial Method for Hyperspectral Image Classification. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*. [\[CrossRef\]](#)
- Kotaridis, I.; Lazaridou, M. Remote sensing image segmentation advances: A meta-analysis. *ISPRS J. Photogramm. Remote Sens.* **2021**, *173*, 309–322. [\[CrossRef\]](#)
- Kislov, D.E.; Korznikov, K.A.; Altman, J.; Vozmishcheva, A.S.; Krestov, P.V.; Disney, M.; Cord, A. Extending deep learning approaches for forest disturbance segmentation on very high-resolution satellite images. *Remote Sens. Ecol. Conserv.* **2021**, *7*, 355–368. [\[CrossRef\]](#)
- Zhou, H.; Yuan, X.; Zhou, H.; Shen, H.; Ma, L.; Sun, L.; Fang, G.; Sun, H. Surveillance of pine wilt disease by high resolution satellite. *J. For. Res.* **2022**, *33*, 1401–1408. [\[CrossRef\]](#)
- Cotrozzi, L. Spectroscopic detection of forest diseases: A review (1970–2020). *J. For. Res.* **2021**, *33*, 21–38. [\[CrossRef\]](#)
- Zarco-Tejada, P.J.; Camino, C.; Beck, P.S.A.; Calderon, R.; Hornero, A.; Hernandez-Clemente, R.; Kattenborn, T.; Montes-Borrego, M.; Susca, L.; Morelli, M.; et al. Previsual symptoms of *Xylella fastidiosa* infection revealed in spectral plant-trait alterations. *Nat. Plants* **2018**, *4*, 432–439. [\[CrossRef\]](#)
- Zarco-Tejada, P.J.; Hornero, A.; Beck, P.S.A.; Kattenborn, T.; Kempeneers, P.; Hernandez-Clemente, R. Chlorophyll content estimation in an open-canopy conifer forest with Sentinel-2A and hyperspectral imagery in the context of forest decline. *Remote Sens. Environ.* **2019**, *223*, 320–335. [\[CrossRef\]](#)
- Kattenborn, T.; Leitloff, J.; Schiefer, F.; Hinz, S. Review on Convolutional Neural Networks (CNN) in vegetation remote sensing. *ISPRS J. Photogramm. Remote Sens.* **2021**, *173*, 24–49. [\[CrossRef\]](#)
- Shen, Q.; Deng, J.; Liu, X.; Huang, H. Prediction of bark beetles pests based on temperature vegetation dryness index. *Trans. Chin. Soc. Agric. Eng.* **2018**, *34*, 167–174.
- Wu, H.; Wang, C.; Miao, Z.; Wang, W.; Wang, X.; Mi, G. Study on Early Warning Technology of Sub-health State of Forest Resources with Spaceborne Remote Sensing. *Remote Sens. Technol. Appl.* **2021**, *36*, 1121–1130.

26. Meng, R.; Gao, R.; Zhao, F.; Huang, C.; Sun, R.; Lv, Z.; Huang, Z. Landsat-based monitoring of southern pine beetle infestation severity and severity change in a temperate mixed forest. *Remote Sens. Environ.* **2022**, *269*, 112847. [\[CrossRef\]](#)
27. Goodwin, N.R.; Coops, N.C.; Wulder, M.A.; Gillanders, S.; Schroeder, T.A.; Nelson, T. Estimation of insect infestation dynamics using a temporal sequence of Landsat data. *Remote Sens. Environ.* **2008**, *112*, 3680–3689. [\[CrossRef\]](#)
28. Zhu, C.; Qu, S.; Zhang, X. Dendrolimus tabulaeformis disaster monitoring and analysis of its influencing factors through remote sensing technology. *J. Remote Sens.* **2016**, *20*, 653–664.
29. Hais, M.; Jonášová, M.; Langhammer, J.; Kučera, T. Comparison of two types of forest disturbance using multitemporal Landsat TM/ETM+ imagery and field vegetation data. *Remote Sens. Environ.* **2009**, *113*, 835–845. [\[CrossRef\]](#)
30. Bárta, V.; Lukeš, P.; Homolová, L. Early detection of bark beetle infestation in Norway spruce forests of Central Europe using Sentinel-2. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *100*, 102335. [\[CrossRef\]](#)
31. Choi, W.-I.L.; Kim, E.-S.; Yun, S.-J.; Lim, J.-H.; Kim, Y.-E. Quantification of One-Year Gypsy Moth Defoliation Extent in Wonju, Korea, Using Landsat Satellite Images. *Forests* **2021**, *12*, 545. [\[CrossRef\]](#)
32. Kern, A.; Marjanović, H.; Csóka, G.; Móríc, N.; Pernek, M.; Hirka, A.; Matošević, D.; Paulin, M.; Kovač, G. Detecting the oak lace bug infestation in oak forests using MODIS and meteorological data. *Agric. For. Meteorol.* **2021**, *306*, 108436. [\[CrossRef\]](#)
33. Xu, Z.H.; Huang, X.Y.; Lin, L.; Wang, Q.F.; Liu, J.; Chen, C.C.; Yu, K.Y.; Zhou, H.K.; Zhang, H.F. Dendrolimus Punctatus Walker Damage Detection Based on Fisher Discriminant Analysis and Random Forest. *Spectrosc. Spectr. Anal.* **2018**, *38*, 2888–2896. [\[CrossRef\]](#)
34. Zou, H. The Effect and Method of Quantitative Analysis of “Red Edge” of Vegetation. *Remote Sens. Inf.* **2010**, *4*, 112–116.
35. Boochs, F.; Kupfer, G.; Dockter, K.; Kühbauch, W. Shape of the red edge as vitality indicator for plants. *Remote Sens.* **1990**, *11*, 1741–1753. [\[CrossRef\]](#)
36. Filella, I.; Penuelas, J. The red edge position and shape as indicators of plant chlorophyll content, biomass and hydric status. *Int. J. Remote Sens.* **1994**, *15*, 1459–1470. [\[CrossRef\]](#)
37. Abdullah, H.; Skidmore, A.K.; Darvishzadeh, R.; Heurich, M.; Pettorelli, N.; Disney, M. Sentinel-2 accurately maps green-attack stage of European spruce bark beetle (*Ips typographus*, L.) compared with Landsat-8. *Remote Sens. Ecol. Conserv.* **2018**, *5*, 87–106. [\[CrossRef\]](#)
38. Eitel, J.U.; Vierling, L.A.; Litvak, M.E.; Long, D.S.; Schulthess, U.; Ager, A.A.; Krofcheck, D.J.; Stoscheck, L. Broadband, red-edge information from satellites improves early stress detection in a New Mexico conifer woodland. *Remote Sens. Environ.* **2011**, *115*, 3640–3646. [\[CrossRef\]](#)
39. Carter, G.A.; Miller, R.L. Early detection of plant stress by digital imaging within narrow stress-sensitive wavebands. *Remote Sens. Environ.* **1994**, *50*, 295–302. [\[CrossRef\]](#)
40. Ahern, F. The effects of bark beetle stress on the foliar spectral reflectance of lodgepole pine. *Int. J. Remote Sens.* **1988**, *9*, 1451–1468. [\[CrossRef\]](#)
41. Carter, G.A.; Knapp, A.K. Leaf optical properties in higher plants: Linking spectral characteristics to stress and chlorophyll concentration. *Am. J. Bot.* **2001**, *88*, 677–684. [\[CrossRef\]](#)
42. Dash, J.P.; Watt, M.S.; Pearce, G.D.; Heaphy, M.; Dungey, H.S. Assessing very high resolution UAV imagery for monitoring forest health during a simulated disease outbreak. *Isprs J. Photogramm. Remote Sens.* **2017**, *131*, 1–14. [\[CrossRef\]](#)
43. Fernandez-Manso, A.; Fernandez-Manso, O.; Quintano, C. SENTINEL-2A red-edge spectral indices suitability for discriminating burn severity. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *50*, 170–175. [\[CrossRef\]](#)
44. Bhattarai, R.; Rahimzadeh-Bajgiran, P.; Weiskittel, A.; MacLean, D.A. Sentinel-2 based prediction of spruce budworm defoliation using red-edge spectral vegetation indices. *Remote Sens. Lett.* **2020**, *11*, 777–786. [\[CrossRef\]](#)
45. Hojas-Gascon, L.; Belward, A.; Eva, H.; Ceccherini, G.; Hagolle, O.; Garcia, J.; Cerutti, P. Potential improvement for forest cover and forest degradation mapping with the forthcoming Sentinel-2 program. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2015**, *40*, 417–423. [\[CrossRef\]](#)
46. Immitzer, M.; Vuolo, F.; Atzberger, C. First Experience with Sentinel-2 Data for Crop and Tree Species Classifications in Central Europe. *Remote Sens.* **2016**, *8*, 166. [\[CrossRef\]](#)
47. Zhou, Z.; Rahman Siddiquee, M.M.; Tajbakhsh, N.; Liang, J. Unet++: A nested u-net architecture for medical image segmentation. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 3–11.
48. Westfall, J.; Ebata, T.; Bains, B. *Summary of Forest Health Conditions in British Columbia*; BC Ministry of Forests and Range, Forest Practices Branch: Victoria, BC, Canada, 2020.
49. Zabihi, K.; Surovy, P.; Trubin, A.; Singh, V.V.; Jakuš, R. A review of major factors influencing the accuracy of mapping green-attack stage of bark beetle infestations using satellite imagery: Prospects to avoid data redundancy. *Remote Sens. Appl. Soc. Environ.* **2021**, *24*, 100638. [\[CrossRef\]](#)
50. Zhang, H.; Wu, C.; Zhang, Z.; Zhu, Y.; Lin, H.; Zhang, Z.; Sun, Y.; He, T.; Mueller, J.; Manmatha, R. Resnest: Split-attention networks. *arXiv* **2020**, arXiv:2004.08955.
51. 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.
52. Roy, A.G.; Navab, N.; Wachinger, C. Recalibrating fully convolutional networks with spatial and channel “squeeze and excitation” blocks. *IEEE Trans. Med. Imaging* **2018**, *38*, 540–549. [\[CrossRef\]](#)

53. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015; pp. 234–241.
54. Chen, L.-C.; Zhu, Y.; Papandreou, G.; Schroff, F.; Adam, H. Encoder-decoder with atrous separable convolution for semantic image segmentation. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 801–818.
55. Lin, T.-Y.; Dollár, P.; Girshick, R.; He, K.; Hariharan, B.; Belongie, S. Feature pyramid networks for object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 2117–2125.
56. Li, H.; Xiong, P.; An, J.; Wang, L. Pyramid attention network for semantic segmentation. *arXiv* **2018**, arXiv:1805.10180.
57. Wang, Z.; Wang, J.; Yang, K.; Wang, L.; Su, F.; Chen, X. Semantic segmentation of high-resolution remote sensing images based on a class feature attention mechanism fused with Deeplabv3+. *Comput. Geosci.* **2022**, *158*, 104969. [[CrossRef](#)]
58. Wang, D.; Zhang, C.; Han, M. MLFC-net: A multi-level feature combination attention model for remote sensing scene classification. *Comput. Geosci.* **2022**, *160*, 105042. [[CrossRef](#)]
59. Wang, D.; Chen, X.; Jiang, M.; Du, S.; Xu, B.; Wang, J. ADS-Net: An Attention-Based deeply supervised network for remote sensing image change detection. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *101*, 102348.
60. Bright, B.C.; Hudak, A.T.; Egan, J.M.; Jorgensen, C.L.; Rex, F.E.; Hicke, J.A.; Meddens, A.J.H. Using Satellite Imagery to Evaluate Bark Beetle-Caused Tree Mortality Reported in Aerial Surveys in a Mixed Conifer Forest in Northern Idaho, USA. *Forests* **2020**, *11*, 529. [[CrossRef](#)]
61. Meigs, G.W.; Kennedy, R.E.; Gray, A.N.; Gregory, M.J. Spatiotemporal dynamics of recent mountain pine beetle and western spruce budworm outbreaks across the Pacific Northwest Region, USA. *For. Ecol. Manag.* **2015**, *339*, 71–86. [[CrossRef](#)]
62. Ye, S.; Rogan, J.; Zhu, Z.; Hawbaker, T.J.; Hart, S.J.; Andrus, R.A.; Meddens, A.J.H.; Hicke, J.A.; Eastman, J.R.; Kulakowski, D. Detecting subtle change from dense Landsat time series: Case studies of mountain pine beetle and spruce beetle disturbance. *Remote Sens. Environ.* **2021**, *263*, 112560. [[CrossRef](#)]
63. Hart, S.J.; Veblen, T.T. Detection of spruce beetle-induced tree mortality using high- and medium-resolution remotely sensed imagery. *Remote Sens. Environ.* **2015**, *168*, 134–145. [[CrossRef](#)]
64. 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](#)]
65. Hong, D.; Yokoya, N.; Chanussot, J.; Zhu, X.X. An augmented linear mixing model to address spectral variability for hyperspectral unmixing. *IEEE Trans. Image Process.* **2018**, *28*, 1923–1938. [[CrossRef](#)] [[PubMed](#)]