



## Article

# A New Individual Tree Species Classification Method Based on the ResU-Net Model

Caiyan Chen <sup>1,2</sup>, Linhai Jing <sup>1,\*</sup>, Hui Li <sup>1</sup>  and Yunwei Tang <sup>1</sup> 

<sup>1</sup> Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China; chencaiyan19@mails.ucas.ac.cn (C.C.); lihui@radi.ac.cn (H.L.); tangyw@aircas.ac.cn (Y.T.)

<sup>2</sup> School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 100049, China

\* Correspondence: jinglh@aircas.ac.cn

**Abstract:** Individual tree species (ITS) classification is one of the key issues in forest resource management. Compared with traditional classification methods, deep learning networks may yield ITS classification results with higher accuracy. In this research, the U-Net and ResNet networks were combined to form a Res-UNet network by changing the structure of the convolutional layer to the residual structure in ResNet based on the framework of the U-Net model. In addition, a second Res-UNet network named Res-UNet2 was further constructed to explore the effect of the stacking of residual structures on network performance. The Res-UNet2 model structure is similar to that of the Res-UNet model, but the convolutional layer in the U-Net model is created with a double-layer residual structure. The two networks proposed in this work were used to classify ITSs in WorldView-3 images of the Huangshan Mountains, Anhui Province, China, acquired in March 2019. The resulting ITS map was compared with the classification results obtained with U-Net and ResNet. The total classification accuracy of the ResU-Net network reached 94.29% and was higher than that generated by the U-Net and ResNet models, verifying that the ResU-Net model can more accurately classify ITSs. The Res-UNet2 model performed poorly compared to Res-UNet, indicating that stacking the residual modules in ResNet does not achieve an accuracy improvement.

**Keywords:** remote sensing; individual tree species classification; deep learning; U-Net; ResU-Net



**Citation:** Chen, C.; Jing, L.; Li, H.; Tang, Y. A New Individual Tree Species Classification Method Based on the ResU-Net Model. *Forests* **2021**, *12*, 1202. <https://doi.org/10.3390/f12091202>

Received: 21 July 2021

Accepted: 1 September 2021

Published: 4 September 2021

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



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Forest resources are among the most important natural resources for humankind, and increasing attention is being paid to the management of forest resources [1–3]. Determining how to classify individual trees, the smallest component of a forest, is of great significance and research value for forest resource management [4,5]. However, the traditional individual tree species (ITS) measurement method is time consuming and laborious, and it is difficult to widely use in rugged areas [6]. With the development of remote sensing technology, high-spatial-resolution remote sensing imagery can be used to determine the positions of individual trees and delineate individual tree crowns, thus providing the possibility for large-scale ITS classification [7–10].

At present, ITS classification technology is mainly based on high-spatial-resolution airborne multispectral or hyperspectral data, high-point-density LiDAR point cloud data, or their combination [11–13]. With these data, a variety of ITSs can be identified, and high classification accuracies can be obtained [14–16]. For example, Guan et al. [17] used deep Boltzmann machines (DBMs) and LiDAR data to obtain the high-level features of individual trees; then, they used a support vector machine (SVM) to classify ten tree species, with an overall accuracy of 86.1%. Zou et al. [18] employed a deep belief network (DBN) and 3D point clouds to classify five tree species, with an overall accuracy of 93.1%. Wang

et al. [19] selected unmanned aerial vehicle (UAV) imagery and a back propagation (BP) neural network to classify six tree species, with an overall accuracy of 89.1%.

However, the acquisition cost of such data is high, so the relevant tree species classification techniques can only be applied to small localized areas [20,21]. With the launch of QuickBird, WorldView, and other satellites, it has become easy to obtain large quantities of high-spatial-resolution satellite imagery and accurately locate the positions of individual trees in images. In addition, the acquisition cost of these data is relatively low, and they can provide data support for ITS classification over large areas [22]. An increasing number of researchers have begun to use high-resolution satellite imagery to perform ITS classification. For example, Korznikov et al. [23] utilized a U-Net-like CNN and RGB imagery from GeoEye-1 to conduct ITS recognition in northern temperate mixed forests, with a mean balanced accuracy of 96%. Roth et al. [24] applied Pléiades 1A satellite data and a circular Hough transform (CHT) to detect individual trees and then used random forests (RF) to classify clove trees, with an overall accuracy of 77.9%.

Traditional image classification methods used for tree species classification mainly include RF and SVM models [25–28]. For example, Immitzer et al. [29] used 8-band WorldView-2 satellite data to classify ten tree species in an Australian temperate forest with pixel-based and object-based RF methods. The overall accuracy of the object-based method was 82.4% ( $\kappa = 0.799$ ), which was better than that of the pixel-based method (overall accuracy = 72.8%,  $\kappa = 0.678$ ). Harikumar et al. [30] performed high-density airborne laser scanning (ALS) and used an SVM to classify deciduous, coniferous, and mixed (both deciduous and coniferous) trees, with overall accuracies of 94.2%, 71.4%, and 72.3%, respectively. Kuzmin et al. [31] selected high-resolution photogrammetric point clouds (PPCs), multispectral (MSP) orthomosaics acquired with an unmanned aerial vehicle (UAV), and an SVM to identify European aspen at the individual tree level in a southern boreal forest, with a classification accuracy of 86% (F1-score). Traditional imagery classification methods require the manual extraction of features; however, most of the features are shallow features and are easily affected by subjective experience and selection [32–35]. With the development of deep learning [36,37], the use of deep learning networks for classification can aid in the extraction of deeper features and produce more accurate classification results [38–41]. For example, Nezami et al. [42] used UAV imagery and 3D convolutional neural networks (3D-CNNs) to classify three major tree species in a boreal forest: pine, spruce, and birch. The producer accuracies of the best 3D-CNN classifier based on a test dataset were 99.6%, 94.8%, and 97.4% for pine, spruce, and birch trees, respectively. Sun et al. [43] employed LiDAR and VGG16 networks to classify eighteen species in Haizhu National Wetland Park, Guangzhou, China, with an overall accuracy of 73.25%. Fricker et al. [44] utilized high-spatial-resolution airborne hyperspectral imagery and a convolutional neural network (CNN) to classify seven dominant tree species in a mixed-conifer forest, with an overall accuracy of 87%. Ferreira et al. [45] applied UAV imagery and a full CNN model to classify three palm trees, with a mean producer's accuracy of 87.8%.

Among the many deep learning classification networks available, the CNN-based U-Net network is widely used by researchers due to its multiscale feature fusion and rich feature classification abilities [46,47]. For example, Giang et al. [48] utilized U-Net and UAV imagery to classify six land cover types, with an overall accuracy of 84.8%. Wagner et al. [49] used U-Net and WorldView-3 imagery to classify natural forests and eucalyptus plantations, with an overall accuracy of 95.4%. Schiefer et al. [50] used U-Net and UAV imagery to classify nine tree species, with an overall accuracy of 89%.

Once set, the depth scale of the U-Net network cannot be changed. If there are too few layers in a deep learning network, training may be insufficient, and the classification accuracy can be low. If there are too many layers in a deep learning network, network degradation can occur, and the classification accuracy may be reduced [51]. The unique residual structure of ResNet [52] can effectively alleviate such problems associated with U-Net.

Combining ResNet with a U-Net network to build a new network structure (ResU-Net network) can solve the performance degradation issue of U-Net under extreme depth conditions [53]. Additionally, this approach enables the U-Net network to contain deeper layers and consider training parameters at the same depth scale. To a certain extent, the insufficient depth problem for U-Net related to setting a fixed depth size can be avoided, and classification performance can be improved [54].

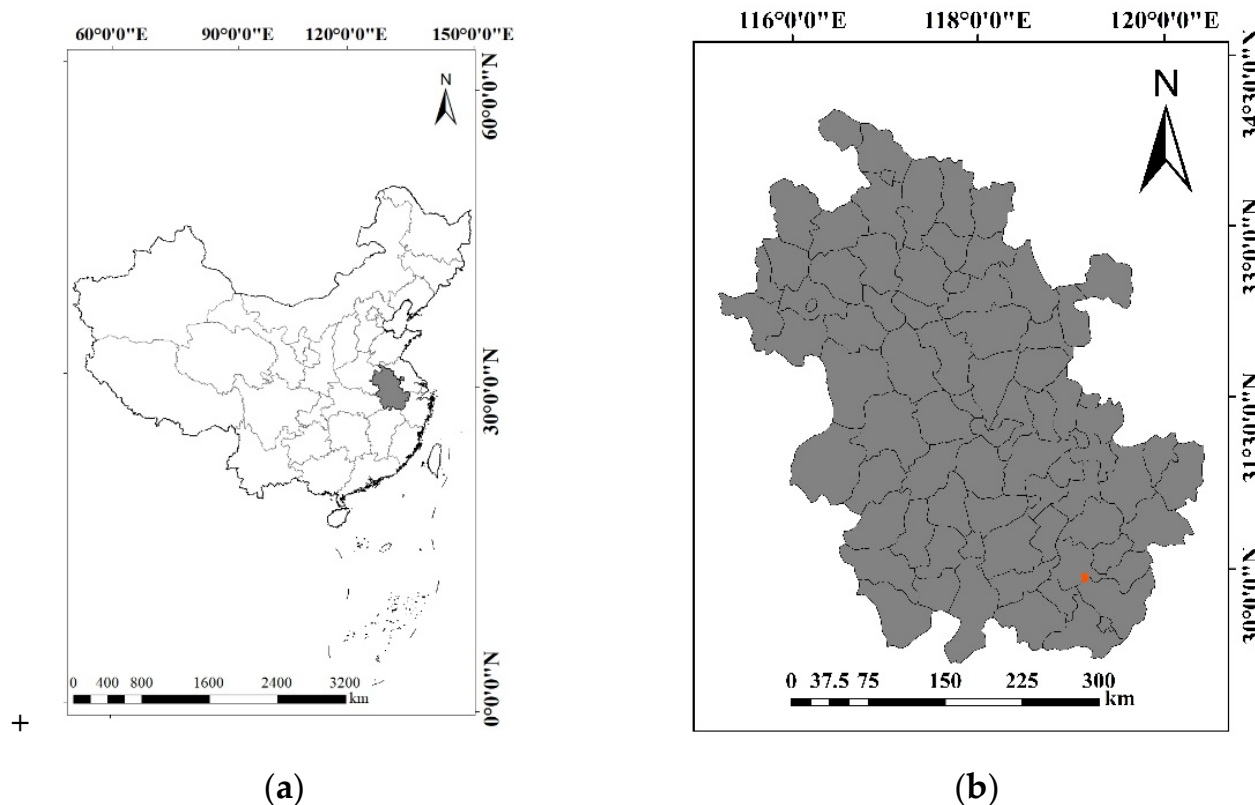
In this study, high-spatial-resolution satellite imagery and an improved U-Net model were used to identify ITSs. An ITS sample set from remote sensing images of the study area was established and enhanced, and U-Net, ResNet, and ResU-Net were used to classify ITSs in the study area. In addition, the effect of ResU-Net models composed of different combinations of ResNet and U-Net to potentially improve ITS classification accuracy is discussed.

The majority of this paper is divided into six parts: Section 2 gives the details of the proposed methods and introduces the experiments, Section 3 presents and discusses the experimental results, Section 4 provides the discussion, and the conclusions are presented in Section 5.

## 2. Materials and Methods

### 2.1. Study Area

The study area is located in the scenic region of the Huangshan Mountains, Anhui Province, as shown in Figure 1, with 56% forest coverage in the scenic area. The gray area in Figure 1a denotes the province where the study area is located, and the red rectangle in Figure 1b represents the study area within the province. The geographical location of the study area is from 118°9′16″ to 118°11′24″ east longitude and 30°7′8″ to 30°10′37″ north latitude, with a total area of 21.6 km<sup>2</sup>. The study area is located in a subtropical monsoon climate zone and on the northern edge of the central subtropics. This area is also located in an evergreen broad-leaved forest with red soil and yellow soil zones. As the topography is characterized by raised peaks and long ravines, the climate varies vertically, and the vertical zoning of plants is obvious in the study area. There are more natural forests than planted forests and more mixed forests than pure forests. The pure forests of *Pinus taiwanensis* (*Pinus taiwanensis* Hayata) are mainly distributed on the top of Huangshan Mountain. Mixed forests of *Pinus taiwanensis* (*Pinus taiwanensis* Hayata) (or *Pinus massoniana* (*Pinus massoniana* Lamb)) and broad-leaved trees are distributed near the peak and on the banks of Huangshan Mountain. Pure broadleaf forests are distributed at the bottom of the mountain. Artificial arbor forests are mainly distributed around the scenic area. The tree species in the scenic area mainly include *Cunninghamia lanceolata* (*Cunninghamia lanceolata* (Lamb.) Hook), pine (*Pinus*), *Phyllostachys pubescens* (*Phyllostachys heterocycla* (Carr.) Mitford cv. *Pubescens*), and arbors (macrophanerophytes).



**Figure 1.** The map of the study area. (a) The gray province denotes where the study area is located. (b) The red rectangle represents the study area within the province.

## 2.2. Experimental Data

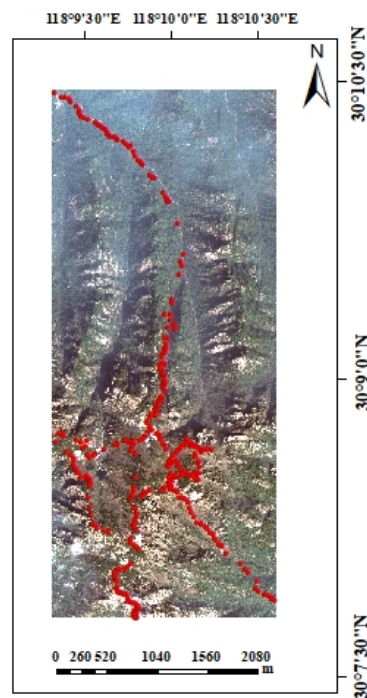
The WorldView-3 imagery of the study area was acquired on 10 March 2019. The imagery has one panchromatic band with a spatial resolution of 0.3 m and eight multispectral bands with a spatial resolution of 1.2 m. The wavelength ranges of the eight multispectral bands are shown in Table 1.

**Table 1.** The WorldView-3 image bands.

Band Order	Band Name	Wavelength (nm)
Band 1	coastal band	400–450
Band 2	blue band	450–510
Band 3	green band	510–580
Band 4	yellow band	585–625
Band 5	red band	630–690
Band 6	red-edge band	705–745

Due to the large extent and complex topography of the Huangshan scenic area, the field sampling area was selected on both sides of the viewing route in the field investigation of the research group, and GPS devices were used to collect sample points in the sampling area.

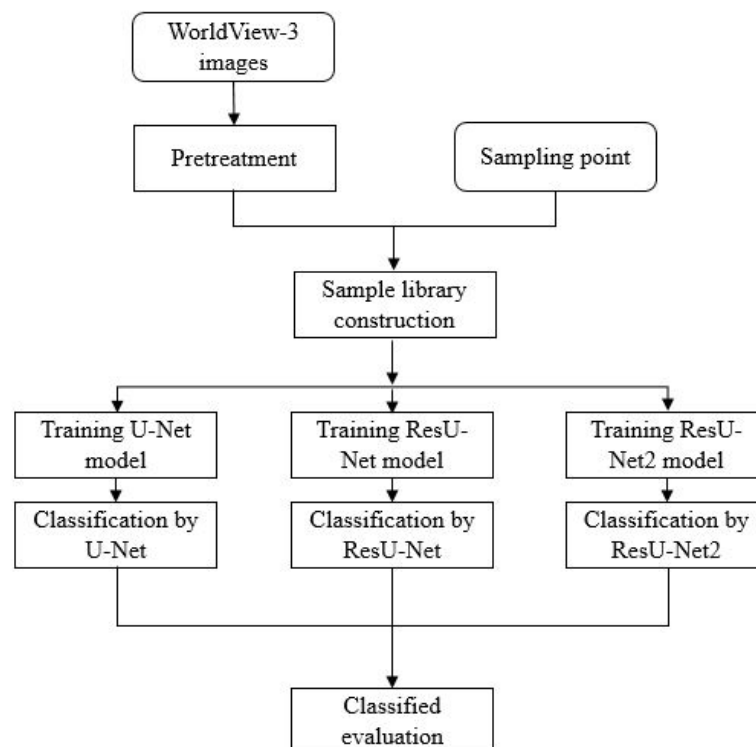
The WorldView-3 imagery and the distribution of sampling points are shown in Figure 2. Bands 5, 3, and 2 were used in the false color composite of WorldView-3.



**Figure 2.** WorldView-3 study area imagery.

### 2.3. Experimental Process

The research process is divided into four main parts: data preprocessing, sample dataset construction, network training and classification, and classification evaluation. This specific process is shown in Figure 3.



**Figure 3.** Flow chart of the experimental process.

#### 2.4. Data Preprocessing

First, the WorldView-3 imagery was processed based on radiometric calibration, orthorectification, and image clipping. The imagery adopted apparent reflection without atmospheric correction. Then, to fully use the spectral and texture information in the remote sensing imagery, the fusion method based on haze and ratio (HR) information proposed by Jing et al. [55] was used to fuse panchromatic imagery and multispectral imagery and obtain 8-band imagery with a spatial resolution of 0.3 m for subsequent experiments. The HR fusion method introduces the haze in the spectral band. The fused image retains the texture information of panchromatic band and greatly reduces the spectral distortion.

#### 2.5. Sample Dataset Construction

##### 2.5.1. Building the Sample Set

To construct the remote sensing imagery sample set of ITSs, it is necessary to first remove the parts of images containing only an individual crown from whole remote sensing images and then determine and label the category of the individual tree crown.

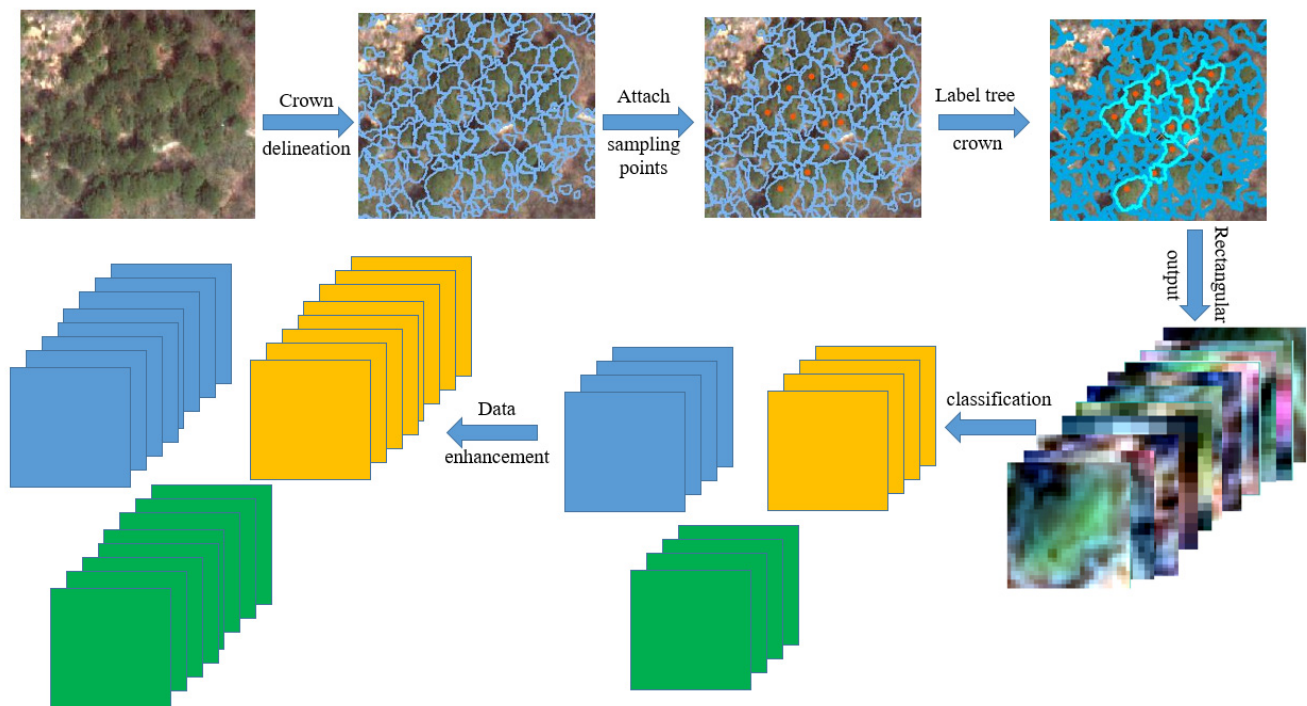
Due to the wide coverage of remote sensing imagery and the complexity of ground objects, it is necessary to use certain methods to extract and describe the tree crowns extracted from remote sensing imagery. In this study, the crown slice from imagery (CSI) algorithm proposed by Jing in 2014 [56] was used for the multiscale segmentation of remote sensing imagery and automatic crown delineation. Due to the small difference in texture information between different tree species in remote sensing images, the CSI algorithm uses the information of crown texture and spectral brightness to delineate the crown, and the crown delineation result is more accurate.

The steps in constructing the remote sensing imagery sample set of ITSs are shown in Figure 4. The main steps are as follows: (1) field survey the study area, and collect samples of tree species categories; (2) use the CSI algorithm to automatically circle individual tree crowns in the remote sensing images of the study area; (3) combine the collected sample points, crown circling results, and remote sensing imagery interpretation results to label individual tree crown species categories; (4) output the labeled tree crowns based on the smallest outer rectangle to obtain remote sensing image maps for individual tree species; and (5) categorize the remote sensing image maps for individual species according to the species categories to build a sample set containing the main tree species in the study area.

The tree species categories were combined according to the field sampling results, and the individual tree crown images were classified after category labeling. The ITS imagery sample set was divided into five categories, as shown in Table 2. There were 112 samples of *Phyllostachys pubescens*, 44 samples of evergreen arbor, 139 samples of *Cunninghamia lanceolata*, 2001 samples of *Pinus taiwanensis*, and 617 samples of deciduous arbor.

**Table 2.** The remote sensing sample set of ITSs.

Tree Species	Species Merged	Shorthand	Field Sampling Points	Classification Labeled Sample Set
<i>Phyllostachys pubescens</i>	<i>Phyllostachys pubescens</i>	Ph. p	18	112
<i>Osmanthus fragrans</i> / <i>Camellia japonica</i> / <i>Manglietia</i> sp./ <i>Rhododendron</i> sp./ <i>Ilex chinensis</i> / <i>Daphniphyllum macropodum</i> / <i>Daphniphyllum oldhamii</i>	Evergreen arbor	Ev.a	113	444
<i>Abies fabri</i> / <i>Taxus</i> sp./ <i>Tsuga chinensis</i>	<i>Cunninghamia lanceolata</i>	Cu.l	67	139
<i>Pinus taiwanensis</i>	<i>Pinus taiwanensis</i>	Pi.t	245	2001
<i>Tilia japonica</i> / <i>Cyclobalanopsis glauca</i> / <i>Castanea seguinii</i> / <i>Emmenopterys henryi</i> / <i>Sorbus</i> sp./ <i>Acer</i> sp.	Deciduous arbor	De.a	327	617
Total			703	3313



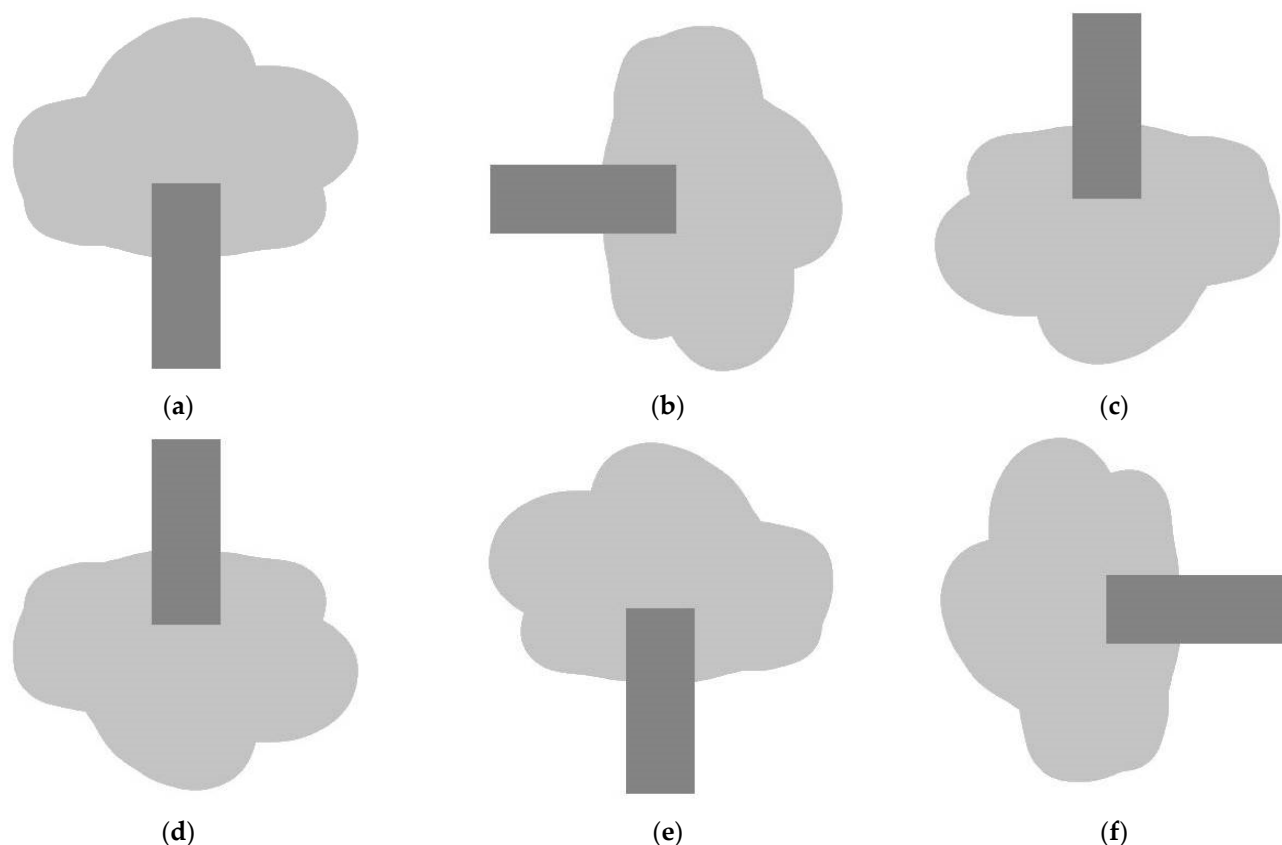
**Figure 4.** Schematic diagram of the steps in building the ITS sample set.

### 2.5.2. Data Augmentation

A deep learning network often requires a large number of samples for training to achieve good accuracy. However, the number of samples labeled in the above process may not be sufficient. Therefore, after each class in the sample set was divided into training sample set, verification sample set, and test sample set in the ratio of 3:1:1, it is necessary to enhance the sample set. In this study, the original ITS samples were rotated by 90 degrees, 180 degrees, and 270 degrees; horizontally flipped; and turned upside down (as shown in Figure 5). The number of sample sets was expanded to six times the original number, as shown in Table 3.

**Table 3.** The enhanced sample set of ITSs.

Tree Species	Training Sample Set	Validation Sample Set	Test Sample Set
Ph. p	396	138	23
Ev.a	1596	534	89
Cu.l	498	168	28
Pi.t	7194	2406	401
De.a	2214	744	124
Total	11,898	3990	665



**Figure 5.** The schematic diagram of data enhancement. (a) Schematic diagram of an individual tree; (b) schematic diagram of 90° rotation; (c) schematic diagram of 180° rotation; (d) schematic diagram of 270° rotation; (e) schematic diagram of horizontal flipping; (f) schematic diagram of vertical flipping.

## 2.6. Network Training and Classification

### 2.6.1. Improved U-Net Model

The U-Net structure is mainly divided into a downsampling stage and an upsampling stage, with only convolutional and pooling layers in the network and no fully connected layers. In the network, a shallow high-resolution layer is used to solve the location problem, and a deeper layer is used to solve the pixel classification problem to promote the subsequent semantic level segmentation and classification of the imagery. In the upsampling stage and downsampling stage of U-Net, convolution operations are performed at the same level. The skip connection structure is used to connect the downsampling layer with the upsampling layer so that the features extracted from the downsampling layer can be directly transferred to the upsampling layer, thereby increasing the accuracy of pixel locations in U-Net and the accuracy of segmentation and classification. Generally, the input imagery of U-Net includes individual-channel grayscale images used in image segmentation operations. In this study, by adding reshape, flattening and softmax operations, the improved U-Net network can complete the classification of 8-channel image data.

The structure of the improved U-Net network is shown in Figure 6. The short red arrows represent the convolution and activation operation. The long red arrows represent the copy and crop operation. The short downward yellow arrows represent the maximum pooling operation. The upward yellow arrows represent the upper convolution operation, and the green arrows represent the convolution operation with a  $1 \times 1$  convolution kernel.



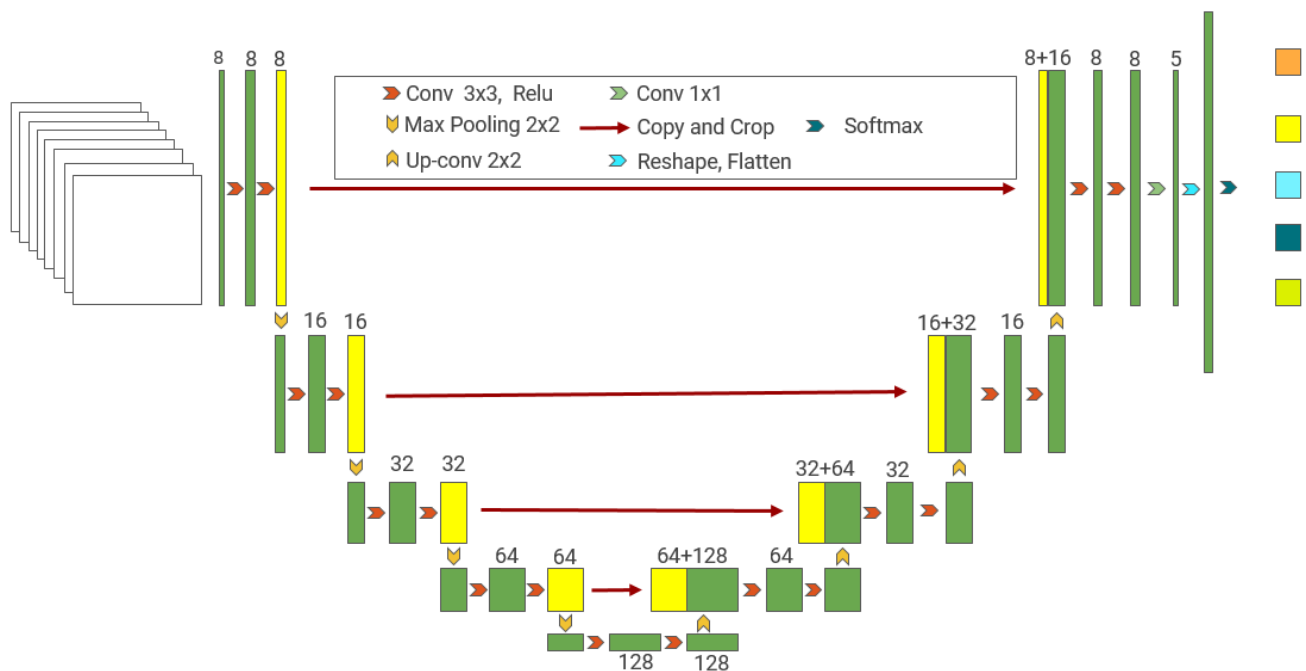


Figure 6. U-Net network structure.

### 2.6.2. ResU-Net Model

The core module of ResNet involves the application of a residual structure and includes the elementwise addition of block inputs and outputs through shortcuts [52]. This simple addition does not add parameters or computations to the network model but can greatly improve the training speed and effectiveness of the model. In addition, the residual structure can solve the degradation problem for deep model layers. At present, there are two main types of residual structures (as shown in Figure 7), and since different numbers of filters are used in this study when combining ResNet and U-Net networks, the latter residual structure is most suitable in the context of maximizing the accuracy of the results.

The second residual structure is combined with U-Net to construct a new ResU-Net network structure, which is used for multichannel imagery classification. The ResU-Net network mainly adopts the U-Net framework with the same upsampling phase and downsampling phase, but unlike that in U-Net, the specific module in the network is implemented as the residual module in ResNet. The specific network architecture is shown in Figure 8, where Figure 8a shows the overall framework of the ResU-Net model and Figure 8b illustrates the specific implementation of each block.

To study whether the stacking of residual modules is helpful to further improve the accuracy of network classification, a second ResU-Net structure (ResU-Net2 network) was designed. Similar to ResU-Net, ResU-Net2 also adopts the U-Net network framework with two stages of upsampling and downsampling. However, two residual modules are stacked in each implementation module to explore whether reasonable stacking of residual modules can make the classification effect more prominent. The specific network architecture is shown in Figure 9, where Figure 9a is the overall framework of the ResU-Net2 model, and Figure 9b is the specific implementation of each block in the overall framework.

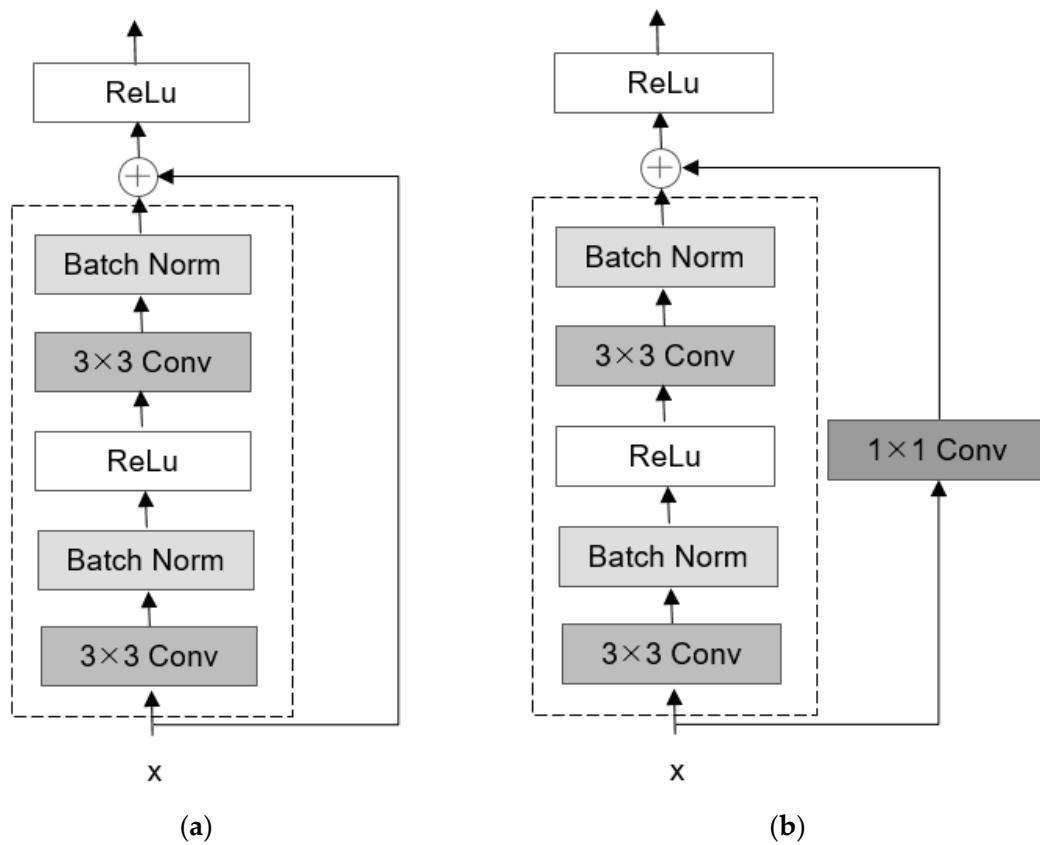


Figure 7. The residual structure without (a) and with (b) a  $1 \times 1$  convolution layer.

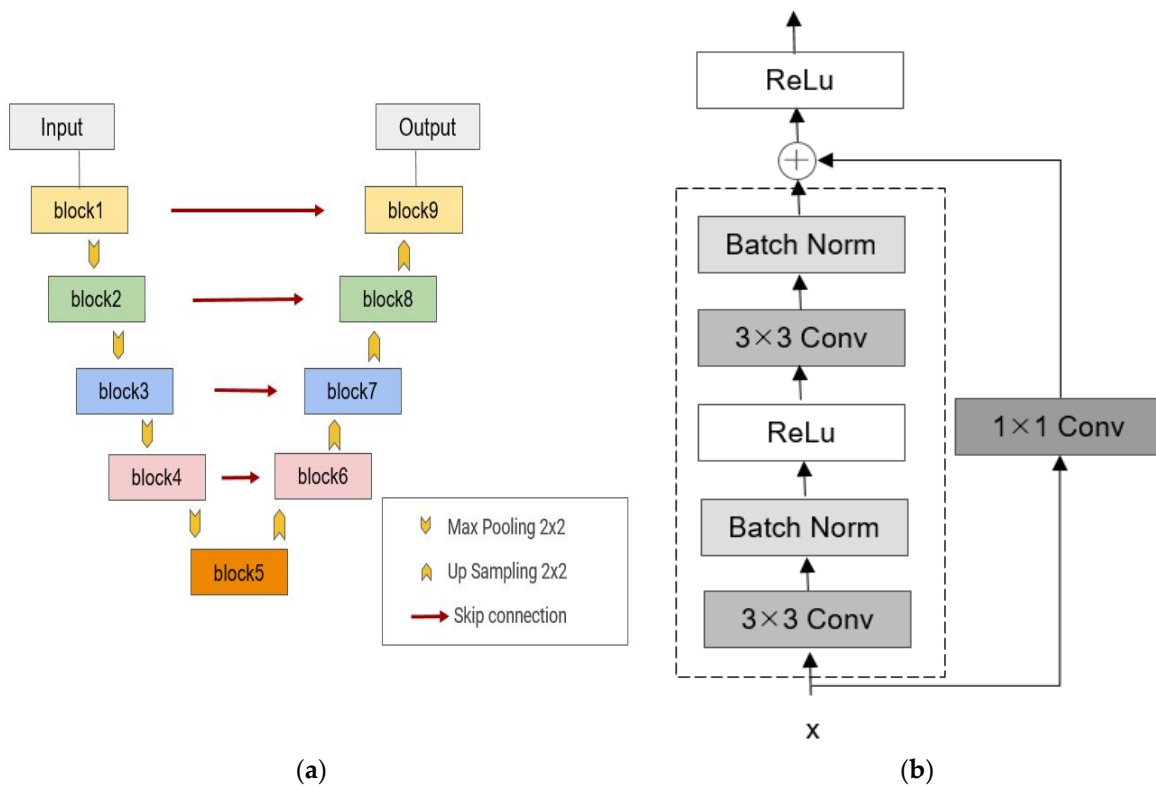


Figure 8. ResU-Net network structure. (a) ResU-Net network framework. (b) The specific implementation of block.

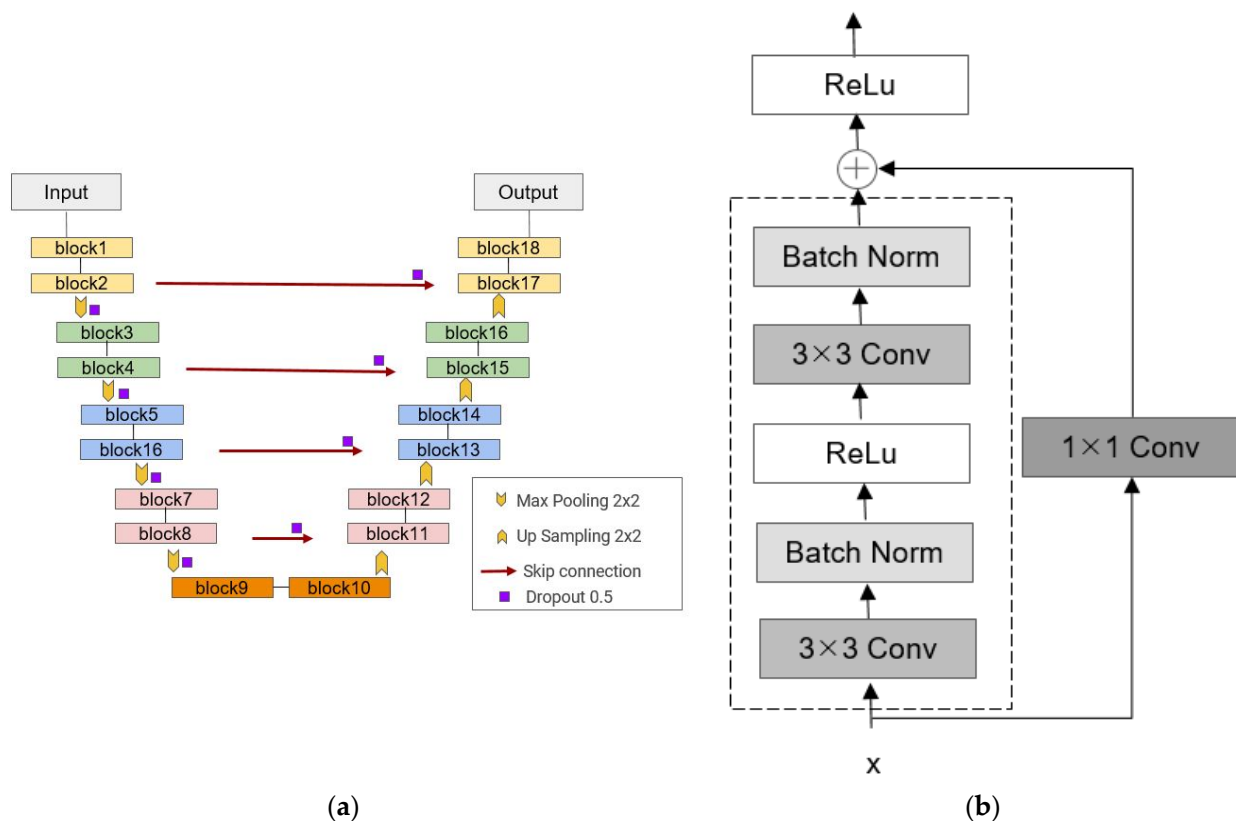


Figure 9. ResU-Net2 network structure: (a) ResU-Net2 network framework and (b) block modules.

### 2.6.3. Experimental Environment

The deep learning environment in this study was built with a Keras 2.2.4 front end and a TensorFlow 1.14.0 back end. The programming language was Python 3.6.8. The operating system was Windows 10, and the graphics card was an NVIDIA GTX1060. Parallel computing was achieved through CUDA10.0.

Based on the software, hardware, and dataset, four combinations of models and data were compared: the improved U-Net model and test dataset, the ResNet model and test dataset, the ResU-Net model and test dataset, and the ResU-Net2 model and test dataset.

### 2.6.4. Training and Prediction

In this study, the ReLU function was chosen as the neuron activation function for the four network models; the loss function was categorical\_crossentropy, and the iterative operator was the Adam optimizer. The batch number in each training cycle was 60, with a total of 500 iterations. In the iterative process, the network was allowed to complete five iterations without improvement (if an accuracy increase was greater than 0.001, it was regarded as an improvement.) If more than 5 iterations passed without an accuracy improvement, the learning rate was reduced by 0.005. The lower limit of the learning rate was  $0.5 \times 10^{-6}$ . After each epoch, the test results were obtained for the verification set. With increasing epochs, if the test error increased or the accuracy improved by less than 0.001 for more than 10 iterations, network training ended. The model with the highest classification accuracy for the validation set was saved and used to produce predictions based on the test set.

## 3. Results

### 3.1. Classification Accuracy

The U-Net, ResNet, ResU-Net, and ResU-Net2 models were trained by using the divided training sample sets. During the training process, the accuracy of the trained model

was verified by using the validation sample set. After iterative training, the convergence period, training accuracy, and verification accuracy of U-Net, ResNet, ResU-Net, and ResU-Net2 models were found and are shown in Table 4.

**Table 4.** The convergence period, training accuracy, and verification accuracy of the models.

Model	Convergence Period	Training Accuracy	Verification Accuracy
U-Net	82	94.28%	94.16%
ResNet	38	98.25%	94.49%
ResU-Net	43	95.77%	95.67%
ResU-Net2	28	93.86%	93.38%

The training results of the four models in Table 4 indicate that the three models (U-Net, ResU-Net, and ResU-Net2) that use the U-Net model framework yield the highest overall training and validation accuracies, with values above 93%, verifying that the U-Net model framework can accurately complete imagery classification tasks. Compared with U-Net, the ResU-Net model formed by combining U-Net and ResNet has a reduced convergence period, faster training speed, and significantly higher training and validation accuracy. Compared with the ResU-Net model, the ResU-Net2 model also has reduced training and validation accuracy, although the convergence period is reduced.

From the training and verification accuracies of the four models, the combination of the ResNet module and U-Net model framework can effectively increase the training depth and improve the training and verification accuracies compared to those of traditional methods, but the excessive stacking of residual modules will decrease the classification accuracy.

To fully illustrate the applicability of the model in the classification and prediction of different tree species, the confusion matrix for the four models was calculated by using the test sample set. The producer's accuracy, user's accuracy, overall accuracy, and kappa coefficient were obtained according to the confusion matrix, as shown in Table 5.

According to Table 5, (1) for *Phyllostachys pubescens* (Ph.p), the ResU-Net model and the ResNet model yielded the highest producer's accuracy values and low sample omission error, while the ResU-Net2 model yielded the lowest producer's accuracy and largest sample omission error. The same user's accuracy of 100% was observed for all four models. (2) For evergreen arbor (Ev.a), the producer's accuracy values of the ResU-Net and ResNet models were the highest, and the producer's accuracy of the ResU-Net2 model was the lowest; the user's accuracy of the ResU-Net2 model was the highest, with the ResU-Net model ranking second, and the user's accuracy of the U-Net model was the worst. (3) For *Cunninghamia lanceolata* (Cu.l), the producer's accuracy of the ResU-Net model was the highest, the producer's accuracy of the U-Net model was the lowest, and the user's accuracy of the ResU-Net2 model was the lowest. (4) For *Pinus taiwanensis* (Pi.t), the ResU-Net model yielded the highest producer's accuracy and user's accuracy, and all models performed well in terms of both of these metrics. (5) For deciduous arbor (De.a), the ResU-Net2 model yielded the highest producer's accuracy and the lowest user's accuracy, and the ResU-Net model displayed the highest user's accuracy and the lowest producer's accuracy. (6) The ResU-Net model exhibits the highest overall accuracy and kappa coefficient, and the overall accuracy and kappa coefficient of the U-Net and ResU-Net2 models are the same. (7) In general, compared with the other three models, the ResU-Net model provides better overall performance and higher stability, and it is more suitable for the classification of ITSs in remote sensing imagery. Although the performance of the U-Net, ResNet, and ResU-Net2 models may be outstanding in terms of the producer's accuracy or user's accuracy for a certain tree species, the overall classification ability and classification stability are inferior to those of the ResU-Net model.

**Table 5.** The evaluation of model classification accuracy.

Model	Evaluating Indicator		Real Tree Species				
			Ph. p	Ev.a	Cu.l	Pi.t	De.a
U-Net	Predicted tree species	Ph. p	21	0	0	0	0
		Ev.a	0	65	4	9	6
		Cu.l	1	0	20	0	0
		Pi.t	1	18	3	390	1
		De.a	0	6	1	2	117
	Producer's accuracy/%	91.30	73.03	71.43	97.26	94.35	
	User's accuracy/%	100.00	77.38	95.24	94.43	92.86	
	Overall accuracy/% Kappa coefficient			92.18 0.86			
ResNet	Predicted tree species	Ph. p	22	0	0	0	0
		Ev.a	0	68	4	8	4
		Cu.l	0	0	22	0	0
		Pi.t	1	13	2	391	5
		De.a	0	8	0	2	115
	Producer's accuracy/%	<b>95.65</b>	<b>76.40</b>	78.57	97.51	92.74	
	User's accuracy/%	100.00	80.95	<b>100.00</b>	94.90	92.00	
	Overall accuracy/% Kappa coefficient			92.93 0.88			
ResU-Net	Predicted tree species	Ph. p	22	0	0	0	0
		Ev.a	0	68	5	8	0
		Cu.l	0	1	23	0	0
		Pi.t	1	15	0	392	2
		De.a	0	5	0	1	122
	Producer's accuracy/%	<b>95.65</b>	<b>76.40</b>	<b>82.14</b>	<b>97.76</b>	98.39	
	User's accuracy/%	100.00	83.95	95.83	<b>95.61</b>	<b>95.31</b>	
	Overall accuracy/% Kappa coefficient			<b>94.29</b> <b>0.90</b>			
ResU-Net2	Predicted tree species	Ph. p	20	0	0	0	0
		Ev.a	0	58	4	5	0
		Cu.l	2	2	21	0	0
		Pi.t	1	18	3	391	1
		De.a	0	11	0	5	123
	Producer's accuracy/%	86.96	65.17	75.00	97.51	<b>99.19</b>	
	User's accuracy/%	100.00	<b>86.57</b>	84.00	94.44	88.49	
	Overall accuracy/% Kappa coefficient			92.18 0.86			

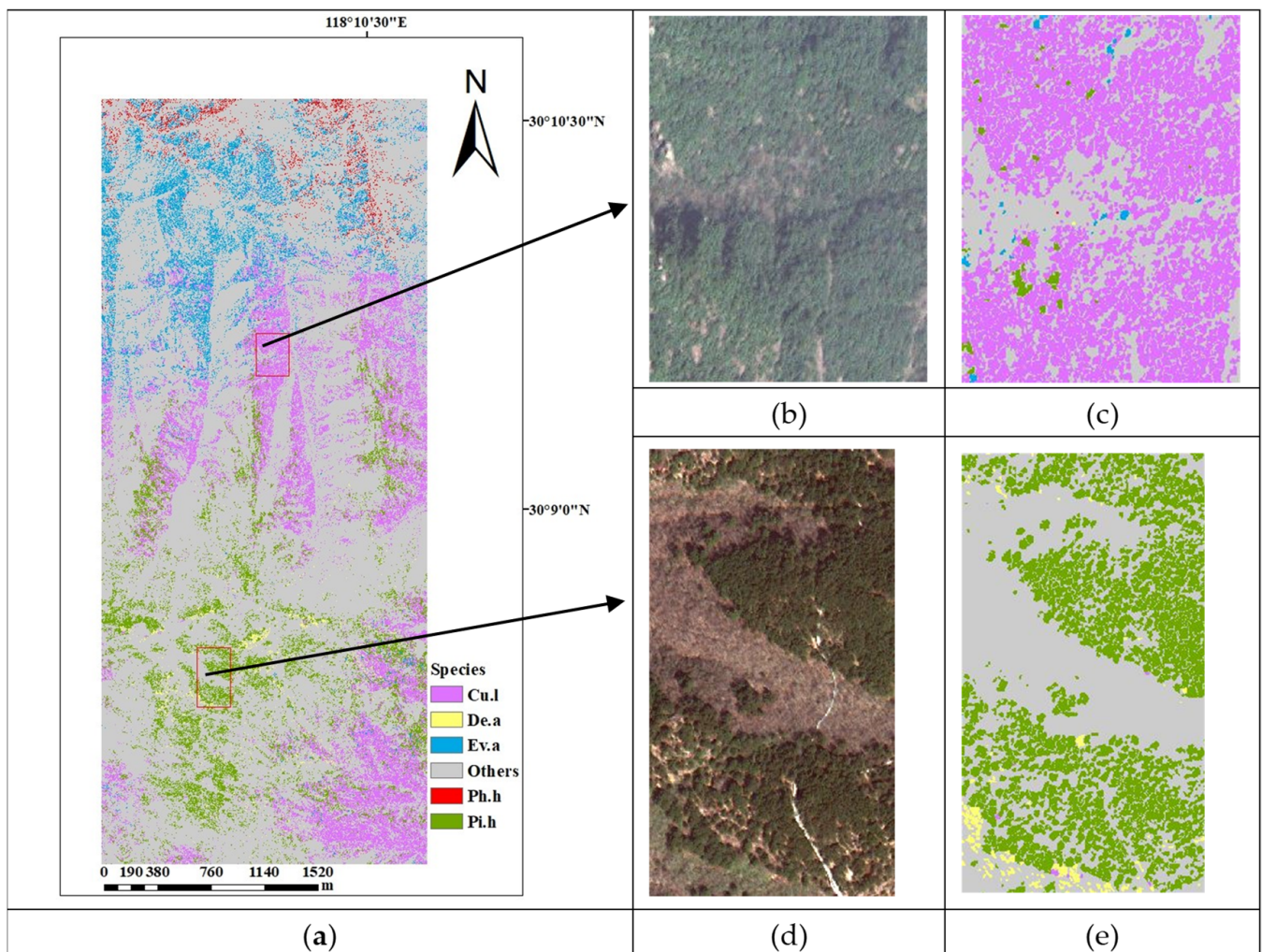
Producer's accuracy, user's accuracy, overall accuracy and kappa coefficient are compared for each of the four models, and the highest model metrics will be bolded.

### 3.2. Classification Map

The comparison in Section 3.1 shows that the ResU-Net model provides higher classification accuracy and stronger applicability than the other models. Therefore, the ResU-Net model was used to complete the classification of ITSs in WorldView-3 imagery in the study area. The following classification result plots were obtained, where the "others" category represent other nontree types.

According to Figure 10, the pure *Phyllostachys pubescens* forests are mainly distributed in the northern part of the research area. From north to south, mixed forests of *Phyllostachys pubescens* and evergreen arbors, evergreen arbor forests, mixed forests of evergreen arbor and *Cunninghamia lanceolata*, pure *Cunninghamia lanceolata* forests, mixed forests of *Cunninghamia lanceolata* and *Pinus taiwanensis*, and mixed forests of *Pinus taiwanensis* and deciduous arbor can be observed. The mixed forests of *Cunninghamia lanceolata* and *Pinus taiwanensis* and some pure forests of *Cunninghamia lanceolata* are distributed in the southeastern part of the study area. These species display an obvious vertical distribution with elevation, which is consistent with the vertical change in climate

in the region. The distribution of the same species in the study area is concentrated. Additionally, the distribution of different species displays transitional trends that are closely related to the climate change characteristics and the preferences of different species in the study area.



**Figure 10.** The tree species distribution map of the study area. (a) The overall tree species distribution in the study area. (b) Enlarged WorldView-3 imagery area 1. (c) Enlarged tree species distribution map area 1. (d) Enlarged WorldView-3 imagery area 2. (e) Enlarged tree species distribution map area 2.

#### 4. Discussion

##### 4.1. The Reliability of the U-Net Model Framework

The experimental results indicate that the three models that use the U-Net framework achieve good accuracy in ITS classification. This result verifies the classification reliability of the U-Net model framework, which improves the accuracy of classification through connecting shallow and deep features. Many scholars have also explored this area, such as Pan et al. [57], who segmented individual buildings using the Worldview satellite image with eight pan-sharpened bands at a 0.5 m spatial resolution and U-Net with segmentation accuracy over 86%. Wang et al. [58] integrated the spatial pyramid pooling module into the U-Net structure to classify 11 feature types with an overall accuracy of more than 86%.

In future research, further utilization of the advantages of the U-Net model framework should be explored, shallow and deep features should be further connected to improve the

modeling accuracy, and the use of other models with this framework should be investigated to improve the model.

#### 4.2. *The Residual Structure and U-Net Model Framework*

The experimental results show that the ResU-Net model can complete the high-precision classification of ITSs and reduce the classification error. Compared with the U-Net model, the ResU-Net model with a residual structure can reduce the degradation issue in the deep learning network, provide sufficient training samples, and improve the classification accuracy for a constant network depth.

The classification performance of traditional residual networks improves as the number of network layers is increased. However, in the framework of the U-Net model, the accuracy of the ResU-Net2 model does not improve with increasing residual structure. Compared with the ResU-Net model, the ResU-Net2 model displayed decreases in both classification accuracy and performance. The results suggest that it is necessary to explore the appropriate combination mode when the two network models are combined. An unsuitable combination mode will lead to performance degradation and network structure redundancy.

At present, some scholars have also explored different structures of ResU-Net models and achieved good results [59,60]. How to find a more suitable ResU-Net model among many combination methods is a problem to be explored in the future.

#### 4.3. *The Separability of Rare Tree Species Samples*

Due to the difficulty of collecting ITS samples in the field and the effects of various factors on the actual collection process, samples of a certain species are often too sparse to fully learn the corresponding classification characteristics and classify individuals. To solve such problems, samples of similar classes are often combined to obtain a large class, expand the sample size, and enhance classification. However, this method can also lead to the loss of ability to recognize rare trees, and only trees in the considered categories can be identified.

In addition, the degree of difficulty in classifying different individual tree species varies from remote sensing images. Individual trees in coniferous forests tend to be independent of each other, making it easy to circle and classify individual tree species. In broadleaf and high-density mixed forests, the canopies cover a wide area, and there are overlapping and cross-linking between canopies, which leads to the individual significance of individual canopies not being obvious, and it is not easy to identify and classify individual tree species [61].

Future research needs to continue to explore the use of small sample sets in network training to accurately discriminate among different tree classes and improve the accuracy of classification results.

#### 4.4. *The Selection of Data Sources*

In this experiment, WorldView-3 data with a spatial resolution of 0.3 m were used as the main research dataset. The WorldView-3 data include eight multispectral bands with abundant spectral and texture information. However, there are still some defects in these data, such as a spatial resolution that is not high enough to support the accurate extraction of individual tree crowns and locations of individual trees. Determining how to improve the positioning accuracy of individual tree crowns is an urgent problem that must be solved.

In addition, with the development of remote sensing technology and the continuous enrichment in remote sensing information sources, increasing amounts of information can be used for ITS classification. High-resolution remote sensing imagery contains a large number of canopy texture features, which can clearly reflect the fine details of the forest, which lays the foundation for high-precision extraction and classification of individual canopies. Hyperspectral images are rich in spectral information and hyperspectral images

have potential in distinguishing tree species of different ages [62,63]. LiDAR data, on the other hand, can effectively reconstruct the structure of individual wood types, thus contributing to the classification of ITS [64].

For the many available types of classification information, determining how to effectively combine multiple pieces of information and improve the accuracy of ITS classification are important tasks for the future.

#### 4.5. Experimental Errors

Although experimental errors were avoided as much as possible during the experimental process, some errors were unavoidable due to objective constraints. For instance, during the sampling process, there was a high possibility of misidentifying tree species due to the similarities among trees and the inadequacy of the knowledge of sampling personnel. In addition, the influence of random noise in the process of remote sensing imagery acquisition can lead to experimental error in tree species classification. The robustness of the classification network must be improved to reduce the influence of the experimental error. Moreover, the abilities of the fully trained classification network to recognize and eliminate mislabeled samples and perform satisfactory training must be researched in the future.

In this area of research, the application of generative adversarial networks (GANs) can be considered. There are two modules in the GAN framework: a generative model and discriminative model. The discriminative module can effectively identify the wrong samples, while the generative module can effectively solve the problem of lack of samples, thus helping to improve the classification ability for ITS [65–67].

## 5. Conclusions

In this study, remote sensing imagery and field sampling data in the study area were used to construct and enhance a remote sensing imagery ITS sample set, and a ResU-Net model was proposed for ITS classification. The following conclusions were obtained.

First, the introduction of the U-Net model framework improved the classification accuracy of ITSs. By comparing the classification results of the improved U-Net model with those of the two ResU-Net models, we found that the training and verification accuracies of the three models were all above 93%, and the overall accuracies were all greater than 92%. These results verify that the U-Net model framework is reliable for ITS classification.

Second, the combination of ResNet and U-Net can improve the accuracy of classification models. The combination of the residual module in ResNet and the U-Net model framework can effectively increase the model training depth and improve the training and verification accuracies. The results indicate that the training and verification accuracies of the ResU-Net model are 95.77% and 95.67%, respectively, with an overall accuracy of 94.29%. Compared with those of the U-Net model, the training and verification accuracies of the ResU-Net model are improved by 1.49% and 1.51%, respectively, and the overall accuracy is improved by 2.11%.

Finally, the excessive stacking of residual modules leads to a decline in the network classification ability. When there are too many residual modules stacked in the U-Net model framework, the classification accuracy decreases. By comparing the classification accuracies of the two ResU-Net models, we find that the training, validation, and overall accuracies of the ResU-Net2 model are 1.91%, 2.29%, and 2.11% lower than those for the ResU-Net model, respectively.

Through this study, a new network structure with high accuracy and suitability for ITS classification is constructed, and the influence of stacking residual modules on network performance is further considered to aid in future experiments.

**Author Contributions:** C.C. designed and completed the experiment; L.J., H.L. and Y.T. provided comments on the method; L.J., H.L. and Y.T. revised the manuscript and provided feedback on the manuscript. All authors have read and agreed to the published version of the manuscript.



**Funding:** This research was funded by the Aerospace Information Research Institute, Chinese Academy of Sciences (Grant No.: Y951150Z2F); the Science and Technology Major Project of Xinjiang Uygur Autonomous Region (2018A03004); and the National Natural Science Foundation of China (41972308 and 42071312).

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

**Acknowledgments:** We thank Guang Ouyang, Haoming Wan, and Xianfei Guo, who graduated from this research group, for their contributions to this research.

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

## References

1. Hyyppä, J.; Hyyppä, H.; Leckie, D.; Gougeon, F.; Yu, X.; Maltamo, M. Review of methods of small-footprint airborne laser scanning for extracting forest inventory data in boreal forests. *Int. J. Remote Sens.* **2008**, *29*, 1339–1366. [[CrossRef](#)]
2. Trivino, M.; Pohjanmies, T.; Mazziotta, A.; Juutinen, A.; Podkopaev, D.; Le Tortorec, E.; Monkkonen, M. Optimizing management to enhance multifunctionality in a boreal forest landscape. *J. Appl. Ecol.* **2017**, *54*, 61–70. [[CrossRef](#)]
3. Creedy, J.; Wurzbacher, A.D. The economic value of a forested catchment with timber, water and carbon sequestration benefits. *Ecol. Econ.* **2001**, *38*, 71–83. [[CrossRef](#)]
4. Yan, S.J.; Jing, L.H.; Wang, H. A New individual tree species recognition method based on a convolutional neural network and high-spatial resolution remote sensing imagery. *Remote Sens.* **2021**, *13*, 479. [[CrossRef](#)]
5. Torabzadeh, H.; Leiterer, R.; Hueni, A.; Schaepman, M.E.; Morsdorf, F. Tree species classification in a temperate mixed forest using a combination of imaging spectroscopy and airborne laser scanning. *Agric. For. Meteorol.* **2019**, *279*, 107744. [[CrossRef](#)]
6. Mey, R.; Stadelmann, G.; Thurig, E.; Bugmann, H.; Zell, J. From small forest samples to generalised uni- and bimodal stand descriptions. *Methods Ecol. Evol.* **2021**, *12*, 634–645. [[CrossRef](#)]
7. Reitberger, J.; Krzystek, P.; Stilla, U. Analysis of full waveform LIDAR data for the classification of deciduous and coniferous trees. *Int. J. Remote Sens.* **2008**, *29*, 1407–1431. [[CrossRef](#)]
8. Suratno, A.; Seielstad, C.; Queen, L. Tree species identification in mixed coniferous forest using airborne laser scanning. *ISPRS J. Photogramm. Remote Sens.* **2009**, *64*, 683–693. [[CrossRef](#)]
9. Yu, X.W.; Litkey, P.; Hyyppä, J.; Holopainen, M.; Vastaranta, M. Assessment of low density full-waveform airborne laser scanning for individual tree detection and tree species classification. *Forests* **2014**, *5*, 1011–1031. [[CrossRef](#)]
10. Hovi, A.; Korhonen, L.; Vauhkonen, J.; Korpela, I. LiDAR waveform features for tree species classification and their sensitivity to tree and acquisition related parameters. *Remote Sens. Environ.* **2016**, *173*, 224–237. [[CrossRef](#)]
11. Zhang, K.W.; Hu, B.X. Individual urban tree species classification using very high spatial resolution airborne multi-spectral imagery using longitudinal profiles. *Remote Sens.* **2012**, *4*, 1741–1757. [[CrossRef](#)]
12. Sun, Y.; Xin, Q.C.; Huang, J.F.; Huang, B.; Zhang, H.S. Characterizing tree species of a tropical wetland in southern China at the individual tree level based on convolutional neural network. *IEEE J-STARS* **2019**, *12*, 4415–4425. [[CrossRef](#)]
13. Zhao, D.; Pang, Y.; Liu, L.J.; Li, Z.Y. Individual tree classification using airborne LiDAR and hyperspectral data in a natural mixed forest of northeast China. *Forests* **2020**, *11*, 303. [[CrossRef](#)]
14. Puttonen, E.; Jaakkola, A.; Litkey, P.; Hyyppä, J. Tree classification with fused mobile laser scanning and hyperspectral data. *Sensors* **2011**, *11*, 5158–5182. [[CrossRef](#)]
15. Puttonen, E.; Litkey, P.; Hyyppä, J. Individual tree species classification by illuminated-shaded area separation. *Remote Sens.* **2010**, *2*, 19–35. [[CrossRef](#)]
16. Zhang, Z.Y.; Liu, X.Y. Support vector machines for tree species identification using LiDAR-derived structure and intensity variables. *Geocarto Int.* **2013**, *28*, 364–378. [[CrossRef](#)]
17. Guan, H.Y.; Yu, Y.T.; Ji, Z.; Li, J.; Zhang, Q. Deep learning-based tree classification using mobile LiDAR data. *Remote Sens. Lett.* **2015**, *6*, 864–873. [[CrossRef](#)]
18. Zou, X.H.; Cheng, M.; Wang, C.; Xia, Y.; Li, J. Tree classification in complex forest point clouds based on deep learning. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 2360–2364. [[CrossRef](#)]
19. Wang, Y.T.; Wang, J.; Chang, S.P.; Sun, L.; An, L.K.; Chen, Y.H.; Xu, J.Q. Classification of street tree species using UAV tilt photogrammetry. *Remote Sens.* **2021**, *13*, 216. [[CrossRef](#)]
20. Ghosh, A.; Fassnacht, F.E.; Joshi, P.K.; Koch, B. A framework for mapping tree species combining hyperspectral and LiDAR data: Role of selected classifiers and sensor across three spatial scales. *Int. J. Appl. Earth Obs. Geoinf.* **2014**, *26*, 49–63. [[CrossRef](#)]
21. Fassnacht, F.E.; Latifi, H.; Sterenczak, K.; Modzelewska, A.; Lefsky, M.; Waser, L.T.; Straub, C.; Ghosh, A. Review of studies on tree species classification from remotely sensed data. *Remote Sens. Environ.* **2016**, *186*, 64–87. [[CrossRef](#)]
22. Mayra, J.; Keski-Saari, S.; Kivinen, S.; Tanhuanpää, T.; Hurskainen, P.; Kullberg, P.; Poikolainen, L.; Viinikka, A.; Tuominen, S.; Kumpula, T.; et al. Tree species classification from airborne hyperspectral and LiDAR data using 3D convolutional neural networks. *Remote Sens. Environ.* **2021**, *256*, 112322. [[CrossRef](#)]

23. Korznikov, K.A.; Kislov, D.E.; Altman, J.; Dolezal, J.; Vozmishcheva, A.S.; Krestov, P.V. Using U-Net-like deep convolutional neural networks for precise tree recognition in very high resolution RGB (red, green, blue) satellite images. *Forests* **2021**, *12*, 66. [[CrossRef](#)]
24. Roth, S.I.B.; Leiterer, R.; Volpi, M.; Celio, E.; Schaepman, M.E.; Joerg, P.C. Automated detection of individual clove trees for yield quantification in northeastern madagascar based on multi-spectral satellite data. *Remote Sens. Environ.* **2019**, *221*, 144–156. [[CrossRef](#)]
25. Somers, B.; Asner, G.P. Tree species mapping in tropical forests using multi-temporal imaging spectroscopy: Wavelength adaptive spectral mixture analysis. *Int. J. Appl. Earth Obs. Geoinf.* **2014**, *31*, 57–66. [[CrossRef](#)]
26. Lee, J.; Cai, X.H.; Lellmann, J.; Dalponte, M.; Malhi, Y.; Butt, N.; Morecroft, M.; Schonlieb, C.-B.; Coomes, D.A. Individual tree species classification from airborne multisensor imagery using robust PCA. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *9*, 2554–2567. [[CrossRef](#)]
27. Le Louarn, M.; Clergeau, P.; Briche, E.; Deschamps-Cottin, M. “Kill two birds with one stone”: Urban tree species classification using bi-temporal pleiades images to study nesting preferences of an invasive bird. *Remote Sens.* **2017**, *9*, 916. [[CrossRef](#)]
28. Mishra, N.B.; Mainali, K.P.; Shrestha, B.B.; Radenz, J.; Karki, D. Species-level vegetation mapping in a himalayan treeline ecotone using unmanned aerial system (UAS) imagery. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 445. [[CrossRef](#)]
29. Immitzer, M.; Atzberger, C.; Koukal, T. Tree species classification with random forest using very high spatial resolution 8-band WorldView-2 satellite data. *Remote Sens.* **2012**, *4*, 2661–2693. [[CrossRef](#)]
30. Harikumar, A.; Paris, C.; Bovolo, F.; Bruzzone, L. A crown quantization-based approach to tree-species classification using high-density airborne laser scanning data. *IEEE Trans. Geosci Remote Sens.* **2021**, *59*, 4444–4453. [[CrossRef](#)]
31. Kuzmin, A.; Korhonen, L.; Kivinen, S.; Hurskainen, P.; Korpelainen, P.; Tanhuanpaa, T.; Maltamo, M.; Vihervaara, P.; Kumpula, T. Detection of european aspen (*Populus tremula* L.) based on an unmanned aerial vehicle approach in boreal forests. *Remote Sens.* **2021**, *13*, 1723. [[CrossRef](#)]
32. Tang, J.G.; Li, S.B.; Liu, P. A review of lane detection methods based on deep learning. *Pattern Recognition.* **2021**, *111*, 107623. [[CrossRef](#)]
33. Wolf, N. Object features for pixel-based classification of urban areas comparing different machine learning algorithms. *Photogramm. Fernerkund. Geoinf.* **2013**, *3*, 149–161. [[CrossRef](#)]
34. Zhou, J.H.; Qin, J.; Gao, K.; Leng, H.B. SVM-based soft classification of urban tree species using very high-spatial resolution remote-sensing imagery. *Int. J. Remote Sens.* **2016**, *37*, 2541–2559. [[CrossRef](#)]
35. Dalponte, M.; Ene, L.T.; Marconcini, M.; Gobakken, T.; Næsset, E. Semi-supervised SVM for individual tree crown species classification. *ISPRS J. Photogramm. Remote Sens.* **2015**, *110*, 77–87. [[CrossRef](#)]
36. Shelhamer, E.; Long, J.; Darrell, T. Fully convolutional networks for semantic segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 640–651. [[CrossRef](#)] [[PubMed](#)]
37. Lecun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [[CrossRef](#)]
38. Liu, B.; Yu, X.C.; Yu, A.Z.; Wan, G. Deep convolutional recurrent neural network with transfer learning for hyperspectral image classification. *J. Appl. Remote Sens.* **2018**, *12*, 17. [[CrossRef](#)]
39. Chen, Y.S.; Jiang, H.L.; Li, C.Y.; Jia, X.P.; Ghamisi, P. Deep feature extraction and classification of hyperspectral images based on convolutional neural networks. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 6232–6251. [[CrossRef](#)]
40. Makantasis, K.; Karantzalos, K.; Doulamis, A.; Doulamis, N. Deep supervised learning for hyperspectral data classification through convolutional neural networks. In Proceedings of the 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Milan, Italy, 26–31 July 2015; pp. 4959–4962.
41. Alipourfard, T.; Arefi, H.; Mahmoudi, S. A novel deep learning framework by combination of subspace-based feature extraction and convolutional neural networks for hyperspectral images classification. In Proceedings of the 38th IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Valencia, Spain, 22–27 July 2018; pp. 4780–4783.
42. Nezami, S.; Khoramshahi, E.; Nevalainen, O.; Pölonen, I.; Honkavaara, E. Tree species classification of drone hyperspectral and RGB imagery with deep learning convolutional neural networks. *Remote Sens.* **2020**, *12*, 1070. [[CrossRef](#)]
43. Sun, Y.; Huang, J.F.; Ao, Z.R.; Lao, D.Z.; Xin, Q.C. Deep Learning approaches for the mapping of tree species diversity in a tropical wetland using airborne LiDAR and high-spatial-resolution remote sensing images. *Forests* **2019**, *10*, 1047. [[CrossRef](#)]
44. Fricker, G.A.; Ventura, J.D.; Wolf, J.A.; North, M.P.; Davis, F.W.; Franklin, J. A convolutional neural network classifier identifies tree species in mixed-conifer forest from hyperspectral imagery. *Remote Sens.* **2019**, *11*, 2326. [[CrossRef](#)]
45. Ferreira, M.P.; Almeida, D.R.A.D.; Papa, D.D.A.; Minervino, J.B.S.; Veras, H.F.P.; Formighieri, A.; Santos, C.A.N.; Ferreira, M.A.D.; Figueiredo, E.O.; Ferreira, E.J.L. Individual tree detection and species classification of Amazonian palms using UAV images and deep learning. *Forest. Ecol. Manag.* **2020**, *475*, 118397. [[CrossRef](#)]
46. 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 October 2015; pp. 234–241.
47. Zhang, W.; Tang, P.; Zhao, L.J. Fast and accurate land cover classification on medium resolution remote sensing images using segmentation models. *Int. J. Remote Sens.* **2021**, *42*, 3277–3301. [[CrossRef](#)]
48. Giang, T.L.; Dang, K.B.; Le, Q.T.; Nguyen, V.G.; Tong, S.S.; Pham, V. U-Net convolutional networks for mining land cover classification based on high-resolution UAV imagery. *IEEE Access* **2020**, *8*, 186257–186273. [[CrossRef](#)]

49. Wagner, F.H.; Sanchez, A.; Tarabalka, Y.; Lotte, R.G.; Ferreira, M.P.; Aidar, M.P.M.; Gloor, E.; Phillips, O.L.; Aragao, L. Using the U-Net convolutional network to map forest types and disturbance in the Atlantic rainforest with very high resolution images. *Remote. Sens. Ecol. Conserv.* **2019**, *5*, 360–375. [[CrossRef](#)]
50. Schiefer, F.; Kattenborn, T.; Frick, A.; Frey, J.; Schall, P.; Koch, B.; Schmidlein, S. Mapping forest tree species in high resolution UAV-based RGB-imagery by means of convolutional neural networks. *ISPRS J. Photogramm. Remote Sens.* **2020**, *170*, 205–215. [[CrossRef](#)]
51. Yang, T.J.; Song, J.K.; Li, L.; Tang, Q. Improving brain tumor segmentation on MRI based on the deep U-Net and residual units. *J. X-ray Sci. Technol.* **2020**, *28*, 95–110. [[CrossRef](#)]
52. He, K.M.; Zhang, X.Y.; Ren, S.Q.; Sun, J. Deep residual learning for image recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 30 June 2016; pp. 770–778.
53. Shia, W.C.; Chen, D.R. Classification of malignant tumors in breast ultrasound using a pretrained deep residual network model and support vector machine. *Comput. Med. Imag. Grap.* **2021**, *87*, 101829. [[CrossRef](#)]
54. Dutta, D.; Chen, G.; Chen, C.; Gagne, S.A.; Li, C.L.; Rogers, C.; Matthews, C. Detecting plant invasion in urban parks with aerial image time series and residual neural network. *Remote Sens.* **2020**, *12*, 3493. [[CrossRef](#)]
55. Jing, L.; Cheng, Q.M. Two improvement schemes of PAN modulation fusion methods for spectral distortion minimization. *Int. J. Remote Sens.* **2009**, *30*, 2119–2131. [[CrossRef](#)]
56. Jing, L.; Hu, B.; Li, J.; Noland, T.; Guo, H. Automated tree crown delineation from Imagery based on morphological techniques. *IOP Conf. Ser. Earth Environ. Sci.* **2014**, *17*, 012066. [[CrossRef](#)]
57. Pan, Z.; Xu, J.; Guo, Y.; Hu, Y.; Wang, G. Deep Learning Segmentation and Classification for Urban Village Using a Worldview Satellite Image Based on U-Net. *Remote Sens.* **2020**, *12*, 1574. [[CrossRef](#)]
58. Wang, Y.; Zhang, D.F.; Dai, G.M. Classification of high resolution satellite images using improved U-Net. *Int. J. Appl. Math. Comput. Sci.* **2020**, *30*, 399–413.
59. Qi, W.; Wei, M.; Yang, W.; Xu, C.; Ma, C. Automatic Mapping of Landslides by the ResU-Net. *Remote Sens.* **2020**, *12*, 2487. [[CrossRef](#)]
60. Yi, Y.; Zhang, Z.; Zhang, W.; Zhang, C.; Li, W.; Zhao, T. Semantic Segmentation of Urban Buildings from VHR Remote Sensing Imagery Using a Deep Convolutional Neural Network. *Remote Sens.* **2019**, *11*, 1774. [[CrossRef](#)]
61. Qiu, L.; Jing, L.; Hu, B.; Li, H.; Tang, Y. A New Individual Tree Crown Delineation Method for High Resolution Multispectral Imagery. *Remote Sens.* **2020**, *12*, 585. [[CrossRef](#)]
62. Ghiyamat, A.; Shafri, H.Z.M. A review on hyperspectral remote sensing for homogeneous and heterogeneous forest biodiversity assessment. *Int. J. Remote Sens.* **2010**, *31*, 1837–1856. [[CrossRef](#)]
63. Ghiyamat, A.; Shafri, H.Z.M.; Mandiraji, G.A.; Shariff, A.R.M.; Mansor, S. Hyperspectral discrimination of tree species with different classifications using single- and multiple-endmember. *Int. J. Appl. Earth Obs. Geoinf.* **2013**, *23*, 177–191. [[CrossRef](#)]
64. Yao, W.; Krzystek, P.; Heurich, M. Tree species classification and estimation of stem volume and DBH based on single tree extraction by exploiting airborne full-waveform LiDAR data. *Remote Sens. Environ.* **2012**, *123*, 368–380. [[CrossRef](#)]
65. Wang, G.; Ren, P. Hyperspectral Image Classification with Feature-Oriented Adversarial Active Learning. *Remote Sens.* **2020**, *12*, 3879. [[CrossRef](#)]
66. Tao, Y.; Xu, M.; Zhong, Y.; Cheng, Y. GAN-Assisted Two-Stream Neural Network for High-Resolution Remote Sensing Image Classification. *Remote Sens.* **2017**, *9*, 1328. [[CrossRef](#)]
67. Wang, X.; Tan, K.; Du, Q.; Chen, Y.; Du, P.J. Caps-TripleGAN: GAN-Assisted CapsNet for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 7232–7245. [[CrossRef](#)]