



Article

Towards Amazon Forest Restoration: Automatic Detection of Species from UAV Imagery

Marks Melo Moura ^{1,*}, Luiz Eduardo Soares de Oliveira ², Carlos Roberto Sanquetta ¹, Alexis Bastos ³, Midhun Mohan ⁴ and Ana Paula Dalla Corte ¹

¹ Department of Forest Engineering, Federal University of Paraná, Av. Lothário Meissner, 900, Curitiba 80270-170, Brazil; sanquetta@ufpr.br (C.R.S.); anacorte@ufpr.br (A.P.D.C.)

² Department of Informatics, Federal University of Paraná, Av. Cel. Francisco H. dos Santos, 100, Curitiba 81530-000, Brazil; luiz.oliveira@ufpr.br

³ Cultural and Environmental Study Center of the Amazon Region—RIOTERRA, Rua Padre Chiquinho, 1651, Porto Velho 76803-786, Brazil; alexis@rioterra.org.br

⁴ Department of Geography, University of California-Berkeley, Berkeley, CA 94709, USA; mid_mohan@berkeley.edu

* Correspondence: marksmoura@ufpr.br

Abstract: Precise assessments of forest species' composition help analyze biodiversity patterns, estimate wood stocks, and improve carbon stock estimates. Therefore, the objective of this work was to evaluate the use of high-resolution images obtained from Unmanned Aerial Vehicle (UAV) for the identification of forest species in areas of forest regeneration in the Amazon. For this purpose, convolutional neural networks (CNN) were trained using the Keras–Tensorflow package with the faster_rcnn_inception_v2_pets model. Samples of six forest species were used to train CNN. From these, attempts were made with the number of thresholds, which is the cutoff value of the function; any value below this output is considered 0, and values above are treated as an output 1; that is, values above the value stipulated in the Threshold are considered as identified species. The results showed that the reduction in the threshold decreases the accuracy of identification, as well as the overlap of the polygons of species identification. However, in comparison with the data collected in the field, it was observed that there exists a high correlation between the trees identified by the CNN and those observed in the plots. The statistical metrics used to validate the classification results showed that CNN are able to identify species with accuracy above 90%. Based on our results, which demonstrate good accuracy and precision in the identification of species, we conclude that convolutional neural networks are an effective tool in classifying objects from UAV images.

Keywords: deep learning; drone; forest identification; unmanned aerial vehicles



Citation: Moura, M.M.; de Oliveira, L.E.S.; Sanquetta, C.R.; Bastos, A.; Mohan, M.; Corte, A.P.D. Towards Amazon Forest Restoration: Automatic Detection of Species from UAV Imagery. *Remote Sens.* **2021**, *13*, 2627. <https://doi.org/10.3390/rs13132627>

Academic Editor: Enrico Tomelleri

Received: 3 June 2021

Accepted: 1 July 2021

Published: 4 July 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

Optical sensors coupled with Unmanned Aerial Vehicles (UAVs; also referred to as drones) are commonly used to acquire geometric characteristics of forests, such as canopy height and diameter. In addition to quantitative information, these sensors can generate qualitative attributes about forest environments, enhancing the possibility of characterizing forest species. However, for such a characterization, solely depending on traditional image processing tools might not be an effective strategy at all times [1,2]. Therefore, additional statistical tests, programming tools, and implementation workflows are deemed necessary for processing these images. In this context, Convolutional Neural Networks (CNN) have been widely used in detecting diverse objects [3–7], especially while characterizing forest environments [8–14].

CNN are artificial intelligence algorithms based on multilayer (feed-forward) neural networks, typically with up to 20 or 30 layers, and are distinguished from other neural networks due to their superior performance with object classification, detection, and

segmentation [15,16]. They have three main types of layers, which are: convolutional layer, pooling layer, and fully-connected layer. Their main characteristic is that they represent the objects in the images and have a high generalization ability, thereby being suitable to classify new examples with reduced data input in the training set accurately [16–19]. CNN are considered as mathematical models of classification inspired by the biological system of visual data processing. In the literature [20–24], these algorithms have satisfactory results for many issues involving images, including plant classification, features extraction and determination of nutrients in plants. These classifications generally rely on morphological descriptors such as leaves, stems, fruits, and flowers [25–29].

Inventories with remote sensors and forest species' composition using an image platform instead of field surveys lead to a reduction in overall cost and acquisition time. Moreover, it favors the analysis of species' composition in large territorial extensions [30–33]. Accurate assessments of forest species' composition promote gains in research and management of these areas since this knowledge is used to analyze biodiversity patterns, estimate wood stocks, and improve stored carbon survey. Various authors [23,34–37] have highlighted the lack of (1) high resolution spatial and spectral images; (2) sensitive machine learning classifiers sophisticated enough to identify lighting, shape, size, and pattern of trees, tops, and leaves, as well as a spectral mix in the tops themselves; and (3) accurate spatial data to train classifiers. However, these difficulties are gradually being solved with new technologies, such as high spatial resolution images obtained through UAVs [38].

UAV images have adequate resolution for the analysis and characterization of forest areas (planted or native). They have been widely used because of their versatility and ease in acquiring data of large areas. In addition, these images are also used as a support for the monitoring and visual analysis of vegetation, especially in areas of native forests and other parts undergoing regeneration in the legal Amazon. In this context, remote sensing is continually improved, nonetheless, new classification tools are needed to bridge the gap between sensor characteristics, remotely sensed high-resolution images and "big data" type information about forests. Therefore, the objective of this study was to test the performance of the application of machine learning algorithms on high-resolution images obtained using UAV for identifying forest species in areas of forest regeneration in the Amazon.

2. Materials and Methods

2.1. Study Area

This study was carried out in a region of the Amazon Forest located in Rondônia, Brazil, between the meridians $62^{\circ}44'05''$ and $63^{\circ}16'54''$, and parallels $9^{\circ}00'00''$ and $9^{\circ}30'00''$, South latitude, as shown in Figure 1. We analyzed areas of forest restoration plantations with plots that are continuously monitored to assess tree increments, new species' natural regeneration, and carbon stock monitoring (Figure 1). In these plots, measurements are conducted on a regular basis to characterize and geolocate individual trees. This information was used as a subsidy to evaluate the results from automated processes related to UAV images.

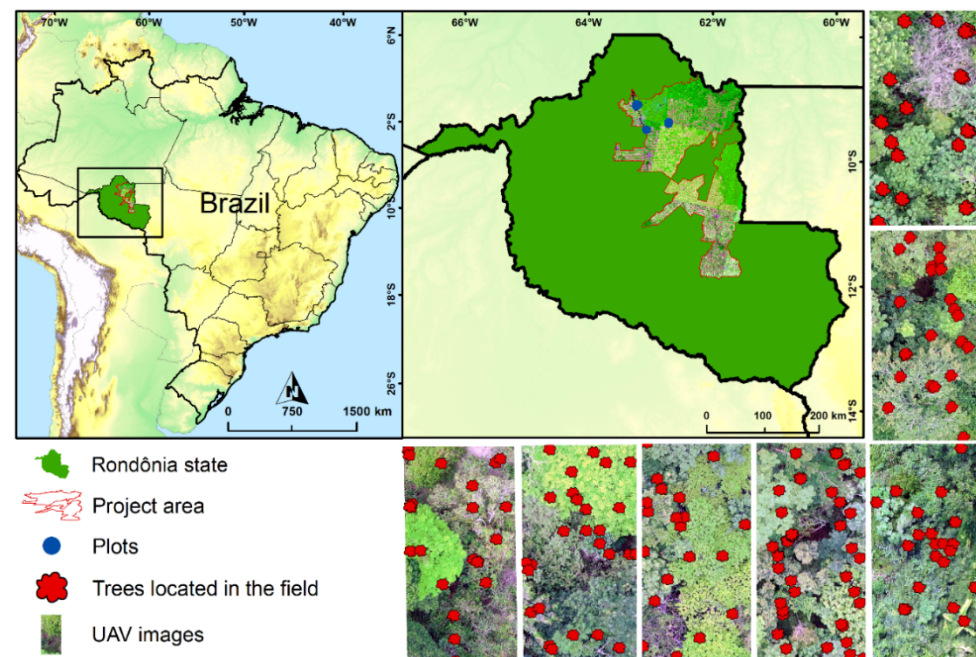


Figure 1. Study area location.

2.2. Image Acquisition and Pre-Processing

The images were acquired by a UAV DJI Phantom 4 Pro, using a RGB camera; the main characteristics are listed in Table 1. Images were obtained in November 2017, 2018, and 2019, always from 8 am to 3 pm, taking advantage of solar lighting and reduced shading in images.

Table 1. Characteristics of images obtained from the UAV DJI PHANTOM 4 PRO.

Characteristics	Type
Width	5.472 pixels
Height	3.648 pixels
Camera	FC6310
Exposure time	1/400s
Aperture value	4.7EV (f/5,6)
Speed rate (ISO)	100
Image type	.jpeg
Focal distance	8.8 mm
Pixel resolution	2.2 cm

We processed images in the Pix4D mapper software to make the orthomosaic of the areas. No tonal corrections or contrasts were made on the raw images or final mosaics.

The species were chosen according to the number of individuals available for sample acquisition, and samples were selected only from plots where forest inventories were conducted in the field, in order to select the species from the UAV images correctly.

The major tree species considered are as follows: *Cecropia juranyiana* C.Mart (Urticaceae) is a pioneer species widely found in forest areas that have been anthropized and in secondary forests. It has an average height of 15 m, and foliolate leaves with up to 14 leaflets [39]. *Hymenaea courbaril* L. (Fabaceae) is a semicaducifolia tree, up to 35 m in height and 80 cm in diameter, with compound leaves with 2 leaflets and a broad crown [40,41]. *Bauhinia acreana* Harms (Fabaceae) is a species with a broad crown with bilobed leaves and pod dehiscent fruit [42]. *Anacardium occidentale* L. (Anacardiaceae) is a tree up to 10 m tall and 40 cm in diameter, with a wide crown and simple petiole leaves, which can change their coloration according to the time of year, with pink

younger leaves and green for mature leaves [43,44]. *Handroanthus serratifolius* (Vahl) S. Grooe (Bignoneaceae) is a tree up to 30 m tall, 80 cm in diameter, with palmate leaves, with 5–7 lobed and a serrated margin. In the flowering months, the species remains only with flowers [44,45]. *Anadenanthera* sp. Speg (Fabaceae) reaches about 25 m in height, with compound leaves [44,46]

2.3. Convolutional Neural Networks Training

A total of 683 trees (samples) were inserted in the *LabelImg* software [47], and this made it possible to select the region of interest using polygons with rectangular and square shapes (subsamples). We obtained 2437 subsamples (Table 2). It was necessary to select subsamples from the same tree, and hence, in a single tree, the largest possible number of samples was selected, depending on canopy size. (Figure 2).

Table 2. Forest species and number of samples for training convolutional networks.

Species	Samples		Subsamples	
	Trees	Test (30%)	Training (70%)	
<i>Cecropia juranyiana</i>	157	158	369	
<i>Hymenaea courbaril</i>	102	126	294	
<i>Bauhinia acreana</i>	96	108	252	
<i>Anacardium occidentale</i>	90	80	186	
<i>Handroanthus serratifolius</i>	106	116	270	
<i>Anadenanthera</i> sp.	132	144	336	
Total	683	731	1706	



Figure 2. Representation of sample selection in an individual image. The red polygon represents tree canopy (*H. serratifolius*), and the yellow rectangles represent subsamples obtained for this species.

The samples here can change size, but the CNN makes the same adjustment for all the inserted images, even if they are of different sizes; the neural network makes the identification using all the parameters referring to the selected samples, texture, contrast, color intensity, brightness, patterns of brightness and darkness, and numbers and arrangements

of brightness and darkness are minimized in the internal layers of the neural network; that is, the samples are standardized internally aiming at mitigating these factors. Regardless of the sample size, the transformations that the CNN make internally are presented in Figure 3. From the characteristics presented (shape, texture, brightness, contrast, etc.), it is possible to distinguish between the species under study in the network.

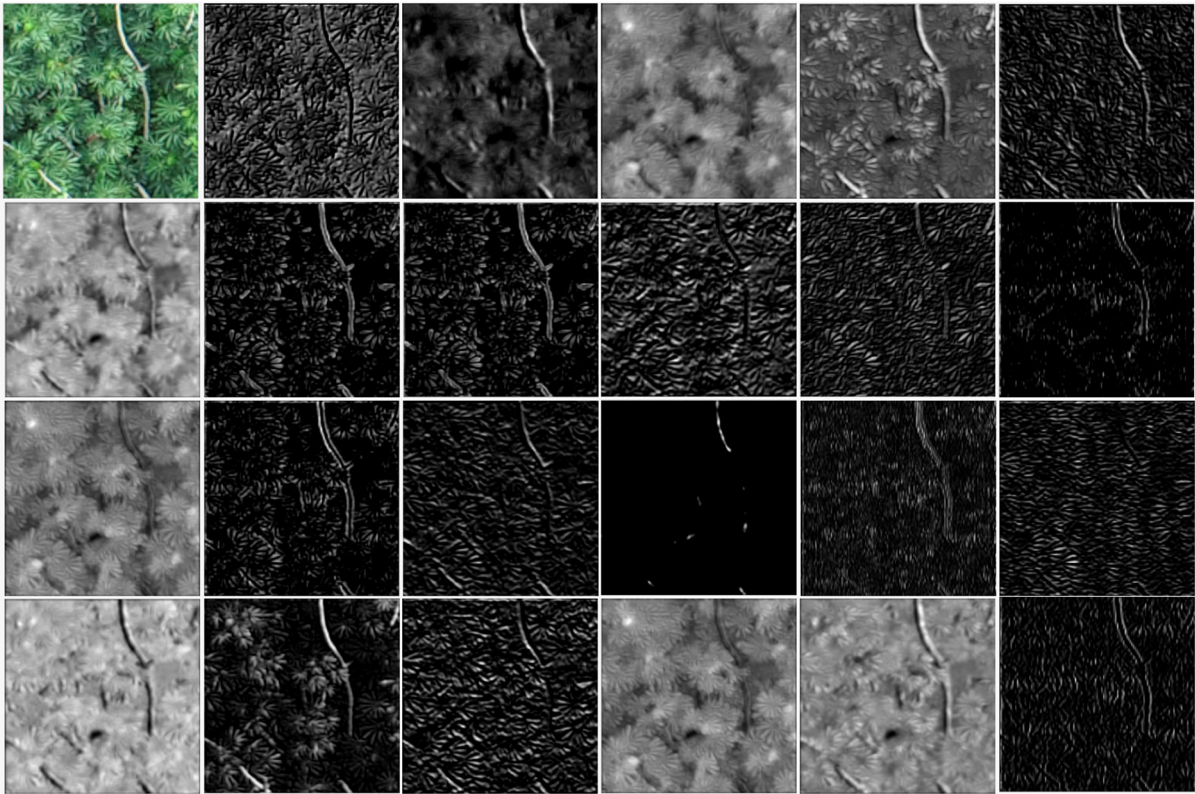


Figure 3. Representation of transformations that CNN make internally for *Cecropia juranyiana* C.Mart (Urticaceae).

At the end of the sample selection process, the software provides a file in .csv format, containing filename, height, width, class, xmin, xmax, ymin, and ymax [47]. From the *Labelimg* .csv file, it was possible to create two sets of samples randomly: Training (70%; 1706 samples) and Test (30%; 731 samples); these samples formed from the Training and Test images were fed into the CNN as input parameters. For validation, 487 images were separated from the training samples to be applied in the validation, guaranteeing that the model did not present overfitting (Table 2).

We used the Python computational language to implement the convolutional neural networks with the Keras–TensorFlow package. This open-source system from Google uses and implements deep learning algorithms of neural networks. The *faster_rcnn_inception_v2_pets* model was utilized from this package, which was modified to train target species samples.

The neural network learns the patterns from the input data by reading the input data set and applying different computations to it. However, the neural network does not just do this once; it learns repeatedly using the input data set and also the results of previous tests. Each step in learning from the input data set is called an epoch. That is, an epoch refers to one cycle in the entire training data set [48,49]. Initially, CNN were trained with a large number of epochs or steps (iterations) to ensure that the smallest loss would be within that step range. After the first training, we determined an ideal number of steps to obtain the least loss to optimize the analyses and repetitions that would be performed; this test served fundamentally to know how many epochs would be necessary for the final model fit; this step was essential because an excessive number of epochs leads the model

to overfitting, i.e., in this case, it will present results with very good statistical metrics, but with erroneous identification of species.

Therefore, tests were performed by reducing or increasing thresholds that are functions that limit the results of the CNN output. Values equal to 1 suggest maximum identification in the model outputs, which can lead to the identification of species that are not recognized in the field, and values of 0 suggest no identification of any species, so manual adjustment of the Threshold is essential for fine-tuning the model output; this will help prevent outputs from presenting results that will be considered false-positive. After initial training, we observed that modifying the threshold resulted in non-recognition of some species or false recognition. This led to a training model with variation in the number of species and thresholds. All results were compared with field data to obtain a model that best reflected real conditions of the studied area. Herein, thresholds modify the output according to input limits. The network receives inputs, then applies a linear combination, and, if that combination is greater or less than the specified limit value, it will produce an output of 1 or 0, respectively. Therefore, results with low probabilities would be rejected. Equation 1 represents the threshold function, where Sigma (Σ) is the sum of input (x) and weight (w) pairs.

$$\text{output} = \begin{cases} 0 & \text{se } \sum_i w_i x_i < \text{Threshold} \\ 1 & \text{se } \sum_i w_i x_i \geq \text{Threshold} \end{cases} \quad (1)$$

2.4. Classification Validation

Validation of the classifications was performed on the dataset for the test. For this, statistical metrics were used to evaluate and test the performance of the adjusted CNN. The metrics are specified in Table 3, of which are calculated according to the results of the classifications. The Kappa index is a measure of agreement used in nominal scales that gives us an idea of how far the observations deviate from those expected, at random, thus indicating to us how legitimate the interpretations are. It measures the percentage of the data values on the main diagonal of the table and then adjusts these values for the amount of agreement that might be expected [50]. Accuracy is an index that reflects the rate at which individuals are correctly classified into the category containing their true score. Ranking accuracy is usually attributed to the appropriateness and validity of your decisions based on the obtained score. A large value for the index indicates a high hit rate of individuals in the correct categories, and a low value indicates a lower rate of correct classification of individuals [51].

Table 3. Statistical indexes used to verify classification results.

Statistical Index	Equation	Description	Reference
Kappa	$\text{Kappa} = \frac{\text{ACC}_{\text{overall}} - \text{RACC}_{\text{overall}}}{1 - \text{RACC}_{\text{overall}}}$	Ranges from 1 to -1 , where 1 indicates a perfect rating and -1 represents an unreliable rating	[50]
Accuracy	$\text{ACC} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$	Number of correct classifications of all predictions	[51]
Adjusted F-score	$\text{AGF} = \sqrt{\left[5 \times \frac{\text{PPV} \times \text{TPR}}{(4 \times \text{PPV}) + \text{TPR}} \right] \times \left[(1 + 0.5^2) \times \frac{\text{NPV} \times \text{TNR}}{(0.5^2 \times \text{NPV}) + \text{TNR}} \right]}$	Use all confusion matrix elements and provide more weights to samples that are correctly classified in the lowest class.	[52]

Table 3. Cont.

Statistical Index	Equation	Description	Reference
Mathews Correlation Coefficient	$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP) \times (TP + FN) \times (TN + FP) \times (TN + FN)}}$	Correlation coefficient of observed and predicted binary classifications, where: Negligible—<0.3 Poor—0.3–0.5 Moderate—0.5–0.7 Strong—0.7–0.9 Very strong—>0.9	[53]
Similarity	$sInd = 1 - \sqrt{\frac{(1-TNR)^2 + (1-TPR)^2}{2}}$	It ranges from 0 to 1, where: 0—no correct ratings 1—perfect classification	[54]

Where, ACCoverall: relative compliance observed between classifiers; RACCoverall: hypothetical probability compliance using the observed data to calculate the probabilities of each classifier to identify each category randomly; TP: true positive; TN: true negative; FP: false positive; FN: false negative; PPV: positive predictive value; TPR: true positive rate; NPV: negative predictive value; TNR: true negative rate.

The adjusted F-score is an index that groups all elements of the original confusion matrix and gives more weight to correctly classified patterns in correctly classified classes [52]. The Matthews Correlation Coefficient has a range from -1 to 1 , and when it has a value of -1 , it indicates a completely wrong binary classifier, while 1 indicates a completely correct binary classifier. This index allows you to evaluate the performance of your classification model [53], whereas similarity is a metric ranging from 0 (wrong classification) to 1 (perfect classification) calculated from the averages of the classifications of the classes of interest [54].

3. Results

From initial tests using the CNN, the training with six species and a varying threshold, we observed that an increase or reduction in threshold parameter interfered in species' characterization accuracy. In this sense, with this parameter equal to 0.9 , the accuracy of species' recognition rates was higher; however, in this case, only two species were identified. When the same value was reduced to 0.8 , the network identified three species; the same happened for 0.7 and 0.6 as well (Figure 4). Decreasing threshold reduced identification accuracy, as well as the overlap of species' identification polygons. However, when comparing the results obtained by the neural network and field data, we observed a relationship between the trees identified by CNN and those observed in the field. Moreover, we observed more classification uncertainty in the areas with more shadows or overlapping of the canopies compared to species with higher canopies and without shade.

Figure 5 shows the loss in insertion of the species, where 1 Species: *C. juranyiana*; 2 Species: *C. juranyiana* + *H. courbaril*; 3 Species: *C. juranyiana* + *H. courbaril* + *B. acreana*; 4 Species: *C. juranyiana* + *H. courbaril* + *B. acreana* + *A. occidentale*; 5 Species: *C. juranyiana* + *H. courbaril* + *B. acreana* + *A. occidentale* + *H. serratifolius*; 6 Species: *C. juranyiana* + *H. courbaril* + *B. acreana* + *A. occidentale* + *H. serratifolius* + *Anadenanthera* sp. At the beginning of the training, the loss was high and decreased at each epoch until it reached the lowest value. From this point, the loss began to increase, showing adjustments with little accuracy in species recognition (Figure 5). From this first training, we observed that recognition of only one species, two species, three species, four species, five species, and six species required 494, 541, 892, 976, 993, and 1042 epochs, respectively. Additionally, on changing the threshold, and including more species (characteristics), the network training time was relatively increased, making this process more time-consuming.

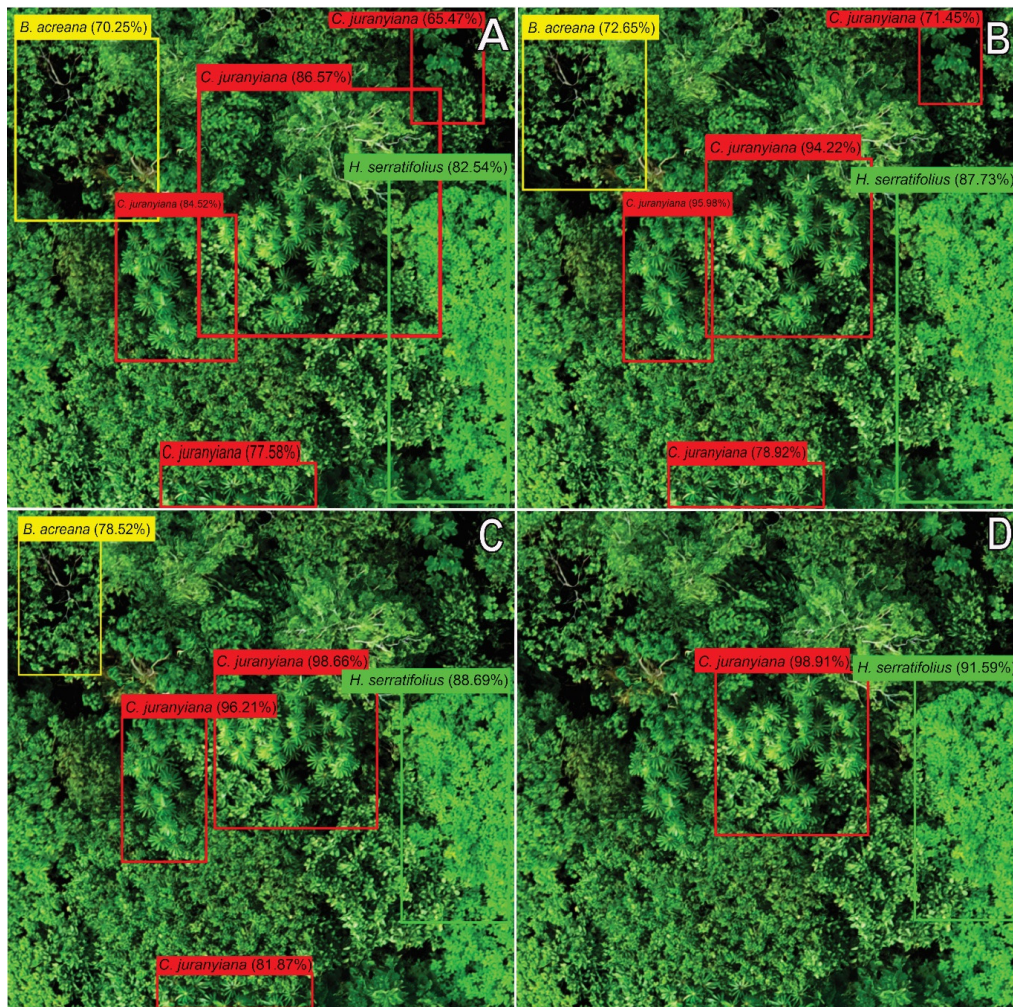


Figure 4. Model application including six species in the training. (A)—Threshold = 0.9, (B)—Threshold = 0.8, (C)—Threshold = 0.7, (D)—Threshold = 0.6.

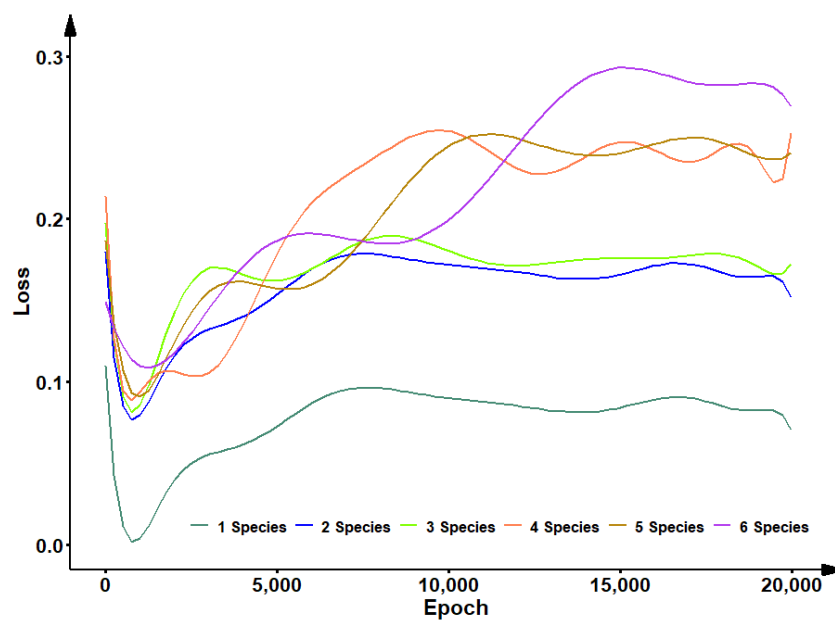


Figure 5. Variation of loss concerning epoch and number of species used in training for convolutional neural networks.

Figure 6 shows loss variation in the validation of the adjusted model using all species. This model eventually converged to its lowest loss value with 439 epochs, demonstrating that with this amount, it is possible to identify the six species used for training in the UAV images.

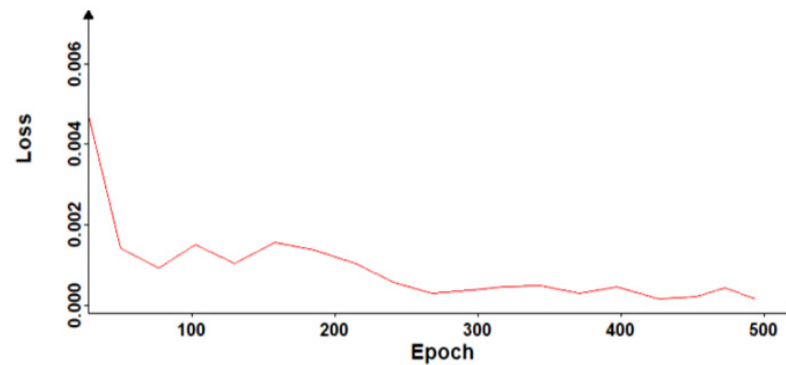


Figure 6. Loss variation of the adjusted convolutional neural networks with 6 Species.

Figure 7 shows the tree location in the plots and the model's functionality in locating trees. We observed that trees with larger canopies are more easily detected when compared with the smaller ones. False recognition is associated with overlapping crowns or similarities between leaf characteristics. In genus identification by the model, the results were more accurate for trees with a smaller canopy circumference. The model did not identify or erroneously classify species whose canopies were of less dense foliage. Canopies with leaves with higher contrast (brightness, color) than the others were identified with greater accuracy. However, we observed some trees with discrepancies concerning the foliage or canopy of the same genre, which had a higher error rate in the classification.

We observed that the training with a threshold of 0.7 provided satisfactory results concerning processing time and species recognition. Therefore, the same value was utilized to calculate the confusion matrix—with all the six species. The overall accuracy achieved was 91.80%. Nevertheless, there occurred false recognition of the characteristics in all species; the biggest classification error (false positive) was found among *C. juranyiana* and *H. serratifolius* with *B. acreana*, in which there was a false characterization in 26 occasions (Table 4).

Table 4. Confusion matrix with the insertion of six species in a convolutional neural network with threshold = 0.7.

	<i>Cecropia</i>	<i>Hymenaea</i>	<i>Bauhinia</i>	<i>Anacardium</i>	<i>Handroanthus</i>	<i>Anadenanthera</i>	Total	User
<i>Cecropia</i>	356	0	26	4	0	0	386	92.23%
<i>Hymenaea</i>	0	281	0	0	5	12	298	94.30%
<i>Bauhinia</i>	4	12	201	4	5	5	235	87.01%
<i>Anacardium</i>	4	0	8	171	0	1	184	92.93%
<i>Handroanthus</i>	0	1	12	7	240	9	269	89.22%
<i>Anadenanthera</i>	5	0	5	0	11	318	335	93.81%
Total	369	294	252	186	261	345	1707	-
Producer accuracy	96.48%	95.58%	79.76%	91.94%	91.95%	92.17%	-	1567
Global accuracy	91.80%					k	g	

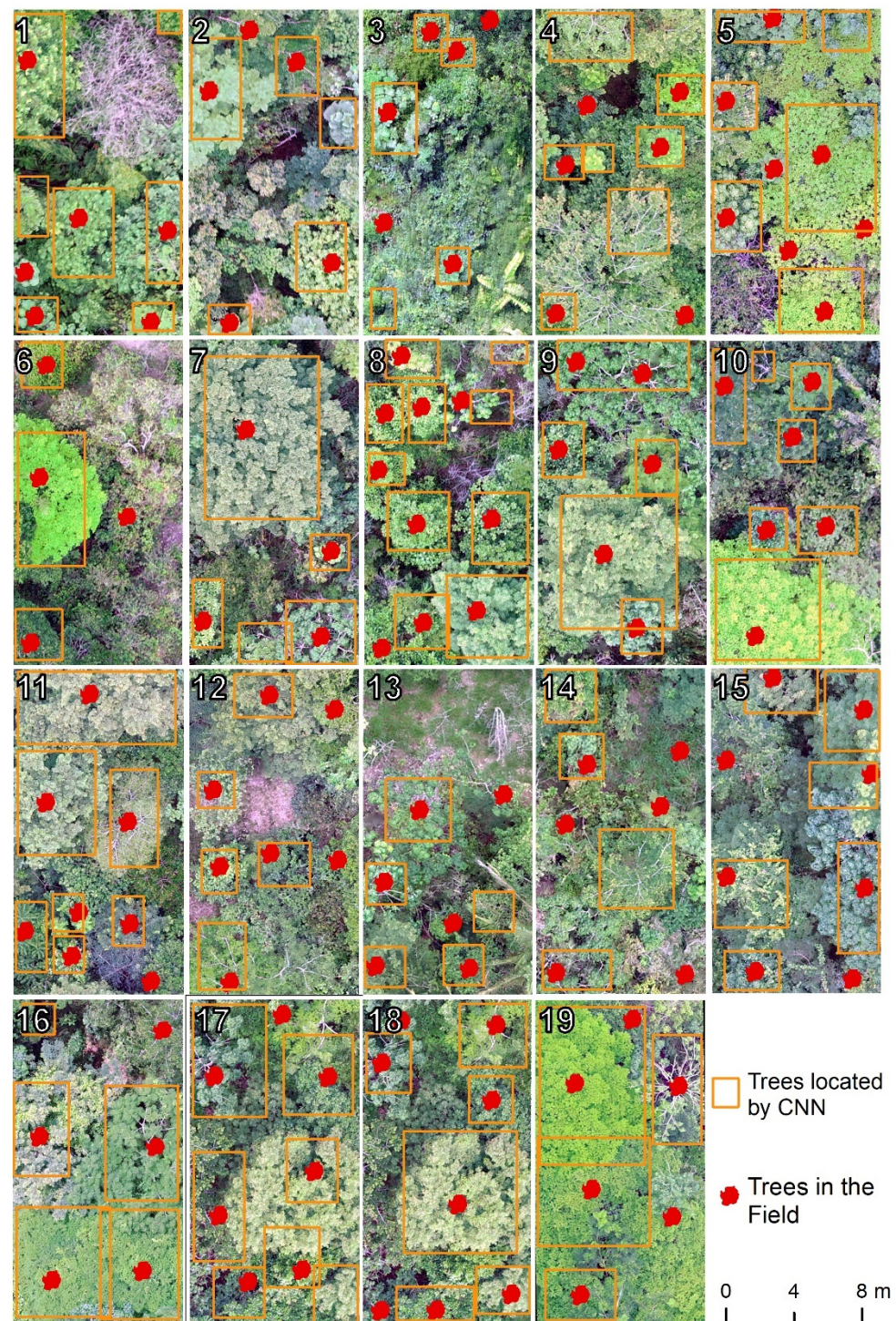


Figure 7. Location of trees from field data and species recognition model.

Within the confusion matrix, the lowest values were observed in the species *B. acreana* with 79.76% and 87.01% for producer and user accuracy, respectively. The highest values observed for producer accuracy were for *C. juranyiana* (96.48%) and *H. courbaril* (95.58%), and in terms of user accuracy, *H. courbaril* (94.30%) and *Anadenanthera* sp. (93.81%) had the best values.

Table 5 shows all the statistical indexes used to validate classifications. The Kappa index obtained was 0.9006, which indicates an adequate classification. We verified results close to this one in Accuracy, adjusted F-score, and Similarity. However, with respect to Matthews' correlation, the species *B. acreana* and *H. serratifolius* obtained a strong classifica-

tion, while for the other species, this index was very strong, just as in Sensitivity the same species obtained results below 0.9.

Table 5. Statistics obtained in the training data set classification process.

Index	<i>Cecropia</i>	<i>Hymenaea</i>	<i>Bauhinia</i>	<i>Anacardium</i>	<i>Handroanthus</i>	<i>Anadenanthera</i>
Kappa				0.9006		
Accuracy	0.9748	0.9824	0.9525	0.9836	0.9707	0.9718
Adjusted F-score	0.9706	0.9667	0.9135	0.9587	0.9594	0.9589
Matthews correlation	0.9272 ¹	0.9387 ¹	0.8057 ²	0.9151 ¹	0.8884 ²	0.9123 ¹
Sensitivity	0.9222	0.9429	0.8701	0.9293	0.8921	0.938

¹—Very strong e ²—Strong.

4. Discussion

The recognition of forest species in the field or using remote sensing techniques is a major challenge for forestry researchers. In our study, particularities related to species' canopy and leaf structure (dendrometric characteristics) were found decisive for achieving satisfiable accuracy of predictive models. The results found in this study demonstrate good accuracy in identifying species using UAV imagery in this regard. All species studied obtained promising results in their identification.

The species *Cecropia* and *Bauhinia* have the smallest canopy perimeters of the species under study. However, statistical indexes showed results above 0.9. This is because their canopy has unevenness concerning branches and leaves, and, as reported in [55], the canopy characteristics directly influence the classifier. *H. courbaril* and *Anadenanthera* sp. have smaller leaves and larger canopies, standing out among the others. *A. occidentale* and *H. serratifolius* have more leaves and canopies that stand out in the forest, and, for this reason, the classifier was efficiently able to distinguish them from the others.

Compared to experiments conducted in environments with fewer tree species, such as in [56], the classification metrics were higher (97.80%—Global accuracy) in this study (91.80%—Global accuracy), as there was less chance of confusion in species' characteristics. Only limited studies have reported their experiences in object detection concerning the selection of the threshold. Freudenberg et al. (2019) [57] used this parameter with a value of 0.5 to detect palm trees, resulting in an F-score ranging from 0.875 to 0.957; these values are close to those observed in our study. In this context, Xiong et al. (2020) [58] compared sensor height with threshold variation in predicting dendrometry variables. The authors concluded that with an increase in sensor distance to the target object, it was necessary to retrain the CNN with an increasing threshold, thus obtaining better results in accuracy and error. In Wang et al. (year) [59], the authors reported that there is no predetermined value of this threshold and argued that the most appropriate value is dependent directly on the database and computational method used.

When analyzing the classifications with different thresholds, our results showed an overlap of CNN detection polygons similar to those obtained by Sarabia et al. (year) [60] and Xiong et al. (2020) [61]. In these studies, the authors reported issues with the interpretation of edges of trees to neighbors; a reduction in threshold value was found crucial to notice this uncertainty. However, this factor did not influence correct species identification; it only changed the visualization of the identification polygon and reduced the value related to the probability of species characterization. Regarding the number of epochs, previous studies suggest that this value is proportional to the number of characteristics inserted in the model [3,25,31,32].

The statistical metrics used to determine model accuracy showed very strong correlations for species detections. However, the results of *B. acreana* demonstrate that the existence of few samples of a given species results in lower values of Accuracy, Adjusted F-score, Matthews correlation, Sensitivity, and Similarity (Table 5). On the other hand, *H. courbaril* had fewer samples than *C. juranyiana* and obtained better results, mainly related to reduced

size of the canopies and leaves of *C. juranyiana* compared to other species. Because it had long branches and leaves only at the tips, the canopies of neighboring trees intersected in these branches, causing detection inaccuracy and false species characterization. This characteristic was also reported by Wagner et al. (2019) [62], who reported accuracy results similar to our study when comparing a consortium of species in the Atlantic Forest biome, which have similar physiognomic characteristics to the Amazon biome.

Matthews' Similarity and Correlation are classification metrics considered to be less biased as they use multiple input variables; they incorporate both the set imbalance and the amount of data referring to classes [34]. In this regard, the inclusion of species caused lower values of those metrics in comparison to Accuracy and the F-score since they have bias related to the number of characteristics, and, specifically, the F-score does not depend on the true negatives (TN). The high similarity observed here demonstrates that the classifier can identify true negatives, just as the high sensitivity indicates that the classifier was able to identify true positives [35].

UAVs generally operate at low altitudes and can acquire images in better spatial resolutions than satellite images, providing more detailed information not only about the forest but also at the species level. However, the methodology presented here requires analysis, processing, and model adjustment to classify the species. Regarding costs, operations with UAVs are more feasible compared to traditional airborne imaging or high-resolution acquisition of orbital sensors. Thus, it is possible to acquire images more frequently, favoring periodic monitoring of the forest in the process of restoration and the study of its dynamics [63].

5. Conclusions

Convolutional neural network training using the `faster_rcnn_inception_v2_pets` model provided satisfactory results in identifying species characteristics. It proved to be an effective tool in classifying objects in UAV images without the need to modify image properties. In this sense, CNN training demonstrated good accuracy and precision in identifying the species under study; however, species insertion made the training time-consuming. Moreover, we noticed that there exists a higher classification error when species' canopies overlap each other, along with shading in images. The quality of UAV images was observed to be an essential component required for obtaining species' characteristic patterns and for network training. It also impacts the frequency for obtaining the images, which expands the number of samples to be acquired. In our case, the number of samples used in CNN training was found sufficient to recognize forest species' characteristics and intrinsic patterns.

The identification of individual trees provides subsidies for new information on forest restoration ecosystems, and we hope our study encourages future researchers to develop methodologies to obtain information regarding distinctions in forest characterization concerning the seasons, vegetation phenology during the months of the year, atmospheric conditions, and other sensors, adding more information to the models.

Author Contributions: All the authors have significantly contributed to the development of this study, being involved in designing, drafting, and giving critical reviews. Conceptualization, M.M.M., L.E.S.d.O., A.B., C.R.S., M.M., A.P.D.C.; methodology, M.M.M., L.E.S.d.O.; validation, A.B., C.R.S., M.M., A.P.D.C.; formal analysis, C.R.S., M.M., A.P.D.C.; data curation, C.R.S., M.M., A.P.D.C.; writing—original draft preparation, M.M.M., M.M., A.P.D.C.; writing—review and editing, C.R.S., M.M., A.P.D.C.; visualization, M.M.M.; supervision, A.B., C.R.S., A.P.D.C.; project administration, M.M.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brasil (CAPES)—Finance Code 001. The authors thank the technical and scientific support of the Center of Excellence in Research on Carbon Fixation in Biomass (BIOFIX) and Center of studies RioTerra. We would also like to thank three anonymous reviewers, MDPI editorial board and staff for their support and guidance.

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

References

1. Natesan, S.; Armenakis, C.; Vepakomma, U. Individual tree species identification using Dense Convolutional Network (DenseNet) on multitemporal RGB images from UAV. *J. Unmanned Veh. Syst.* **2020**, *8*, 310–333. [[CrossRef](#)]
2. Casapia, X.T.; Falen, L.; Bartholomeus, H.; Cárdenas, R.; Flores, G.; Herold, M.; Coronado, E.N.H.; Baker, T.R. Identifying and Quantifying the Abundance of Economically Important Palms in Tropical Moist Forest Using UAV Imagery. *Remote Sens.* **2019**, *12*, 9. [[CrossRef](#)]
3. Cao, D.; Chen, Z.; Gao, L. An improved object detection algorithm based on multi-scaled and deformable convolutional neural networks. *Hum. Cent. Comput. Inf. Sci.* **2020**, *10*, 1–22. [[CrossRef](#)]
4. Wu, Y.; Wang, J.; Qian, K.; Liu, Y.; Guo, R.; Hu, S.; Yuancong, W.; Chen, T.; Rong, L. An energy-efficient deep convolutional neural networks coprocessor for multi-object detection. *Microelectron. J.* **2020**, *98*, 104737. [[CrossRef](#)]
5. Zhang, M.; Chen, Y.; Liu, X.; Lv, B.; Wang, J. Adaptive Anchor Networks for Multi-Scale Object Detection in Remote Sensing Images. *IEEE Access* **2020**, *8*, 57552–57565. [[CrossRef](#)]
6. Yao, Q.; Hu, X.; Lei, H. Multiscale Convolutional Neural Networks for Geospatial Object Detection in VHR Satellite Images. *IEEE Geosci. Remote Sens. Lett.* **2021**, *18*, 23–27. [[CrossRef](#)]
7. Pi, Y.; Nath, N.D.; Behzadan, A.H. Convolutional neural networks for object detection in aerial imagery for disaster response and recovery. *Adv. Eng. Inform.* **2020**, *43*, 101009. [[CrossRef](#)]
8. Qiao, R.; Ghodsi, A.; Wu, H.; Chang, Y.; Wang, C. Simple weakly supervised deep learning pipeline for detecting individual red-attacked trees in VHR remote sensing images. *Remote Sens. Lett.* **2020**, *11*, 650–658. [[CrossRef](#)]
9. Xiao, C.; Qin, R.; Huang, X. Treetop detection using convolutional neural networks trained through automatically generated pseudo labels. *Int. J. Remote Sens.* **2019**, *41*, 3010–3030. [[CrossRef](#)]
10. Roslan, Z.; Awang, Z.; Husen, M.N.; Ismail, R.; Hamzah, R. Deep Learning for Tree Crown Detection in Tropical Forest. In Proceedings of the 14th International Conference on Ubiquitous Information Management and Communication (IMCOM), Taichung, Taiwan, 3–5 January 2020; pp. 1–7.
11. Fromm, M.; Schubert, M.; Castilla, G.; Linke, J.; McDermid, G. Automated Detection of Conifer Seedlings in Drone Imagery Using Convolutional Neural Networks. *Remote Sens.* **2019**, *11*, 2585. [[CrossRef](#)]
12. Syrris, V.; Hasenohr, P.; Delipetrev, B.; Kotsev, A.; Kempeneers, P.; Soille, P. Evaluation of the Potential of Convolutional Neural Networks and Random Forests for Multi-Class Segmentation of Sentinel-2 Imagery. *Remote Sens.* **2019**, *11*, 907. [[CrossRef](#)]
13. Sanquetta, C.R.; Behling, A.; Corte, A.P.D.; Cadore, G.C.; Junior, S.C.; Macedo, J.H.P. Eficiência de conversão da radiação fotossintética interceptada em Fitomassa de mudas de *Eucalyptus dunii* Maiden em função da densidade de plantas e do ambiente de cultivo. *Sci. For. Sci.* **2014**, *42*, 573–580.
14. Sanquetta, C.R.; Behling, A.; Dalla Corte, A.P.; Péllico Netto, S.; Rodrigues, A.L.; Simon, A.A. A model based on environmental factors for diameter distribution in black wattle in Brazil. *PLoS ONE* **2014**, *9*, 100093. [[CrossRef](#)] [[PubMed](#)]
15. Albawi, S.; Mohammed, T.A.; Al-Zawi, S. Understanding of a convolutional neural network. In Proceedings of the 2017 International Conference on Engineering and Technology (ICET), Antalya, Turkey, 21–23 August 2017; pp. 1–6.
16. Shakya, A.; Biswas, M.; Pal, M. Parametric study of convolutional neural network based remote sensing image classification. *Int. J. Remote Sens.* **2021**, *42*, 2663–2685. [[CrossRef](#)]
17. Chen, Z.; Zhang, J.; Tao, D. Recursive Context Routing for Object Detection. *Int. J. Comput. Vis.* **2021**, *129*, 142–160. [[CrossRef](#)]
18. Jia, F.; Liu, J.; Tai, X.-C. A regularized convolutional neural network for semantic image segmentation. *Anal. Appl.* **2021**, *19*, 147–165. [[CrossRef](#)]
19. Shivappriya, S.; Priyadarsini, M.; Stateczny, A.; Puttamadappa, C.; Parameshchhari, B. Cascade Object Detection and Remote Sensing Object Detection Method Based on Trainable Activation Function. *Remote Sens.* **2021**, *13*, 200. [[CrossRef](#)]
20. Safonova, A.; Guirado, E.; Maglinets, Y.; Alcaraz-Segura, D.; Tabik, S. Olive Tree Biovolume from UAV Multi-Resolution Image Segmentation with Mask R-CNN. *Sensors* **2021**, *21*, 1617. [[CrossRef](#)]
21. Kattenborn, T.; Leitloff, J.; Schiefer, F.; Hinz, S. Review on Convolutional Neural Networks (CNN) in vegetation remote sensing. *ISPRS J. Photogramm. Remote Sens.* **2021**, *173*, 24–49. [[CrossRef](#)]
22. Flores, D.; González-Hernández, I.; Lozano, R.; Vazquez-Nicolas, J.M.; Toral, J.L.H. Automated Agave Detection and Counting Using a Convolutional Neural Network and Unmanned Aerial Systems. *Drones* **2021**, *5*, 4. [[CrossRef](#)]
23. Yan, S.; Jing, L.; 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](#)]
24. Korznikov, K.A.; Kislov, D.E.; Altman, J.; Doležal, 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](#)]

25. Jassmann, T.J.; Tashakkori, R.; Parry, R.M. Leaf classification utilizing a convolutional neural network. In Proceedings of the SoutheastCon 2015, Fort Lauderdale, FL, USA, 9–12 April 2015.
26. Gu, J.; Wang, Z.; Kuen, J.; Ma, L.; Shahroudy, A.; Shuai, B.; Liu, T.; Wang, X.; Wang, G.; Cai, J.; et al. Recent advances in convolutional neural networks. *Pattern Recognit.* **2018**, *77*, 354–377. [[CrossRef](#)]
27. Woźniak, M. Advanced Computational Intelligence for Object Detection, Feature Extraction and Recognition in Smart Sensor Environments. *Sensors* **2020**, *21*, 45. [[CrossRef](#)] [[PubMed](#)]
28. Pullanagari, R.; Dehghan-Shoar, M.; Yule, I.J.; Bhatia, N. Field spectroscopy of canopy nitrogen concentration in temperate grasslands using a convolutional neural network. *Remote Sens. Environ.* **2021**, *257*, 112353. [[CrossRef](#)]
29. Jaihuni, M.; Khan, F.; Lee, D.; Basak, J.K.; Bhujel, A.; Moon, B.E.; Park, J.; Kim, H.T. Determining Spatiotemporal Distribution of Macronutrients in a Cornfield Using Remote Sensing and a Deep Learning Model. *IEEE Access* **2021**, *9*, 30256–30266. [[CrossRef](#)]
30. Wulder, M. Optical remote-sensing techniques for the assessment of forest inventory and biophysical parameters. *Prog. Phys. Geogr. Earth Environ.* **1998**, *22*, 449–476. [[CrossRef](#)]
31. Hyyppä, J.; Hyyppä, H.; Inkinen, M.; Engdahl, M.; Linko, S.; Zhu, Y.-H. Accuracy comparison of various remote sensing data sources in the retrieval of forest stand attributes. *For. Ecol. Manag.* **2000**, *128*, 109–120. [[CrossRef](#)]
32. Silva, C.A.; Klauber, C.; Hentz, A.M.K.; Corte, A.P.D.; Ribeiro, U.; Liesenberg, V. Comparing the performance of ground filtering algorithms for terrain modeling in a forest environment using airborne LiDAR data. *Floresta Ambiente* **2018**, *25*, e20160150. [[CrossRef](#)]
33. Da Cunha Neto, E.M.; Rex, F.E.; Veras, H.F.P.; Moura, M.M.; Sanquetta, C.R.; Käfer, P.S.; Sanquetta, M.N.I.; Zambrano, A.M.A.; Broadbent, E.N.; Corte, A.P.D. Using high-density UAV-Lidar for deriving tree height of *Araucaria Angustifolia* in an Urban Atlantic Rain Forest. *Urban For. Urban Green* **2021**, *63*, 127197. [[CrossRef](#)]
34. Miraki, M.; Sohrabi, H.; Fatehi, P.; Kneubuehler, M. Individual tree crown delineation from high-resolution UAV images in broadleaf forest. *Ecol. Inform.* **2021**, *61*, 101207. [[CrossRef](#)]
35. Zheng, J.; Fu, H.; Li, W.; Wu, W.; Yu, L.; Yuan, S.; Tao, W.Y.W.; Pang, T.K.; Kanniah, K.D. Growing status observation for oil palm trees using Unmanned Aerial Vehicle (UAV) images. *ISPRS J. Photogramm. Remote Sens.* **2021**, *173*, 95–121. [[CrossRef](#)]
36. Kopačková-Strnadová, V.; Koucká, L.; Jelének, J.; Lhotáková, Z.; Oulehle, F. Canopy Top, Height and Photosynthetic Pigment Estimation Using Parrot Sequoia Multispectral Imagery and the Unmanned Aerial Vehicle (UAV). *Remote Sens.* **2021**, *13*, 705. [[CrossRef](#)]
37. Neuville, R.; Bates, J.; Jonard, F. Estimating Forest Structure from UAV-Mounted LiDAR Point Cloud Using Machine Learning. *Remote Sens.* **2021**, *13*, 352. [[CrossRef](#)]
38. Fassnacht, F.; Latifi, H.; Stereńczak, K.; Modzelewska, A.; Lefsky, M.; Waser, L.; Straub, C.; Ghosh, A. Review of studies on tree species classification from remotely sensed data. *Remote Sens. Environ.* **2016**, *186*, 64–87. [[CrossRef](#)]
39. Berg, C.C. Espécies de *Cecropia* da Amazônia Brasileira. *Acta Amaz.* **1978**, *8*, 149–182. [[CrossRef](#)]
40. Pinto, R.B.; Tozzi, A.M.G.A.; Mansano, V. *Hymenaea* in Flora do Brasil. Available online: <http://floradobrasil.jbrj.gov.br/reflora/floradobrasil/FB22976> (accessed on 18 June 2021).
41. CNCFlora. *Hymenaea courbaril* in Lista Vermelha da Flora Brasileira Versão. Available online: <http://cncflora.jbrj.gov.br/portal/pt-br/profile/Hymenaea%20courbaril> (accessed on 18 June 2021).
42. Vaz, A.M.S.F. *Bauhinia* in Flora do Brasil. 2020. Available online: <http://floradobrasil.jbrj.gov.br/reflora/PrincipalUC/PrincipalUC.do> (accessed on 18 June 2021).
43. Silva-Luz, C.L.; Pirani, J.R.; Pell, S.K.; Mitchell, J.D. *Anacardiaceae* in Flora do Brasil. Available online: <http://floradobrasil.jbrj.gov.br/reflora/PrincipalUC/PrincipalUC.do> (accessed on 18 June 2021).
44. Lorenzi, H. *Árvores Brasileiras: Manual de Identificação e Cultivo de Plantas Arbóreas Nativas do Brasil*; Instituto Plantarum: Nova Odessa, Brazil, 2000.
45. Lohmann, L.G. *Handroanthus* in Flora do Brasil. Available online: <http://floradobrasil.jbrj.gov.br/reflora/PrincipalUC/PrincipalUC.do> (accessed on 18 June 2021).
46. Morim, M.P. *Anadenanthera* in Flora do Brasil. Available online: <http://floradobrasil.jbrj.gov.br/reflora/PrincipalUC/PrincipalUC.do> (accessed on 18 June 2021).
47. Tzotalin Git Code. Available online: <https://github.com/tzotalin/labelImg> (accessed on 30 March 2021).
48. Frankle, J.; Schwab, D.J.; Morcos, A.S. The Early Phase of Neural Network Training. *arXiv* **2020**, arXiv:2002.10365.
49. Taqi, A.M.; Awad, A.; Al-Azzo, F.; Milanova, M. The Impact of Multi-Optimizers and Data Augmentation on TensorFlow Convolutional Neural Network Performance. In Proceedings of the 2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), Miami, FL, USA, 10–12 April 2018; pp. 140–145.
50. Cohen, J. A Coefficient of Agreement for Nominal Scales. *Educ. Psychol. Meas.* **1960**, *20*, 37–46. [[CrossRef](#)]
51. Sammut, C.; Webb, G.I. *Encyclopedia of Machine Learning*; Springer: New York, NY, USA, 2010.
52. Maratea, A.; Petrosino, A.; Manzo, M. Adjusted F-measure and kernel scaling for imbalanced data learning. *Inf. Sci.* **2014**, *257*, 331–341. [[CrossRef](#)]
53. Matthews, B. Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochim. Biophys. Acta (BBA) Protein Struct.* **1975**, *405*, 442–451. [[CrossRef](#)]
54. Ballabio, D.; Grisoni, F.; Todeschini, R. Multivariate comparison of classification performance measures. *Chemom. Intell. Lab. Syst.* **2018**, *174*, 33–44. [[CrossRef](#)]

55. Ferreira, M.P.; Wagner, F.H.; Aragão, L.E.; Shimabukuro, Y.E.; Filho, C.R.D.S. Tree species classification in tropical forests using visible to shortwave infrared WorldView-3 images and texture analysis. *ISPRS J. Photogramm. Remote Sens.* **2019**, *149*, 119–131. [[CrossRef](#)]
56. Mubin, N.A.; Nadarajoo, E.; Shafri, H.Z.M.; Hamedianfar, A. Young and mature oil palm tree detection and counting using convolutional neural network deep learning method. *Int. J. Remote Sens.* **2019**, *40*, 7500–7515. [[CrossRef](#)]
57. Freudenberg, M.; Nölke, N.; Agostini, A.; Urban, K.; Wörgötter, F.; Kleinn, C. Large Scale Palm Tree Detection in High Resolution Satellite Images Using U-Net. *Remote Sens.* **2019**, *11*, 312. [[CrossRef](#)]
58. Gollob, C.; Ritter, T.; Wassermann, C.; Nothdurft, A. Influence of Scanner Position and Plot Size on the Accuracy of Tree Detection and Diameter Estimation Using Terrestrial Laser Scanning on Forest Inventory Plots. *Remote Sens.* **2019**, *11*, 1602. [[CrossRef](#)]
59. Wang, D.; Feng, Y.; Attwood, K.; Tian, L. Optimal threshold selection methods under tree or umbrella ordering. *J. Biopharm. Stat.* **2018**, *29*, 98–114. [[CrossRef](#)]
60. Sarabia, R.; Aquino, A.; Ponce, J.M.; López, G.; Andújar, J.M. Automated Identification of Crop Tree Crowns from UAV Multispectral Imagery by Means of Morphological Image Analysis. *Remote Sens.* **2020**, *12*, 748. [[CrossRef](#)]
61. Xiong, J.; Liu, Z.; Chen, S.; Liu, B.; Zheng, Z.; Zhong, Z.; Yang, Z.; Peng, H. Visual detection of green mangoes by an unmanned aerial vehicle in orchards based on a deep learning method. *Biosyst. Eng.* **2020**, *194*, 261–272. [[CrossRef](#)]
62. Wagner, F.H.; Sanchez, A.; Tarabalka, Y.; Lotte, R.G.; Ferreira, M.P.; Aïdar, M.P.M.; Gloor, E.; Phillips, O.; Aragão, L.E.O.C. 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](#)]
63. Mohan, M.; Richardson, G.; Gopan, G.; Aghai, M.M.; Bajaj, S.; Galgamuwa, G.A.P.; Vastaranta, M.; Arachchige, P.S.P.; Amorós, L.; Corte, A.P.D.; et al. UAV-Supported Forest Regeneration: Current Trends, Challenges and Implications. *Remote Sens.* **2021**, *13*, 2596. [[CrossRef](#)]