




Article

Rubber Tree Recognition Based on UAV RGB Multi-Angle Imagery and Deep Learning

Yuying Liang^{1,†}, Yongke Sun^{1,†}, Weili Kou^{1,2}, Weiheng Xu^{1,2}, Juan Wang³, Qiuhua Wang⁴, Huan Wang¹ and Ning Lu^{1,2,*}

¹ College of Big Data and Intelligence Engineering, Southwest Forestry University, Kunming 650223, China; liangyy082844@163.com (Y.L.); sunyongke@swfu.edu.cn (Y.S.); kw_l_eric@163.com (W.K.); weihengx@gmail.com (W.X.); wanghuan@swfu.edu.cn (H.W.)

² Key Laboratory of National Forestry and Grassland Administration on Forestry and Ecological Big Data, Kunming 650223, China

³ Eco-Development Academy, Southwest Forestry University, Kunming 650223, China; schima@163.com

⁴ College of Civil Engineering, Southwest Forestry University, Kunming 650223, China; qhwang2010@swfu.edu.cn

* Correspondence: ninglu@swfu.edu.cn; Tel.: +86-0871-6386-3018

† These authors contributed equally to this work.

Abstract: The rubber tree (*Hevea brasiliensis*) is an important tree species for the production of natural latex, which is an essential raw material for varieties of industrial and non-industrial products. Rapid and accurate identification of the number of rubber trees not only plays an important role in predicting biomass and yield but also is beneficial to estimating carbon sinks and promoting the sustainable development of rubber plantations. However, the existing recognition methods based on canopy characteristic segmentation are not suitable for detecting individual rubber trees due to their high canopy coverage and similar crown structure. Fortunately, rubber trees have a defoliation period of about 40 days, which makes their trunks clearly visible in high-resolution RGB images. Therefore, this study employed an unmanned aerial vehicle (UAV) equipped with an RGB camera to acquire high-resolution images of rubber plantations from three observation angles (-90° , -60° , 45°) and two flight directions (SN: perpendicular to the rubber planting row, and WE: parallel to rubber planting rows) during the deciduous period. Four convolutional neural networks (multi-scale attention network, MAnet; Unet++; Unet; pyramid scene parsing network, PSPnet) were utilized to explore observation angles and directions beneficial for rubber tree trunk identification and counting. The results indicate that Unet++ achieved the best recognition accuracy (precision = 0.979, recall = 0.919, F-measure = 94.7%) with an observation angle of -60° and flight mode of SN among the four deep learning algorithms. This research provides a new idea for tree trunk identification by multi-angle observation of forests in specific phenological periods.

Keywords: rubber plantation; UAV; deep learning; defoliation period; multi-angle images; recognition



Citation: Liang, Y.; Sun, Y.; Kou, W.; Xu, W.; Wang, J.; Wang, Q.; Wang, H.; Lu, N. Rubber Tree Recognition Based on UAV RGB Multi-Angle Imagery and Deep Learning. *Drones* **2023**, *7*, 547. <https://doi.org/10.3390/drones7090547>

Academic Editor: Diego

González-Aguilera

Received: 10 July 2023

Revised: 17 August 2023

Accepted: 22 August 2023

Published: 24 August 2023



Copyright: © 2023 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

Since natural latex is an important raw material for fiberboard, tires, and plywood, the rubber tree (*Hevea brasiliensis*) is widely spread in tropical regions such as Southeast Asia and the tropical region of China [1]. Rubber plantations have significantly impacted local species diversity, carbon storage, land use conditions, the livelihoods of local communities, and government revenue [2–4]. Thus, the growth monitoring of rubber plantations has attracted great attention. Although a number of research studies have reported change detection, disease detection, stand age, biomass and carbon, and leaf area index (LAI) estimation in rubber plantations [5], the automatic rubber tree counting method based on remote sensing imagery still remains unclear due to the high canopy coverage. It is undeniable that accurately identifying the number of rubber trees is essential for predicting

the yield, analyzing the driving force, and understanding the carbon storage and carbon sequestration potential of rubber plantations in tropical areas.

The traditional rubber tree identification method is mainly manual counting in the field, which is time-consuming and labor-intensive, and it is difficult to apply large-scale rubber plantation monitoring. Nowadays, remote sensing techniques with the advantages of obtaining large-scale, rapid, objective, and accurate forest information have been proven to be an effective approach to forest growth monitoring [5,6]. Although previous studies have obtained high estimation accuracy in large-scale forest monitoring with satellite data, they could not meet the requirement of forest growth monitoring at the tree level due to their low spatial-temporal resolution and inflexible maneuverability, and especially their difficult operation when monitoring specific times and locations [7]. Compared to satellite platforms, unmanned aerial vehicles (UAVs), with their flexible revisit time and easy operation, have become a new cost-effective manner to obtain remote sensing data [8]. UAVs have been popular in various applications since 2010 [9], especially in precision agriculture, forestry, and photogrammetry. UAV-based images acquired from user-defined sensors, flight height, and flight times have been used to extract various forest parameters. For example, Han et al. [10] proposed a real-time orthophoto mosaicking-based tree counting framework to detect trees using sequential aerial images, which is very effective for the fast detection of large areas. Vélez et al. [11] proposed a novel technique using planar area and ground shadows calculated from UAV RGB imagery to estimate pistachio tree (*Pistacia vera* L.) canopy volume. Gan et al. [12] demonstrated the favorable and robust performance of the Detectree2 method for tree crown detection and canopy width extraction in high-mountain temperate deciduous forests with UAV imagery. These studies indicated that UAV-based high-resolution RGB imagery has great potential in retrieving the biophysical parameters of forests.

In recent years, image recognition based on deep learning (DL) algorithms has provided an outstanding technique in identifying and detecting observed objects, which has been widely applied in crop automatic counting, such as rice (*Oriza sativa* L.) panicle counting [13], wheat (*Triticum aestivum* L.) ears counting [14], corn (*Zea mays* L.) plant counting [15], sorghum (*Sorghum bicolor* L. Moench) panicle counting [16], and ramie (*Boehmeria nivea*) plant counting [17]. These studies found that the precision of crop recognition depends not only on the DL algorithms but also on the acquired image resolution. There is no doubt that higher-resolution images are beneficial for distinguishing crops and weeds in a complex field environment [18]. Recently, many experts employed a UAV equipped with RGB cameras to acquire high-resolution imagery for crop and tree counting. These RGB images combined with DL technology can successfully extract vegetation features and segment individual plants [19,20]. Additionally, Chen and Liao [21] provided fast region-based convolutional neural network (Fast R-CNN) architecture to detect and classify palm trees. Li et al. [22] developed a low-cost approach for counting rapeseed inflorescences using YOLOv5 with the convolutional block attention module based on UAV RGB imagery. A similar study was also conducted in the counting of rice seedlings and achieved a promised recognition accuracy [23].

The DL algorithm has greater advantages in processing digital images, video, voice, natural language, and other high-dimensional data [24]. The convolutional neural network (CNN) is a special deep feedforward network, which its components consist of an input layer, a convolutions layer, fully connected layers, and an output layer [25]. CNN excels at image-based tasks due to its advantages in learning the spatial structure of the pixels automatically [26], which produces multiple DL network architectures, such as Unet [27], Unet++ [28], multi-scale attention network (MANet) [29], and pyramid scene parsing network (PSPnet) [30]. In addition, these CNNs have been widely applied in medical image segmentation [28] and the target detection of vehicles [31], crops [13], and trees [32], which received an outstanding performance.

To achieve the automatic counting of trees from digital images, researchers have investigated the potential of the CNN approach for tree counting using high-resolution

images obtained from UAV-based sensors [33–35]. Previous studies have found that CNN achieved high recognition precision values of 0.95 and 0.92 for citrus [36] and navel orange trees [37], respectively. The improved CNN method called Fast R-CNN successfully achieved average accuracies of 99.8%, 100%, and 91.4% in identifying palm trees from young, mature, and mixed vegetation areas, respectively [21]. Although these studies have retrieved high accuracy in counting trees by CNN with UAV-based images, the proposed DL algorithms could not be directly migrated to other tree types because plant stand counting is affected by the plant size, leaf overlap, and variable spacing factors [38].

At present, canopy characteristics and canopy height models (CHM) extracted from UAV-based high-resolution images are mainly used for tree recognition. For example, Aliero et al. [39] detected and delineated the palm oil crown using convolution and morphological analysis based on the concept of crown geometry and vegetation response to radiation. An image threshold method was employed to create the palm oil tree centroids with an accuracy of 96.5%. Juan et al. [40] utilized the watershed algorithm to detect a single tree by the height value difference between pixels from a two-dimensional gray image of CHM. These studies mainly focus on canopy characteristic extraction in sparse forest environments; few studies pay more attention to the automatic counting of rubber trees using UAV-based RGB images due to high canopy coverage. Additionally, the main trunk of a rubber tree will split into multiple secondary trunks during the growth stage [41], so there is no significant difference in rubber tree crown architecture during the leaf flourishing period, making the current tree identification methods based on canopy characteristics not suitable for rubber tree recognition. Even though some studies attempted to segment tree crowns using ground-based mobile LiDAR, there are still challenges in distinguishing individual rubber trees in surroundings with self-occluded canopy architectures [42]. Fortunately, the rubber tree trunk can be observed with high-resolution RGB images acquired from deciduous periods because of special trunk characteristics, including a bright color, distinctive shape, and large size compared to the surrounding background. However, DL methods combining trunk-based features in identifying and counting rubber trees remain unclear.

To address the aforementioned issues, this study proposed an approach based on CNN to cope with the challenge of estimating the number of rubber trees in highly dense plantations from UAV-based RGB images. Given that it is difficult to identify the number of rubber trees from the canopy characteristics during the flourishing period, this study verifies the possibility of detecting rubber tree trunks by using UAV-based RGB images acquired from multiple observation angles and flight directions in the deciduous period. Four CNNs (MAnet; Unet++; Unet; PSPnet) were used to explore beneficial observation angles and directions for rubber tree trunk identification and counting. Furthermore, this study was the first attempt to detect rubber tree trunks by combining DL methods with UAV-based RGB images.

2. Materials and Methods

2.1. Study Area

The experiment was performed in Jinghong city located in Xishuangbanna Dai Autonomous Prefecture, Yunnan, China, in March 2021, at the coordinates 100°25' E–101°31' E longitude, 21°27' N–22°36' N latitude (Figure 1). Jinghong belongs to a humid tropical monsoon climate with obvious dry and wet seasons. Its annual precipitation and annual average temperature are 1200–1700 mm and 23.5 °C, respectively. Rubber plantations have become an important economical species in this region. All rubber trees in this study area were planted in 2000, with a row distance of 7–10 m and column distance of 1.5–2.5 m according to planting field inventory. Rubber tree cultivation exhibits distinct row and column patterns except for the position of corners. The field investigation and flight campaign were conducted during the deciduous period of rubber plantations, when only tree trunks and branches remain on the ground. This study area covered an area of 142.6 ha, planted with 614 rubber plants.

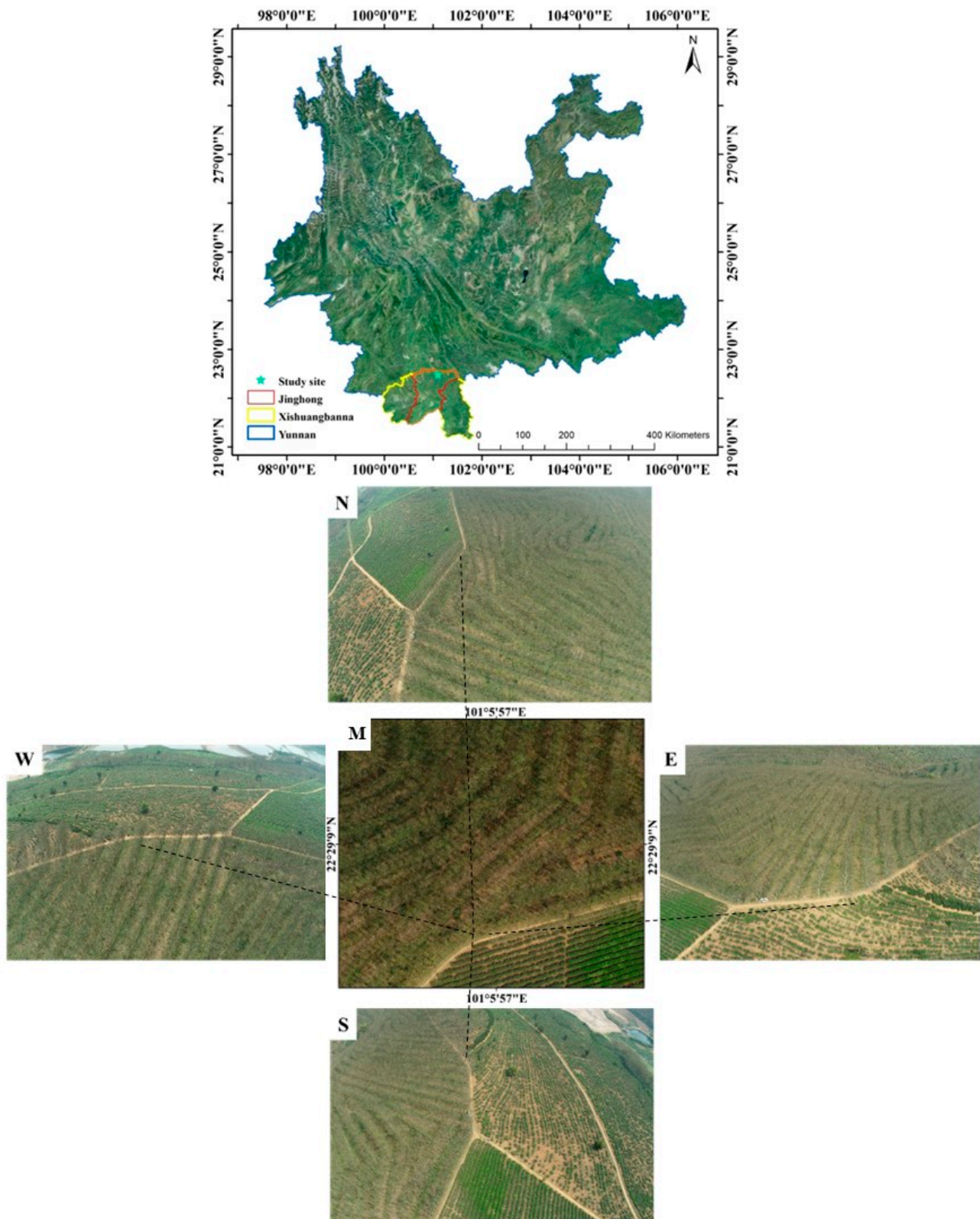


Figure 1. The location of the study area and demonstration of the sampling images acquired from different directions (E: east, S: south, W: west, N: north, M: nadir observation) at the same position.

2.2. Data Acquisition and Preprocessing

2.2.1. Data Acquisition

UAV-based RGB images were collected by the DJI Phantom 4 RTK after the leaves of rubber trees fell in March 2021. The UAV was equipped with an RGB camera with a resolution of 5472×3648 pixels. Flight campaigns were conducted in clear weather from 11:00 a.m. to 1:00 p.m. local time. To generate an orthophoto image covering the study area, a predefined flight path was designed to acquire imagery with 80% forward overlapping and 70% side overlapping. Multi-angle RGB images of rubber trees at a flight altitude of 100 m were acquired from three observation angles (-45° , -60° , -90°). All flight parameters related to the flight plan were set using the DJI GS RTK app on the DJI remote controller. In addition, 75 rubber trees in 3 regions of interest (ROI, $20 \text{ m} \times 25 \text{ m}$) of this study area were measured by the ZHD V200 (RTK, Guangzhou Hi-Target Navigation Tech Co., Ltd., Guangzhou, China) to obtain the location, which was utilized to evaluate the recognition accuracy of rubber tree trunks with DL techniques.

2.2.2. UAV Image Processing

The high-quality UAV-based RGB images were used to generate a digital orthophoto map (DOM) with Agisoft Metashape software (<https://www.agisoft.com/> accessed on 17 October 2022). Firstly, the generated DOM from RGB original images was utilized as a reference to visually interpret rubber trees within the study area. Secondly, DL algorithms were employed to generate binary mask images of the rubber trees based on original images. To integrate the identification results of individual images of rubber trees and cover a large study area, the recognition resultant images of rubber trees were mosaicked to an orthophoto using Agisoft Metashape software. The detailed orthophoto generation process can be found in the research of Lu et al. [43]. Lastly, ArcGIS software was used to evaluate the identification performance of rubber trees with DL techniques in a visual interpretation method on two overlapping DOMs. A total of 614 trees were manually identified according to the generated orthophoto of the study area.

To meet the requirement of the DL architecture, UAV-based high-resolution RGB images were resized to 320×320 pixels to match the size of a rubber tree and DL algorithms. A total of 600 images including their horizontal flip images, vertical flip images, and brightness contrast change images were randomly divided into training images of 80% and validation images of 20% based on a dataset partitioning strategy from Ariza-Sentís et al. [44]. The ground truths of rubber tree trunks were labeled as a polygon by using the publicly available “Labelme” (<https://github.com/wkentaro/labelme> accessed on 10 October 2022) tool. Each labeled image has an additional JSON file containing the coordinates of the annotated bounding boxes.

2.3. Convolutional Neural Network

This study employed four convolutional neural networks (MANet, Unet++, Unet, PSP-net) to recognize rubber tree trunks using single images collected from different observation angles. To improve the efficiency of CNN, ResNet50 was utilized to optimize the network framework. The four CNN architectures used for this study are shown in Figure 2.

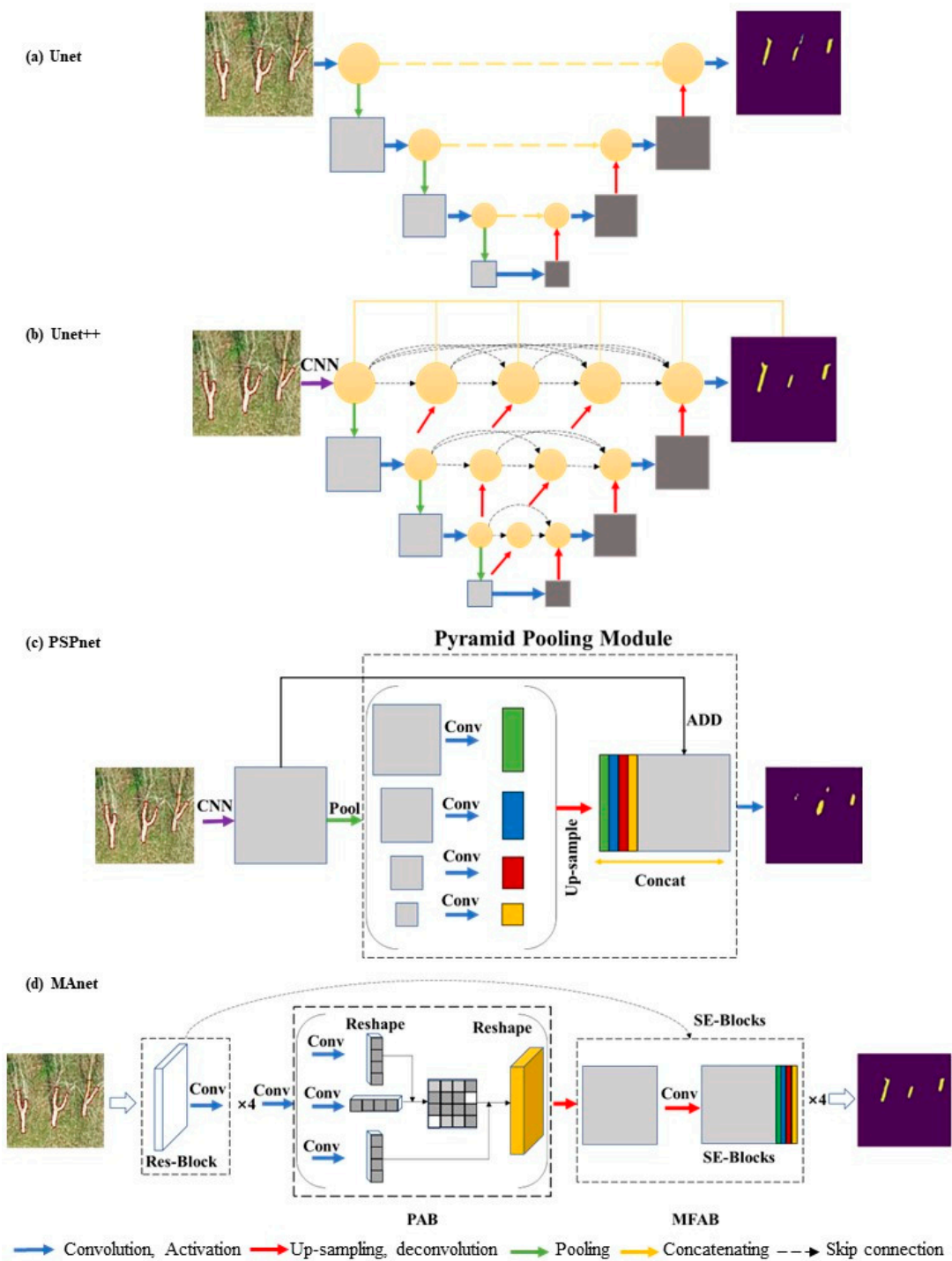


Figure 2. The architecture of four CNNs used for this study: (a) Unet, (b) Unet++, (c) pyramid scene parsing network (PSPnet), and (d) multi-scale attention network (MAnet). The left RGB images are the labeled images, and the right images are the detected images. Trees are highlighted in yellow, while the background is depicted in purple.

2.3.1. Unet and Unet++ Network

Unet is named based on its U-shaped structure and is a network with encoder and decoder architecture. This network has been applied excellently in medical image segmentation [27]. Figure 2a shows the overall architecture of Unet. This method, by optimally using the fully convolutional (FC) network’s key ideas (skip connections, adding convolutions and non-linearities between each up-sampling step), makes FC methods suitable

for semantic segmentation. In this way, the high-resolution feature mapping of the input was directly extracted from the encoder to the decoder network through a simple skip connection. However, the output result after encoding and decoding has the problem of incomplete correspondence with the original input image.

To confirm the appropriate depth of the network and to reduce the semantic gap between the encoder and decoder of the feature maps, Zhou et al. [28] proposed the Unet++ network (Figure 2b). This method can restore the extracted feature map size as the original image size after the decoding process and add more skip connections between encoding and decoding. Thus, the semantics of the encoder are closer to those of the decoder, and the network depth can be optimized.

2.3.2. Pyramid Scene Parsing Network

The pyramid scene parsing network (PSPnet) was proposed by Zhao et al. [30] to obtain appropriate global characteristics. This algorithm exploits the capability of global context information by different-region-based context aggregation with a pyramid pooling module. PSPnet provides a complete understanding of the scene, which can be used to predict the label, location, and shape of each scene element. Meanwhile, using both local and global information makes prediction more reliable. This method integrates the information from different scales and different subregions and plays a crucial role in the parsing task of complex scenes. In addition, more attention will be paid to the tiny targets in the scene to enhance the recognition effectiveness. This network uses four different pyramid scale pool modules to collect context information in the process of the training. Then, pyramid pooling module combines the characteristics of the pyramid of four different scales as the global prior. Finally, the global prior features with the original figure and the subsequent convolution are used to generate the final prediction figure (Figure 2c).

2.3.3. Multi-Scale Attention Network

The multi-scale attention network (MANet) proposed by Fan et al. [29] was first used to segment the liver and tumors (Figure 2d). This method models the feature interdependencies in spatial dimensions using a position-wise attention block, which is able to obtain the spatial dependencies between pixels in a global view. In addition, it has the great advantage of using a multi-scale fusion attention block to capture the channel dependencies among any feature maps by multi-scale semantic feature fusion via combining the high- and low-level feature maps. Thus, local features and their global correlations are adaptively integrated to improve the performance of the network.

2.3.4. ResNet50 Network

The residual network (ResNet) is one of the most popular image segmentation architectures [45]. ResNet50 has been proven to produce the best performance in the Timm open-source datasets due to its advantages in the subsampling part of feature map extraction [46]. Compared to ResNet50, the traditional method is to deepen or widen the network to improve the recognition accuracy, which leads to an increase in the number of hyperparameters and an increased computational burden for DL networks [47]. The ResNet's cross-layer connection addressed the issue of diminishing gradient return values, which will make training deep convolutional neural networks easier. In this study, all four networks used ResNet50 with the first four encoder layers to obtain the feature representation and optimize the accuracy without increasing the complexity parameters.

The Unet is a classic architecture that provides a baseline for image segmentation tasks, but it has some limitations in capturing detailed context information due to the limited receptive field. The Unet++ extends these capabilities by incorporating nested skip connections. The involved nested skip pathways increase its computational complexity compared to the original Unet. The PSPnet can more effectively capture context information at various scales, allowing it to handle objects of different sizes and improve segmentation accuracy, while it may not perform as well when it comes to fine-grained segmentation tasks.

Although the above methods can obtain context fusion information, they cannot describe the spatial and channel relationships among objects in a global view. The MANet considers the spatial and channel dependencies, which fuse the high- and low-level semantic features in the channel-wise dependencies [29]. Although all DL networks lack interpretability, most of the four DL techniques have achieved promising recognition accuracy in plant counting and tree detection [48,49].

2.4. Rubber Tree Trunks Identification from CNNs

To address the issue of rubber tree counting automatically and efficiently, this study proposes an approach based on DL techniques using UAV-based multi-angle RGB imagery. All DL algorithms described in Section 2.3 were implemented with the Pytorch DL framework. The workflow of the proposed approach is presented in Figure 3. Rubber tree images were obtained from a UAV-based RGB sensor with three multi-angle observations of -45° , -60° , and -90° during the defoliation period. These original images from different angles were spliced and generated a DOM. Meanwhile, the UAV original images were resized to smaller sizes and labeled manually using the “labelme” tool. The tree trunk label dataset was then divided into training (80%) and validation (20%) datasets. Four DL algorithms combined with the ResNet50 network were employed to build the rubber tree trunk recognition model. Finally, the detection mask images of rubber tree trunks were used to generate a DOM again and compare it with the visual interpretation results.

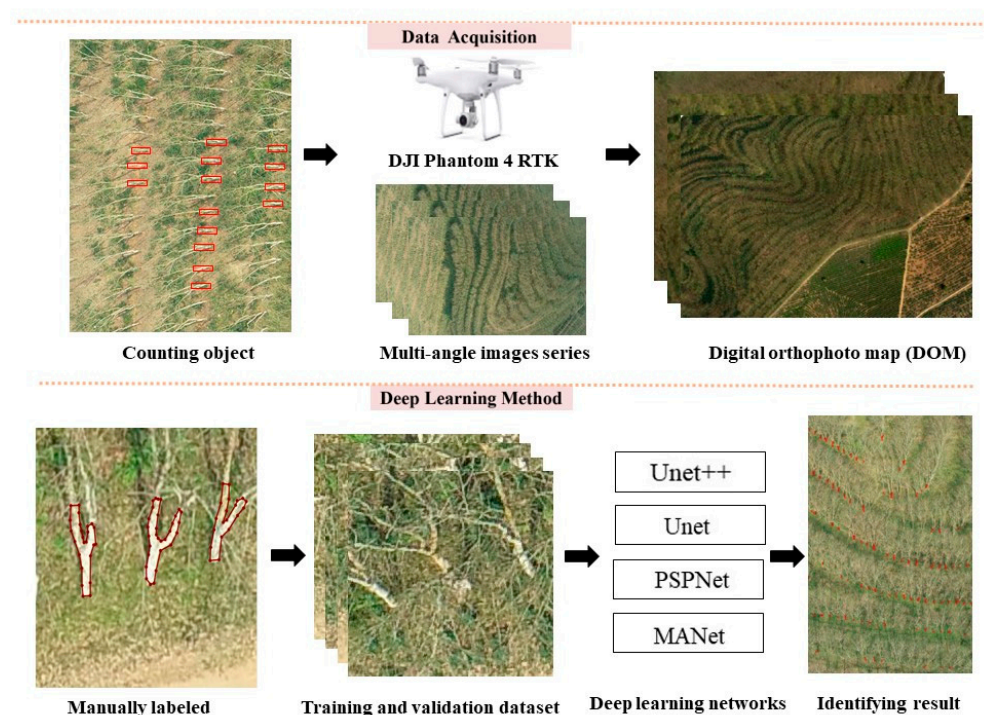


Figure 3. Experimental methodology and procedure of rubber tree recognition used in this study. Data acquisition included field measurement in rubber tree counting, multi-angle image series acquisition, and digital orthophoto map (DOM) generation. The procedure of using deep learning methods for rubber tree trunk recognition included the first step of manually labeling RGB images, the second step of dividing these labeled RGB images into training and validation datasets, followed by training four deep learning networks (Unet, Unet++, pyramid scene parsing network (PSPnet), and multi-scale attention network (MANet)) using the training dataset, and finally, evaluating the recognition accuracy using the validation dataset.

2.5. Accuracy Analysis

To evaluate the learning effect of the CNN network, this research employed intersection over union (IOU) loss and dice loss as two assessment indicators. IOU measures the

quality of target detection by calculating the ratio of the intersection and union between the prediction box and the real box (Figure 4). In addition, dice loss is used to evaluate the similarity between two labels (Equations (1) and (2)). It is worth noting that the smaller the values of IOU loss and dice loss, the better the detection effect.

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

$$\text{dice loss} = 1 - \text{dice} \quad (2)$$

where X is the predicted box, while Y represents the real box. The intersection between X and Y is denoted by $|X \cap Y|$. Additionally, $|X|$ and $|Y|$ refer to the pixel numbers of X and Y , respectively.

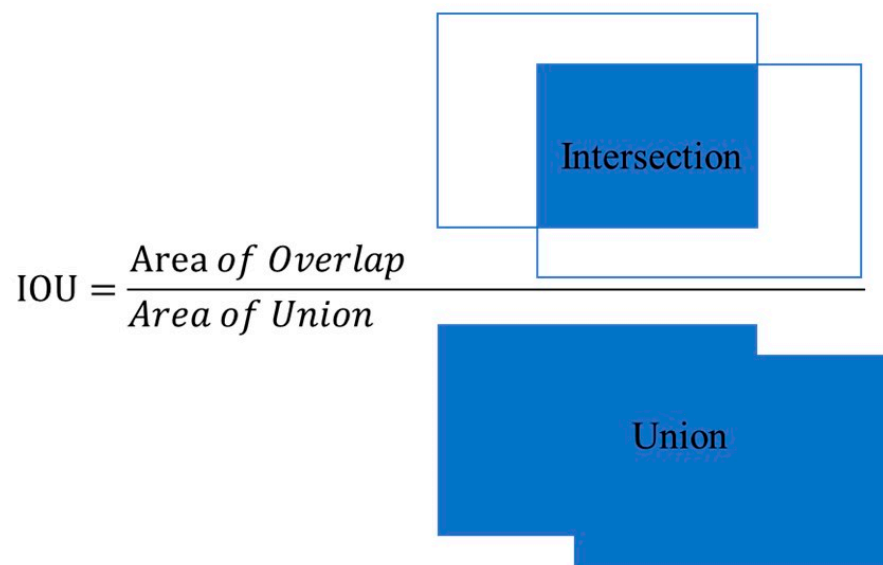


Figure 4. Illustration of the intersection over union (IOU) used to assess the performance metric of the CNN network in image segmentation.

The accuracy of detection was evaluated by three generally used quality factors: precision, recall, and F-measure [13]. Precision represents the percentage of examples classified as true positive examples that are positive examples. Recall represents the percentage of examples classified as true positive examples of all samples, and the F-measure is the evaluation index that integrates these two indexes and reflects the overall index comprehensively. The equations are defined as follows:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (3)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (4)$$

$$\text{F-measure} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \% \quad (5)$$

where true positive (TP) represents the correct classification of a region as a rubber tree, and false positive (FP) represents the error classification of a background region as a rubber tree. False negative (FN) indicates an incorrect classification of a tree as a background region.

3. Results

3.1. Comparison of Learning Effect of Different DL Algorithms

The numbers of epochs and loss function scores obtained from all the models were shown in Figure 5. During the training process, the dice loss value of the models gradually decreased as the number of iterations increased. Among the proposed networks, the dice loss value of the PSPnet decreased faster than the others and the dice loss tended to be stable around 3 epochs with its score of 0.997, while the other networks initially decreased slowly but sharply after 40 epochs and obtained the smallest dice loss between 42 and 46 epochs. It is noteworthy that the PSPnet had a slightly smaller dice loss score compared to the other networks. As for the IOU loss values, the dice loss value of the PSPnet snapped towards stability at around three epochs. Meanwhile, the other three DL algorithms, after experiencing an initial upward trend followed by stabilization, sharply decreased to their minimum values and remained constant after 40 epochs. Different from dice loss, all networks gained the same score for IOU loss.

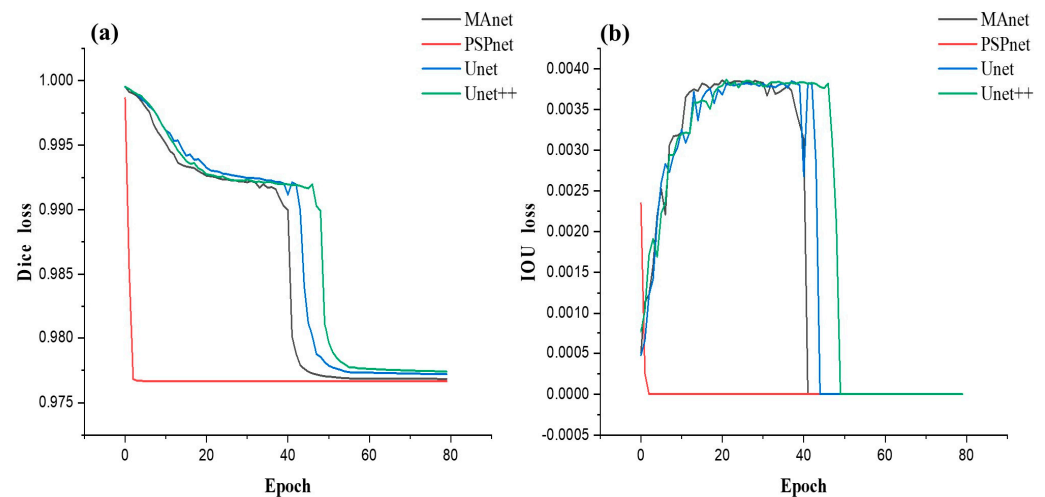


Figure 5. The comparison of dice loss (a) and IOU loss (b) derived from four distinct CNN algorithms: Unet, Unet++, pyramid scene parsing network (PSPnet), and multi-scale attention network (MAnet).

3.2. Identification and Counting of Rubber Trees

For visual comparison, this study utilized original images to visualize the detection results based on different recognition methods. Figure 6 shows the results of rubber trunk prediction based on four DL methods. Unet++ and Unet can more accurately recognize the trunk outline and identify more trees than the other two DL architectures. However, PSPnet and MAnet had obviously missing detection in some cases. In particular, some rubber tree trunk outlines identified by PSPnet were not integrated, which resulted in the possibility of multiple detections.

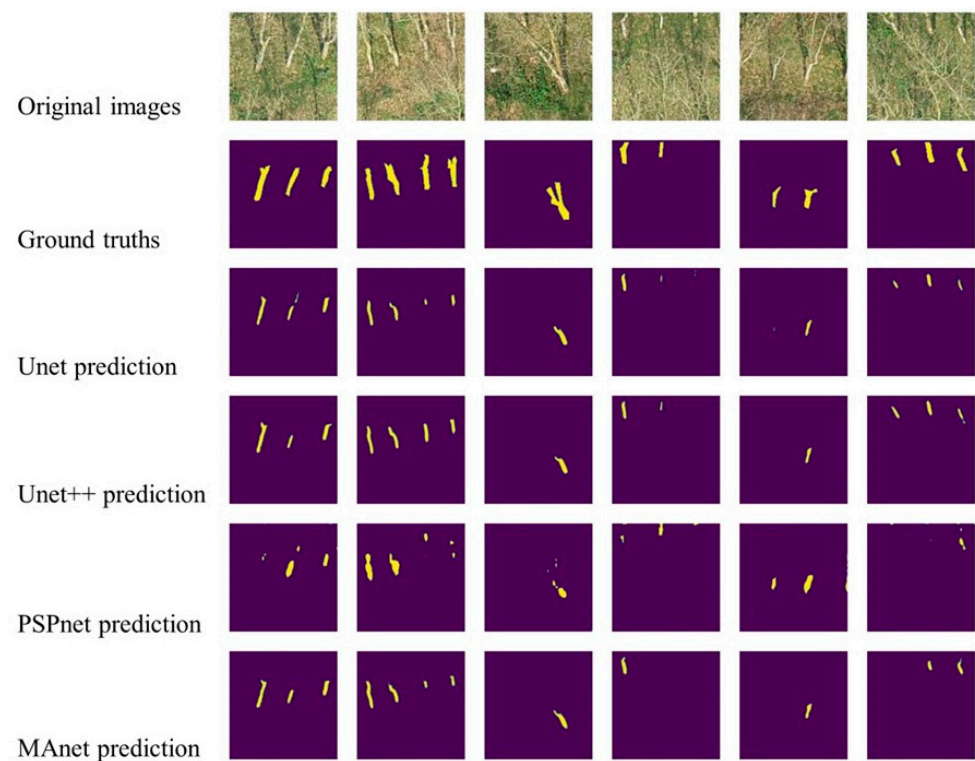


Figure 6. Detection results of rubber tree trunks using four distinct CNN algorithms: Unet, Unet++, pyramid scene parsing network (PSPnet), and multi-scale attention network (MAnet). Rubber tree trunks are highlighted in yellow, while the background is depicted in purple. “Original images” are the image sources in RGB format. “Ground truths” are the labeled images.

Different from previous recognition research based on a single UAV image, in this study, the number of rubber trees in the study area was obtained by generating the detected mask orthophoto images. As demonstrated in Figure 7, the ability to accurately identify rubber tree trunks was affected by the distance and angle of the UAV-based camera shot. The tree trunks were closer and the center of the photo can be identified with greater precision. As the UAV moves forward in flight, the rubber tree trunks in the back rows can also be identified row by row, which makes it possible to recognize rubber tree trunks in a large region by mosaicking identified binary mask images.

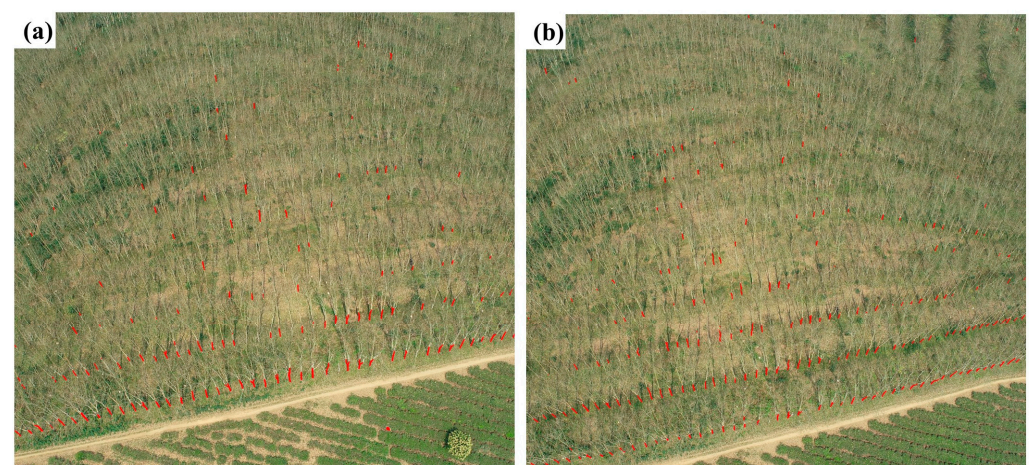


Figure 7. The rubber tree trunk identification results from the same flight path with adjacent images. (a) Current image captured by UAV, (b) next image captured by UAV.

To assess the capability of recognizing rubber tree trunks based on multi-angle original UAV images, this research compared the number of rubber tree trunks identified in the images captured from the same location. As Table 1 shows, it was found that 226 rubber trees could be identified using images at the observation angle of -90° ; 216 plants could be identified at -45° SN; 75 plants at -45° WE, 410 plants at -60° SN, and 159 plants at -60° WE. It is obvious that the -60° SN observation angle can retrieve more trunk information in images of a rubber forest.

Table 1. Comparison of the identification of rubber tree trunk counts in images from different observation angles at the same location.

Direction and Angle	Tree Number
-90°	226
-45° SN	216
-45° WE	75
-60° SN	410
-60° WE	159

Note: SN represents that the UAV acquired the images in the direction of flight perpendicular to the rubber planting rows; WE represents that the UAV acquired the images in the direction of flight parallel to the rubber planting rows; -45° , -60° , -90° represent the three observation angles.

3.3. Performance of DL Techniques with Multi-Angle Observation in Rubber Tree Identification

Table 2 shows the identification accuracy of rubber tree trunks in terms of precision, recall, and F-measure based on the proposed method by using images from different observation angles and orientations. It can be seen that the recognition precision of rubber tree trunks from different angles and directions was more than 0.93 by applying different recognition methods. But the recall values fluctuated greatly ranging from 0.088 to 0.919. The F-measure values also fluctuated greatly. The best F-measure values of the ResNet-Unet++, ResNet-Unet, ResNet-PSPnet, and ResNet-MANet were 94.7%, 90.0%, 42.3%, and 90.8%, respectively.

Table 2. The recognition accuracy comparison of rubber tree trunks using four distinct CNN algorithms (Unet, Unet++, PSPnet: pyramid scene parsing network, and MANet: multi-scale attention network) with UAV-based RGB images acquired from different observation angles and orientations. SN: images taken perpendicular to the rubber planting row flight, WE: images taken parallel to the rubber planting rows.

Method	Angle	Precision	Recall	F-Measure
ResNet-Unet++	-90°	0.979	0.901	93.8%
	-60° SN	0.979	0.919	94.7%
	-60° WE	0.977	0.687	80.7%
	-45° SN	0.972	0.862	91.4%
	-45° WE	0.979	0.756	85.3%
ResNet-Unet	-90°	0.983	0.824	90.0%
	-60° SN	0.974	0.730	83.4%
	-60° WE	0.977	0.546	70.0%
	-45° SN	0.961	0.764	85.1%
	-45° WE	0.970	0.427	59.3%
ResNet-PSPnet	-90°	0.965	0.090	16.4%
	-60° SN	0.967	0.241	38.6%
	-60° WE	0.968	0.148	25.7%
	-45° SN	0.971	0.270	42.3%
	-45° WE	0.931	0.088	16.1%
ResNet-MANet	-90°	0.985	0.842	90.8%
	-60° SN	0.985	0.839	90.5%
	-60° WE	0.974	0.611	75.1%
	-45° SN	0.950	0.711	81.3%
	-45° WE	0.982	0.531	68.9%

Our results also show that the ResNet-Unet++ achieved the best recognition accuracy and the highest recall with a UAV perpendicular to the rubber planting row flight mode (SN) and an observation angle of -60° (precision = 0.979; recall = 0.919; F-measure = 94.7%). In addition, the recall values of the rubber tree trunks identified by the images taken in parallel to the rubber planting rows (WE) were always lower than that of another shooting direction. By comparing the F-measure accuracy, the ResNet-Unet++ algorithm substantially outperformed the other DL algorithms, followed by ResNet-MANet (Figure 8).

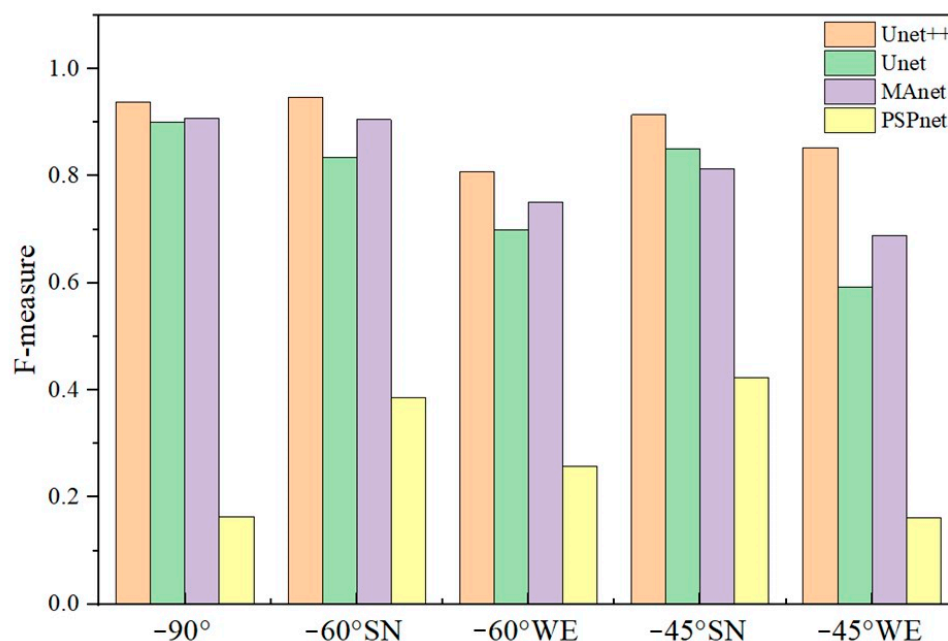


Figure 8. The F-measure results of rubber tree trunk identification with different observation angles and orientations using four CNN algorithms (Unet, Unet++, pyramid scene parsing network (PSPnet), and multi-scale attention network (MANet)). SN: images taken perpendicular to the rubber planting rows, WE: images taken parallel to the rubber planting rows.

To further validate the robustness and effectiveness of the proposed methods, the identified results were compared and analyzed with the field investigation of rubber trees (75 plants) obtained by the RTK, which represents the real position and the number of rubber trees (Table 3). The results show that ResNet-Unet++ performed the best (F-measure = 96.6%), while the ResNet-PSPnet had the lowest recognition accuracy among the four DL networks (F-measure = 35.2%) in this study area. The ResNet-Unet and ResNet-MANet achieved similar recognition accuracy (F-measure = 89.7% and F-measure = 87.6%, respectively).

Table 3. The accuracy assessment of four CNN algorithms (Unet, Unet++, pyramid scene parsing network (PSPnet), and multi-scale attention network (MANet)) compared with the field measured result. The UAV RGB images were taken perpendicular to the rubber planting rows with an observation angle of -60° .

Method	Detect	TP	FP	FN	Precision	Recall	F-Measure
ResNet-Unet++	74	72	2	3	0.973	0.960	96.6%
ResNet-Unet	61	61	0	14	1.000	0.813	89.7%
ResNet-PSPnet	16	16	0	59	1.000	0.213	35.2%
ResNet-MANet	62	60	2	15	0.968	0.800	87.6%

4. Discussion

4.1. Challenges in Canopy Segmentation of Single Trees in Rubber Forests

The conventional approach to tree identification based on canopy feature segmentation is not suitable for rubber forests. The main reasons can be attributed to two aspects: On the one hand, rubber forests have a high canopy coverage rate and high canopy overlap during their flourishing period, making it difficult for existing canopy segmentation methods to distinguish individual trees. Previous studies have focused on tree recognition of coniferous forests or orchards using UAV-based high-resolution RGB images of tree canopy [33,50,51]. These characteristics of tree canopy related to sparse growth scenarios and non-overlapping tree crowns are significantly different from rubber forests (Figure 9a). On the other hand, most rubber trees have more than one crown, which leads to large recognition errors in the crown-based segmentation algorithm. As shown in Figure 9b, the trunk of a rubber tree will generate multiple branches when it grows to a certain height (about 3 m). This means a rubber tree has one to three corresponding crowns, which makes it a great challenge to count the number of trees from UAV-based canopy RGB images.

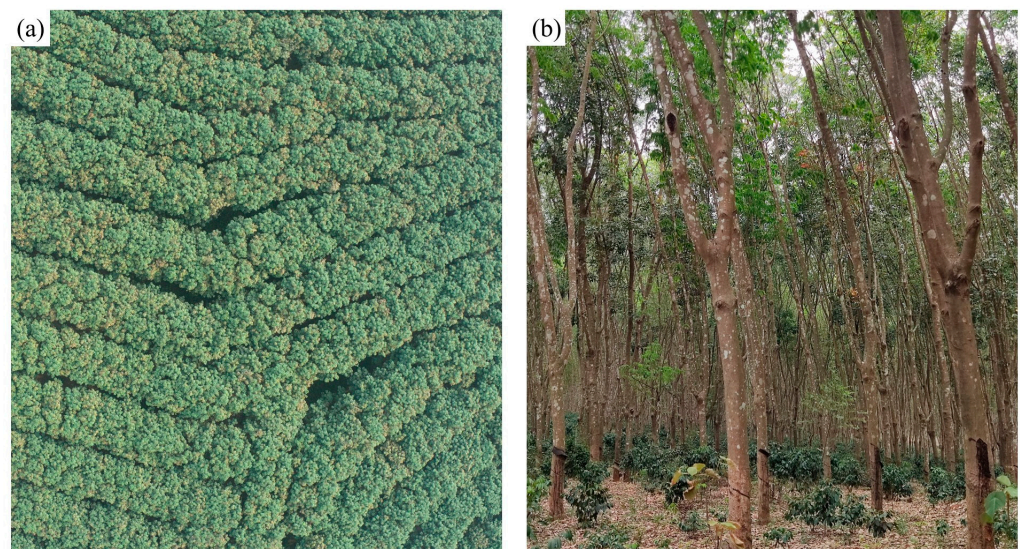


Figure 9. Images of the rubber tree crown and trunk in the stationary phase. (a) Canopy images of rubber trees acquired from UAV-based RGB sensors; (b) trunks of rubber trees captured by a phone sensor.

Even though the optical sensor has no penetrability compared to light detection and ranging (LiDAR), the trunk of a rubber tree can be seen clearly in UAV-based RGB images without the cover of leaves during the deciduous period (Figure 10b). Given the special characteristics of rubber forests with a 40 days defoliation period per year, this study attempted to identify individual rubber tree trunks in UAV-based high-resolution RGB images with four DL algorithms. This idea is consistent with a previous study that utilized the phenological characteristics of different vegetation in mixed forests for tree crown recognition. Chadwick et al. [52] delineated and measured the regenerating conifer crowns located beneath deciduous overstorey using UAV imagery. They found that regenerating conifer crowns can be well recognized under leaf-off conditions. However, their study only utilized images acquired from a vertical view and did not evaluate the performance of images from different observation angles. Our results prove the feasibility of rubber tree recognition through the use of multi-angle UAV images. These findings open a new perspective for the detection of tall deciduous trees.

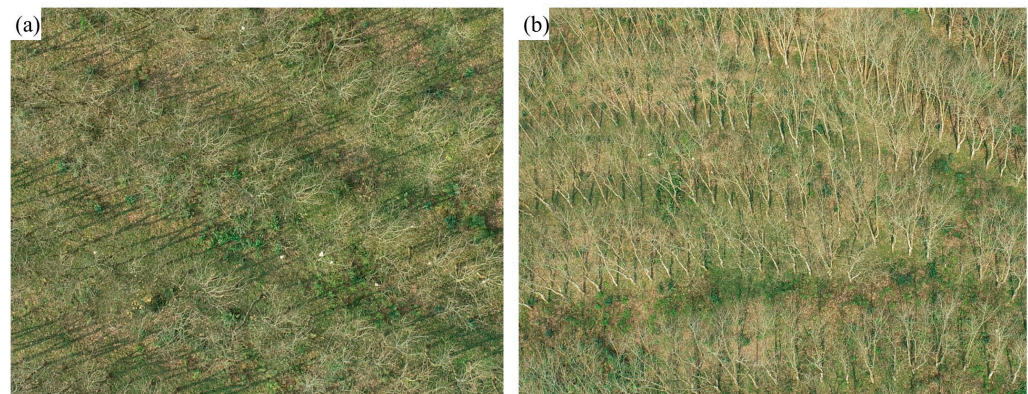


Figure 10. The bare crowns and trunks of rubber forest during the deciduous period. (a) Bare crowns of rubber tree images acquired from an observation angle of -90° ; (b) rubber tree trunk images acquired from an observation angle of -90° on the complex terrain.

4.2. Effects of UAV Multi-Angle Observations and Orientations on Tree Identification in a Rubber Forest

This study assessed the detection performance of rubber tree trunks using UAV-based high-resolution RGB images obtained from three observation angles (-90° , -60° , -45°) and two flight orientations (SN: perpendicular to rubber planting rows, and WE: parallel to rubber planting rows) during a defoliation period, with four deep learning techniques. For the two flight directions, we found that more rubber tree trunks could be detected in the SN mode than the WE mode based on the same research area (Table 3). This can be explained explicitly by the fact that the flight direction was highly correlated with the rubber forest planting orientation. According to the field investigation, the row spacing (7–10 m) of rubber planting is larger than the column spacing (1.5–2.5 m). Since the observation from south to north (SN) was perpendicular to the planting row direction of the rubber forest, more rubber trees can be seen in the UAV-based RGB images. Conversely, the smaller column spacing in the WE mode limited the visibility of tree trunks due to tree occlusion. Therefore, when the UAV collected images in the SN mode, the observed occlusion between the front and back rows was much smaller than the occlusion between adjacent trees, which could explain why the SN observation was superior to the WE observation.

For different observation angles, it was found that the observation angle of -60° was the best, followed by -90° , and then -45° , in detecting the number of rubber tree trunks at the same location. A reasonable explanation is that the view angle of -60° observed more rubber forest profile information than -90° , while -45° may be due to an excessive inclination angle in the shooting process, resulting in increased occlusion between the rubber trees. Despite a larger oblique angle and larger coverage of the image, the obtained oblique image has a lower spatial resolution when it is far away from the acquisition position [53]. Thus, using -45° observation angle images to generate orthophotos may lead to inaccurate calculations of the proportions and angles between the camera space and the real space. This research employed four DL methods due to the discrepancy in pixel resolution within an image [54]. This large offset resulted in multiple feature matching points to filter out in the calculation of the camera position and direction in different positions. As a result, the final mosaicked orthophoto became fuzzy due to the inclusion of more details of the tree branches, which lowered the recognition accuracy of rubber tree trunks. In addition, our finding is in agreement with a recent study by Hati and Singh [55], which reported that side-view images were more effective in identifying species and tracking growth with a deep neural network approach.

This study found that the oblique observation only slightly improved the recognition accuracy of the rubber tree in a deciduous period compared to the conventional vertical observation. The reasons may be as follows: Firstly, although the study area is a plantation with relatively fixed row and column distance, its arrangement is not absolutely

south–north or west–east. As shown in Figure 10b, different from crops, the row and column distribution of rubber tree planting is irregular and uneven due to complex terrain. Therefore, there was still a certain observation angle between the rubber forest and the aerial images acquired from the vertical observation angle of -90° , which yielded a promised recognition accuracy. Secondly, only three fixed observation angles were used in this study, which may not include an optimal oblique observation angle for rubber tree trunk identification. Furthermore, this study focused on recognizing rubber tree trunks using UAV-based RGB images collected during the defoliation period from a single observation angle. A previous study proved that the integration of multi-angle information improved the estimation accuracy of crop N nutrition status [53]. Thus, the performance of multi-angle observation information fusion for rubber tree trunk identification should be evaluated in further research.

4.3. Comparison of CNN Methods for Rubber Tree Identification

This research employed four DL methods (Unet++, Unet, PSPnet, MAnet) to recognize rubber tree trunks using UAV-based RGB imagery acquired from different observation directions and angles. In order to reduce the model degradation caused by the increase in network depth and minimize training errors, the residual network was utilized to improve the performance of DL models [56]. The study results show that Unet++ achieved the best recognition accuracy among the four DL algorithms for rubber tree trunk identification from UAV-based RGB images. This finding is in line with previous studies that reported that the Unet++ performed well in tasks like X-ray microscopy image segmentation, crowd counting, and liver CT image segmentation. [57–59]. The possible reason for this is that the Unet++ added more dense skip links to reduce the semantic gap between encoding and decoding. In addition, the Unet++ used network pruning to improve learning ability and reduced prediction time, which could be easier to deal with learning tasks than the Unet [28]. Although the MAnet achieved optimal target detection on datasets related to the medical field, it achieved moderate accuracy in rubber tree trunk identification. The plausible explanation is that compared to the medical images, the branches of the rubber forest in the study area are disordered and overlapped, and the branches in the obtained rubber tree images are irregular, which affected the final recognition precision [29]. In contrast, the PSPnet exhibited the poorest performance among the four deep learning techniques. The reason for this is that the pyramid pool module in the PSPnet algorithm expanded the identification boundary and ignored the edge details, resulting in unclear boundaries and blurry edges of the rubber tree trunks in the recognition result image. Even if the secondary filtering with a threshold of 0.2 was utilized to obtain a distinct rubber trunk profile with the PSPnet (Figure 11), the recognition effect of the PSPnet was not as good as the other algorithms used in this research. The secondary filtering filtered out image boundary pixels with smaller gray values to retrieve clear rubber trunks. However, it increased the loss value so that the detection accuracy was the lowest. This also resulted in significant false negatives in rubber tree trunk detection. A similar finding was also reported by Torbati-Sarraf et al. [57], who demonstrated that the Unet++ outperformed the PSPnet in medical image segmentation.

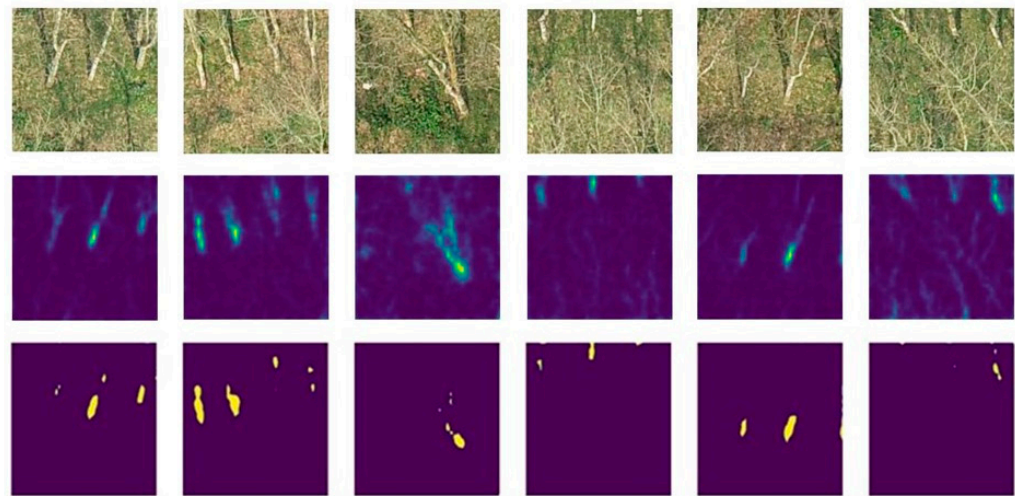


Figure 11. The comparison of the rubber tree trunk identification from the PSPnet algorithm before and after secondary filtering (**top** row is the image source, RGB input; **middle** row is the rubber tree trunk identification from the PSPnet algorithm directly; **bottom** row is the rubber tree trunk identification from the PSPnet algorithm with secondary filtering). Trees are highlighted in blue and yellow, while the background is depicted in purple.

4.4. Implications of Tree Trunk Identification and Counting

The homogeneous and highly overlapping canopy characteristics of rubber trees make existing crown-based segmentation methods for counting the number of individual trees impractical. Given the distinctive shape and large size of rubber tree trunks compared to small branches and the surrounding background during the defoliation period, it is easy to recognize tree trunks from UAV-based RGB images. This study proposes an innovative method to count the number of rubber trees by trunk recognition with DL techniques. Our findings demonstrate that high-resolution RGB images obtained by UAV combined with DL algorithms are feasible and effective detecting rubber tree trunks, which can be used to overcome the individual tree segmentation challenges in dense forests. Similarly, a recent study attempted to utilize tree trunk information extracted from UAV-based LiDAR for single-tree detection in a dense cedar plantation forest and achieved promising recognition accuracy [60]. Although their study differed from ours in terms of sensors, tree species characteristics, and acquisition time, the common idea of both is to extract tree trunks to calculate the number of trees. Compared to the expensive cost and complex data processing of LiDAR, this study provides a new perspective for tree trunk recognition using cost-effective UAV RGB imagery acquired during a deciduous period.

This study assessed the performance of UAV-based multi-angle RGB images for rubber tree trunk detection, but we did not consider the impact of tree shadows caused by sunlight from multi-angle observations. Although our flight campaigns were conducted from 11:00 am to 1:00 pm, shadows are still present in some angle images due to terrain influence. This study proposes a tree trunk identification approach based on the flight path, which makes the rubber tree trunk bright and has the advantage of mitigating shadow effects. The idea can be applied to extract the variables (e.g., tree height, crown diameter) of specific tree species in mixed forests based on tree phenological characteristics. In addition, the integration of multi-angle observation images and the impact of different resolution images on the accuracy of tree trunk recognition should also be explored. Meanwhile, the robustness and transferability of this proposed approach should be validated for different tree species and sites in future research.

5. Conclusions

This study evaluated the performance of four DL techniques in recognizing rubber tree trunks using UAV-based RGB images acquired from three view angles (-90° , -60° ,

45°) and two flight directions during a deciduous period. Our results demonstrate that the Unet++ outperformed the other three DL algorithms with the best recognition accuracy (precision = 0.979, recall = 0.919, F-measure = 94.7%) from a -60° observation angle and the SN (SN: perpendicular to the rubber planting rows) flight mode. Compared to conventional vertical observation, UAV-based images acquired from an oblique angle only slightly improved the accuracy of rubber tree identification due to its irregular planting caused by complex terrain. This implied that the oblique observation did not provide a significant advantage in detecting rubber tree trunks. However, this study demonstrated substantial potential in identifying rubber tree trunks using UAV-based high-resolution RGB images acquired during the defoliation period and in conjunction with DL algorithms. These findings provide new insights into tree identification using UAV-based RGB images acquired from specific growth stages and oblique observation angles based on tree characteristics.

Author Contributions: Y.L.: Investigation, data curation, writing—original draft; Y.S.: data curation, methodology; W.K.: conceptualization, methodology; J.W.: conceptualization; Q.W.: conceptualization; W.X.: supervision; H.W.: software, validation; N.L.: conceptualization, writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China (32160368, 32060320, 32260391), the Key Laboratory of National Forestry and Grassland Administration on Forestry and Ecological Big Data, Southwest Forestry University (2022-BDK-02), the Joint Special Project for Agriculture of Yunnan Province (202101BD070001-59), the Scientific Research Foundation for Ph.D. of Southwest Forestry University (110222004), the Research Foundation for Basic Research of Yunnan Province (202101AT070039), and the Youth Top Talents of Yunnan Ten Thousand Talents Program (YNWR-QNBJ-2019-270).

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Acknowledgments: We would like to thank Hongyan Lai, Maojia Gong, Xiong Yin, Yue Chen, Yuguo Zhang, Xiaoqing Li, LiMin Fuyang, and Guiliang Chen for their help in the data collection. We also thank the anonymous reviewers for their constructive comments.

Conflicts of Interest: The authors declare that they have no known competing financial interest or personal relationship that could have appeared to influence the work reported in this paper.

References

1. Chen, B.; Yun, T.; Ma, J.; Kou, W.; Li, H.; Yang, C.; Xiao, X.; Zhang, X.; Sun, R.; Xie, G.; et al. High-precision stand age data facilitate the estimation of rubber plantation biomass: A case study of Hainan Island, China. *Remote Sens.* **2020**, *12*, 3853. [[CrossRef](#)]
2. Liang, Y.; Kou, W.; Lai, H.; Wang, J.; Wang, Q.; Xu, W.; Wang, H.; Lu, N. Improved estimation of aboveground biomass in rubber plantations by fusing spectral and textural information from UAV-based RGB imagery. *Ecol. Indic.* **2022**, *142*, 109286. [[CrossRef](#)]
3. Tang, J.; Pang, J.; Chen, M.; Guo, X.; Zeng, R. Biomass and its estimation model of rubber plantations in Xishuangbanna, Southwest China. *Chin. J. Ecol.* **2009**, *28*, 1942–1948.
4. Kou, W.; Dong, J.; Xiao, X.; Hernandez, A.J.; Qin, Y.; Zhang, G.; Chen, B.; Lu, N.; Doughty, R. Expansion dynamics of deciduous rubber plantations in Xishuangbanna, China during 2000–2010. *GISci. Remote Sens.* **2018**, *55*, 905–925. [[CrossRef](#)]
5. Azizan, F.A.; Kiloos, A.M.; Astuti, I.S.; Abdul Aziz, A. Application of optical remote sensing in rubber plantations: A systematic review. *Remote Sens.* **2021**, *13*, 429. [[CrossRef](#)]
6. Kou, W.; Liang, C.; Wei, L.; Hernandez, A.; Yang, X. Phenology-based method for mapping tropical evergreen forests by integrating of MODIS and Landsat imagery. *Forests* **2017**, *8*, 34. [[CrossRef](#)]
7. Zhang, X.; Zhang, K.; Sun, Y.; Zhao, Y.; Zhuang, H.; Ban, W.; Chen, Y.; Fu, E.; Chen, S.; Liu, J.; et al. Combining spectral and texture features of UAS-based multispectral images for maize leaf area index estimation. *Remote Sens.* **2022**, *14*, 331. [[CrossRef](#)]
8. Wang, X.; Wang, Y.; Zhou, C.; Yin, L.; Feng, X. Urban forest monitoring based on multiple features at the single tree scale by UAV. *Urban For. Urban Green.* **2021**, *58*, 126958. [[CrossRef](#)]
9. Jin, W.; Ge, H.-L.; Du, H.-Q.; Xu, X.-J. A review on unmanned aerial vehicle remote sensing and its application. *Remote Sens. Inf.* **2009**, *1*, 88–92. [[CrossRef](#)]
10. Han, P.; Ma, C.; Chen, J.; Chen, L.; Bu, S.; Xu, S.; Zhao, Y.; Zhang, C.; Hagino, T. Fast tree detection and counting on UAVs for sequential aerial images with generating orthophoto mosaicing. *Remote Sens.* **2022**, *14*, 4113. [[CrossRef](#)]
11. Vélez, S.; Vacas, R.; Martín, H.; Ruano-Rosa, D.; Álvarez, S. A novel technique using planar area and ground shadows calculated from UAV RGB imagery to estimate pistachio tree (*Pistacia vera* L.) canopy volume. *Remote Sens.* **2022**, *14*, 6006. [[CrossRef](#)]

12. Gan, Y.; Wang, Q.; Iio, A. Tree crown detection and delineation in a temperate deciduous forest from UAV RGB imagery using deep learning approaches: Effects of spatial resolution and species characteristics. *Remote Sens.* **2023**, *15*, 778. [[CrossRef](#)]
13. Zhou, C.; Ye, H.; Hu, J.; Shi, X.; Hua, S.; Yue, J.; Xu, Z.; Yang, G. Automated counting of rice panicle by applying deep learning model to images from unmanned aerial vehicle platform. *Sensors* **2019**, *19*, 3106. [[CrossRef](#)] [[PubMed](#)]
14. Zaji, A.; Liu, Z.; Xiao, G.; Bhowmik, P.; Sangha, J.S.; Ruan, Y. Wheat spike localization and counting via hybrid UNet architectures. *Comput. Electron. Agric.* **2022**, *203*, 107439. [[CrossRef](#)]
15. Kitano, B.T.; Mendes, C.C.T.; Geus, A.R.; Oliveira, H.C.; Souza, J.R. Corn plant counting using deep learning and UAV images. *IEEE Geosci. Remote Sens. Lett.* **2019**, 1–5. [[CrossRef](#)]
16. Lin, Z.; Guo, W. Sorghum panicle detection and counting using unmanned aerial system images and deep learning. *Front. Plant Sci.* **2020**, *11*, 534853. [[CrossRef](#)]
17. Fu, H.-Y.; Yue, Y.-K.; Wang, W.; Liao, A.; Xu, M.-Z.; Gong, X.; She, W.; Cui, G.-X. Ramie plant counting based on UAV remote sensing technology and deep learning. *J. Nat. Fibers* **2023**, *20*, 2159610. [[CrossRef](#)]
18. Khan, S.; Tufail, M.; Khan, M.T.; Khan, Z.A.; Iqbal, J.; Alam, M. A novel semi-supervised framework for UAV based crop/weed classification. *PLoS ONE* **2021**, *16*, e0251008. [[CrossRef](#)]
19. Zhang, J.; Zhao, B.; Yang, C.; Shi, Y.; Liao, Q.; Zhou, G.; Wang, C.; Xie, T.; Jiang, Z.; Zhang, D.; et al. Rapeseed stand count estimation at leaf development stages with UAV imagery and convolutional neural networks. *Front. Plant Sci.* **2020**, *11*, 617. [[CrossRef](#)]
20. Jiang, X.; Wu, Z.; Han, S.; Yan, H.; Zhou, B.; Li, J. A multi-scale approach to detecting standing dead trees in UAV RGB images based on improved faster R-CNN. *PLoS ONE* **2023**, *18*, e0281084. [[CrossRef](#)]
21. Chen, Z.Y.; Liao, I.Y. Improved Fast R-CNN with fusion of optical and 3D data for robust palm tree detection in high resolution UAV images. *Int. J. Mach. Learn. Comput.* **2020**, *10*, 122–127. [[CrossRef](#)]
22. Li, J.; Li, Y.; Qiao, J.; Li, L.; Wang, X.; Yao, J.; Liao, G. Automatic counting of rapeseed inflorescences using deep learning method and UAV RGB imagery. *Front. Plant Sci.* **2023**, *14*, 1101143. [[CrossRef](#)] [[PubMed](#)]
23. Wu, J.; Yang, G.; Yang, X.; Xu, B.; Han, L.; Zhu, Y. Automatic counting of in situ rice seedlings from UAV images based on a deep fully convolutional neural network. *Remote Sens.* **2019**, *11*, 691. [[CrossRef](#)]
24. Zheng, Y.; Li, G.; Li, Y. Survey of application of deep learning in image recognition. *Comput. Eng. Appl.* **2019**, *55*, 20–36. [[CrossRef](#)]
25. Nakajima, K.; Tanaka, Y.; Katsura, K.; Yamaguchi, T.; Watanabe, T.; Shiraiwa, T. Biomass estimation of world rice (*Oryza sativa* L.) core collection based on the convolutional neural network and digital images of canopy. *Plant Prod. Sci.* **2023**, *26*, 187–196. [[CrossRef](#)]
26. Su, J.; Zhu, X.; Li, S.; Chen, W.-H. AI meets UAVs: A survey on AI empowered UAV perception systems for precision agriculture. *Neurocomputing* **2023**, *518*, 242–270. [[CrossRef](#)]
27. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention 2015*; Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F., Eds.; Springer International Publishing: Cham, Switzerland, 2015; Volume 9351, pp. 234–241. [[CrossRef](#)]
28. Zhou, Z.; Siddiquee, M.M.R.; Tajbakhsh, N.; Liang, J. *UNet++: A nested U-Net architecture for medical image segmentation*. *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*; Springer: Berlin/Heidelberg, Germany, 2018. [[CrossRef](#)]
29. Fan, T.; Wang, G.; Li, Y.; Wang, H. MA-Net: A multi-scale attention network for liver and tumor segmentation. *IEEE Access* **2020**, *8*, 179656–179665. [[CrossRef](#)]
30. Zhao, H.; Shi, J.; Qi, X.; Wang, X.; Jia, J. Pyramid scene parsing network. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017.
31. Momin, M.A.; Junos, M.H.; Mohd Khairuddin, A.S.; Abu Talip, M.S. Lightweight CNN model: Automated vehicle detection in aerial images. *Signal Image Video Process.* **2022**, *17*, 1209–1217. [[CrossRef](#)]
32. Csillik, O.; Cherbini, J.; Johnson, R.; Lyons, A.; Kelly, M. Identification of citrus trees from unmanned aerial vehicle imagery using convolutional neural networks. *Drones* **2018**, *2*, 39. [[CrossRef](#)]
33. Chen, G.; Shang, Y. Transformer for tree counting in aerial images. *Remote Sens.* **2022**, *14*, 476. [[CrossRef](#)]
34. Djerriri, K.; Ghabi, M.; Karoui, M.S.; Adjoudj, R. Palm trees counting in remote sensing imagery using regression convolutional neural network. In Proceedings of the IGARSS 2018—2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018; pp. 2627–2630. [[CrossRef](#)]
35. Li, W.; Fu, H.; Yu, L.; Cracknell, A. Deep learning based oil palm tree detection and counting for high-resolution remote sensing images. *Remote Sens.* **2017**, *9*, 22. [[CrossRef](#)]
36. Osco, L.P.; Arruda, M.d.S.d.; Marcato Junior, J.; da Silva, N.B.; Ramos, A.P.M.; Moryia, É.A.S.; Imai, N.N.; Pereira, D.R.; Creste, J.E.; Matsubara, E.T.; et al. A convolutional neural network approach for counting and geolocating citrus-trees in UAV multispectral imagery. *ISPRS J. Photogramm. Remote Sens.* **2020**, *160*, 97–106. [[CrossRef](#)]
37. Chen, Y.; Zhang, X.; Chen, X. Identification of navel orange trees based on deep learning algorithm YOLOv4. *Sci. Surv. Mapp.* **2022**, *47*, 135–144. [[CrossRef](#)]
38. Bai, Y.; Nie, C.; Wang, H.; Cheng, M.; Liu, S.; Yu, X.; Shao, M.; Wang, Z.; Wang, S.; Tuohuti, N.; et al. A fast and robust method for plant count in sunflower and maize at different seedling stages using high-resolution UAV RGB imagery. *Precis. Agric.* **2022**, *23*, 1720–1742. [[CrossRef](#)]

39. Aliero, M.M.; Bunza, R.M.; Al-Doksi, J. The usefulness of unmanned airborne vehicle (UAV) imagery for automated palm oil tree counting. *Res. J. For.* **2014**, *1*, 1–12.
40. Wang, J.; Zhang, C.; Chen, Q.; Li, H.; Peng, X.; Bai, M.; Xu, Z.; Liu, H.; Chen, Y. The method of extracting information of *Cunninghamia lanceolata* crown combined with RGB and LiDAR based on UAV. *J. Southwest For. Univ.* **2022**, *42*, 133–141.
41. Kouadio, Y.J.; Obouayeba, S.; N'guessan, A.A.; Vouli, B.B.N.B.; Soumahin, E.F. Agromorphological characterization of a rubber tree-teak agroforestry system in central Côte d'Ivoire. *Asian J. Res. Agric. For.* **2022**, *8*, 273–292. [[CrossRef](#)]
42. Yun, T.; Jiang, K.; Hou, H.; An, F.; Chen, B.; Li, W.; Xue, L. Rubber Tree Crown segmentation and property retrieval using ground-based mobile LiDAR after natural disturbances. *Remote Sens.* **2019**, *11*, 903. [[CrossRef](#)]
43. Lu, N.; Zhou, J.; Han, Z.; Li, D.; Cao, Q.; Yao, X.; Tian, Y.; Zhu, Y.; Cao, W.; Cheng, T. Improved estimation of aboveground biomass in wheat from RGB imagery and point cloud data acquired with a low-cost unmanned aerial vehicle system. *Plant Methods* **2019**, *15*, 17. [[CrossRef](#)]
44. Ariza-Sentís, M.; Valente, J.; Kooistra, L.; Kramer, H.; Múcher, S. Estimation of spinach (*Spinacia oleracea*) seed yield with 2D UAV data and deep learning. *Smart Agric. Technol.* **2023**, *3*, 100129. [[CrossRef](#)]
45. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA; 27–30 June 2016; pp. 770–778. [[CrossRef](#)]
46. Wightman, R.; Touvron, H.; Jégou, H.E. ResNet strikes back: An improved training procedure in timm. *arXiv* **2021**, arXiv:2110.00476.
47. Shafiq, M.; Gu, Z. Deep residual learning for image recognition: A survey. *Appl. Sci.* **2022**, *12*, 8972. [[CrossRef](#)]
48. Farjon, G.; Liu, H.; Edan, Y. Deep-learning-based counting methods, datasets, and applications in agriculture: A review. *Precis. Agric.* **2023**, *24*, 1683–1711. [[CrossRef](#)]
49. Hao, Z.; Lin, L.; Post, C.J.; Mikhailova, E.A.; Li, M.; Chen, Y.; Yu, K.; Liu, J. Automated tree-crown and height detection in a young forest plantation using mask region-based convolutional neural network (Mask R-CNN). *ISPRS J. Photogramm. Remote Sens.* **2021**, *178*, 112–123. [[CrossRef](#)]
50. Neupane, B.; Horanont, T.; Hung, N.D. Deep learning based banana plant detection and counting using high-resolution red-green-blue (RGB) images collected from unmanned aerial vehicle (UAV). *PLoS ONE* **2019**, *14*, e0223906. [[CrossRef](#)]
51. Zhu, M.; Zhou, Z.; Zhao, X.; Huang, D.; Jiang, Y.; Wu, Y.; Cui, L. Recognition and extraction method of single dragon fruit plant in Plateau-Canyon areas based on UAV remote sensing. *Trop. Geogr.* **2019**, *39*, 502–511. [[CrossRef](#)]
52. Chadwick, A.J.; Goodbody, T.R.H.; Coops, N.C.; Hervieux, A.; Bater, C.W.; Martens, L.A.; White, B.; Roeser, D. Automatic delineation and height measurement of regenerating conifer crowns under leaf-off conditions using UAV imagery. *Remote Sens.* **2020**, *12*, 4104. [[CrossRef](#)]
53. Lu, N.; Wang, W.; Zhang, Q.; Li, D.; Yao, X.; Tian, Y.; Zhu, Y.; Cao, W.; Baret, F.; Liu, S.; et al. Estimation of nitrogen nutrition status in winter wheat from unmanned aerial vehicle based multi-angular multispectral imagery. *Front. Plant Sci.* **2019**, *10*, 1601. [[CrossRef](#)]
54. Xu, H.; Zhou, X.; Huang, H.; Chen, M. Single tree structure parameter extraction of structure-from-motion with multi-view stereophotogrammetry. *Sci. Surv. Mapp.* **2018**, *43*, 108–114. [[CrossRef](#)]
55. Hati, A.J.; Singh, R.R. AI-driven pheno-parenting: A deep learning based plant phenotyping trait analysis model on a novel soilless farming dataset. *IEEE Access* **2023**, *11*, 35298–35314. [[CrossRef](#)]
56. Zhou, F.-Y.; Jin, L.; Dong, J. Review of convolutional neural network. *Chin. J. Comput.* **2017**, *40*, 23. [[CrossRef](#)]
57. Torbati-Sarraf, H.; Niverty, S.; Singh, R.; Barboza, D.; De Andrade, V.; Turaga, P.; Chawla, N. Machine-learning-based algorithms for automated image segmentation techniques of transmission X-ray microscopy (TXM). *JOM* **2021**, *73*, 2173–2184. [[CrossRef](#)]
58. Marcellino; Cenggoro, T.W.; Pardamean, B. UNET++ with scale pyramid for crowd counting. *ICIC Express Lett.* **2022**, *16*, 75–82. [[CrossRef](#)]
59. Li, J.; Liu, K.; Hu, Y.; Zhang, H.; Heidari, A.A.; Chen, H.; Zhang, W.; Algarni, A.D.; Elmannai, H. Eres-UNet++: Liver CT image segmentation based on high-efficiency channel attention and Res-UNet+. *Comput. Biol. Med.* **2023**, *158*, 106501. [[CrossRef](#)] [[PubMed](#)]
60. Zhang, Y.; Tan, Y.; Onda, Y.; Hashimoto, A.; Gomi, T.; Chiu, C.; Inokoshi, S. A tree detection method based on trunk point cloud section in dense plantation forest using drone LiDAR data. *For. Ecosyst.* **2023**, *10*, 100088. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.