remote sensing

Article

Recognition of Sago Palm Trees Based on Transfer Learning

Sri Murniani Angelina Letsoin [1,2,*], Ratna Chrismiari Purwestri [3], Fajar Rahmawan [4] and David Herak [1]

[1] Department of Mechanical Engineering, Faculty of Engineering, Czech University of Life Sciences Prague, Kamýcká 129, Suchdol, 16500 Praha, Czech Republic
[2] Faculty of Engineering, University of Musamus, Merauke 99611, Indonesia
[3] Department of Excellent Research EVA 4.0, Faculty of Forestry and Wood Sciences, Czech University of Life Sciences Prague, Kamýcká 129, Suchdol, 16500 Praha, Czech Republic
[4] INTSIA Foundation of Papua Province, Furia 3 Number 116 Abepura, Jayapura City 99225, Indonesia
* Correspondence: letsoin@tf.czu.cz

Abstract: Sago palm tree, known as *Metroxylon Sagu Rottb*, is one of the priority commodities in Indonesia. Based on our previous research, the potential habitat of the plant has been decreasing. On the other hand, while the use of remote sensing is now widely developed, it is rarely applied for detection and classification purposes, specifically in Indonesia. Considering the potential use of the plant, local farmers identify the harvest time by using human inspection, i.e., by identifying the bloom of the flower. Therefore, this study aims to detect sago palms based on their physical morphology from Unmanned Aerial Vehicle (UAV) RGB imagery. Specifically, this paper endeavors to apply the transfer learning approach using three deep pre-trained networks in sago palm tree detection, namely, SqueezeNet, AlexNet, and ResNet-50. The dataset was collected from nine different groups of plants based on the dominant physical features, i.e., leaves, flowers, fruits, and trunks by using a UAV. Typical classes of plants are randomly selected, like coconut and oil palm trees. As a result, the experiment shows that the ResNet-50 model becomes a preferred base model for sago palm classifiers, with a precision of 75%, 78%, and 83% for sago flowers (SF), sago leaves (SL), and sago trunk (ST), respectively. Generally, all of the models perform well for coconut trees, but they still tend to perform less effectively for sago palm and oil palm detection, which is explained by the similarity of the physical appearance of these two palms. Therefore, based our findings, we recommend improving the optimized parameters, thereby providing more varied sago datasets with the same substituted layers designed in this study.

Keywords: classification; deep learning; models; sago

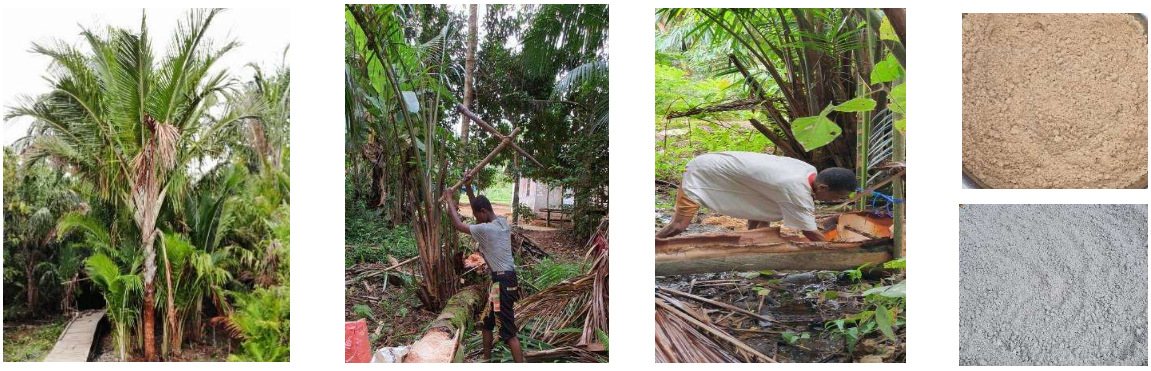
Citation: Letsoin, S.M.A.; Purwestri, R.C.; Rahmawan, F.; Herak, D. Recognition of Sago Palm Trees Based on Transfer Learning. *Remote Sens.* **2022**, *14*, 4932. https://doi.org/10.3390/rs14194932

Academic Editors: Amin Beiranvand Pour, Laura Crispini, Arindam Guha and Snehamoy Chatterjee

Received: 21 August 2022
Accepted: 28 September 2022
Published: 2 October 2022

1. Introduction

Sago palm from the genus *Metroxylon* grows naturally in Asian countries such as Indonesia, specifically in Papua or West Papua Province. This palm has become more important recently since the Indonesian Government is concerned about the role of this palm in various sectors, such as the food industry, as well as other uses [1,2]. Nevertheless, the detection of sago palm trees tends to be challenging due to their comparable features with other plants, for instance, coconut tree or oil palm tree, especially in natural sago forests, where they commonly live together with other particular plants. Therefore, appropriate assessment should be based on their spatial need considerations [3,4]. Results of the previous research regarding land cover changes in the Papua Province of Indonesia and the impact on sago palm areas in the region confirmed that 12 of 20 districts of Merauke Regency in Papua Province tended to lose their potential sago palm habitats. Therefore, one of the recommendations is to attempt to detect and recognize the sago palm [5]. The palm has made significant contributions to supporting local households, for instance, low bioethanol, particularly waste from washing and purifying of sago processing, and food security, specifically from the starches [4,6]. When the harvest time begins, as indicated by the flowers on the top center of the tree, local people will cut the tree and remove the bark, followed by processing to extract the starch. Figure 1a describes the sago palm tree that was captured by using a UAV from our fieldwork, whereas (b) and (c) represent the traditional processes of local farmers at Mappi Regency of Papua Province in Indonesia. The general activity consists of bark removal, pulping, washing, purifying, and subsequent sieving. The visual interpretation will be more demanding as a consequence of the height of the plant, which can be more than 15 m in swampy areas, along with swampy shrubs [7].



(**a**) Sago palm tree from the fieldwork

(**b**) Traditional process during the harvest time

(**c**) Traditional washing and purifying

(**d**) Dry starches

Figure 1. (**a**) Sago palm tree in the fieldwork. (**b**) Traditional bark removal and pulping. (**c**) Washing, purifying, and sieving to get the starch. (**d**) Dried sago starches are ready to use.

Conversely, the advancement of remote sensing technology is quite preferable for solving particular situations, such as detection or recognition. For example, using Unmanned Aerial Vehicle (UAV) data to identify multiple land crops and then classifying these data according to their area, or utilizing nonproprietary satellite imagery tools such as Google Earth Pro to detect the selected object [8,9]. The previous study developed object-based image analysis (OBIA) and image processing using high-resolution satellite imagery for sago palm classification. Nevertheless, the study pointed out some challenges for sago palm classification, for instance, asymmetrical spatial shape due to the semi-wild stand palm, various clumps, and overlapped palm trees [3]. Remote sensing technology using satellite imagery or UAV has been combined with artificial intelligence algorithms or image analytics, supported by various methods, including a deep learning model. As established by [10], detecting the particular species in wetland areas using transfer learning in the stem density system for potatoes [11], or applying deep learning in UAV images to obtain the expected attributes from kiwi fruit, such as location and canopy chain [12], are possible. The deep learning and transfer learning environment have not only been applied in the agricultural sector, but have also been applied towards other objectives, for instance, discovering turbine blade damage [13], crime monitoring systems based on face recognition [14], or energy forecasting using transfer learning and particle swarm optimization (PSO) [15]. Deep learning is a sub-field of machine learning where the model is derived by an artificial neural network structure [16]. Recently, deep learning has been successfully applied in plant detection, for instance, tree crown detection, which generally could be performed by three approaches, i.e., Convolutional Neural Network (CNN), semantic segmentation, and object detection using YOLO, etc. [17]. This current study uses deep learning based on a CNN, which consists of an input layer, convolution layers, pooling layers, fully connected layers, and an output layer. This network's detection system proved superior to other machine learning methods [18,19]. In an image classification task, the machine learning model takes different feature extraction of images, for instance, shape, height, etc., then moves to the classification step. Meanwhile, a medium or huge amount of dataset should be available. In contrast, deep learning obtains the images without a manual feature extraction step. Feature extraction and classification work through model layers; therefore, deep learning requires a large amount of data to achieve a good performance. It takes a long time to train the model and less time to test it. Since machine learning and deep learning need a more extensive dataset for training, they will require a higher hardware memory capacity [20]. To overcome the lack of data and the time cost consumed during training, transfer learning can be applied by using a deep learning model. Several earlier studies defined transfer learning as a technique using a model that has been trained for one task, which is then used as a baseline to train a model for another typical assignment; as long as the target model is in the same domain [21]. There are three main strategies for doing transfer learning on deep learning models, i.e., using pre-trained models, applying feature extraction by discarding fully connected output layers, and fine-tuning the last layers on pre-trained models [22–24].

Numerous deep learning networks based on CNN have been widely elaborated, for instance, GoogLeNet and DenseNet. Nonetheless, as explained before, the existing model can be modified for other purposes but has not yet been investigated for sago palm detection. Furthermore, transfer learning is acceptable with fewer data and could reduce training time and computer resources, as concluded by an earlier study. Therefore, the current study will use a transfer learning strategy to predict the plants based on their physical appearances, such as leaves, trunks, flowers, or fruits. In order to address this, three different pre-trained networks based on CNN were customized for detection and prediction; namely, SqueezeNet, AlexNet, and ResNet-50, were applied in this study. We modified the last layer and discarded the fully connected output layer to achieve our new task. The study's dataset consists of data training and testing of three plants: sago palm, coconut tree, and oil palm tree. Each class is categorized based on features such as leaves, fruits or flowers, and trunks. The study aims (1) to obtain the prior model based on classification performance, i.e., precision, F1-score, and sensitivity; and (2) to evaluate the transfer learning task in sago palm detection based on leaves, flowers, and trunks.

2. Materials and Methods

2.1. Study Region

The study was performed in Merauke Regency (Figure 2), which is located in the southern part of Papua Province of Indonesia (137°38′52.9692″E–141°0′13.3233″E and 6°27′50.1456″S–9°10′1.2253″S). In the last decade, the population growth in Papua Province was around 18.28%, and approximately 1.20 million people there are economically active in agriculture. Based on weather data, the annual minimum and maximum temperature average deviates between 16–32 °C, while the average rainfall registered is 2900 mm with high humidity from 62% to 95%.

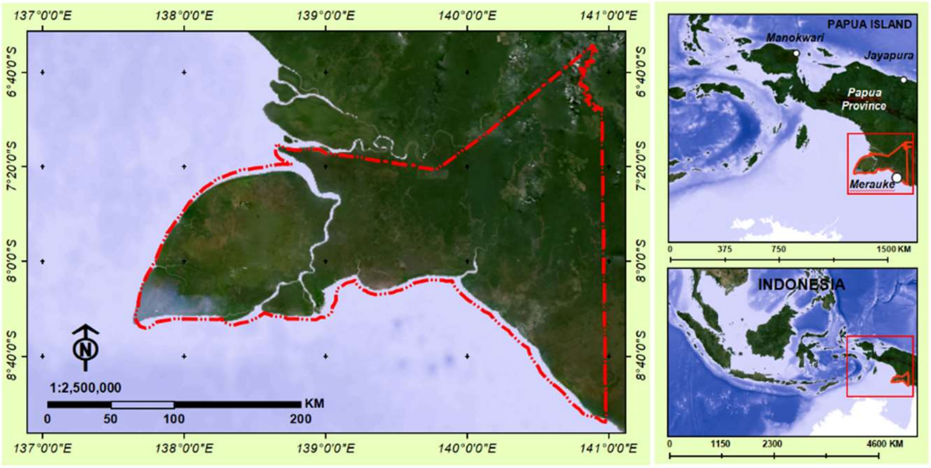
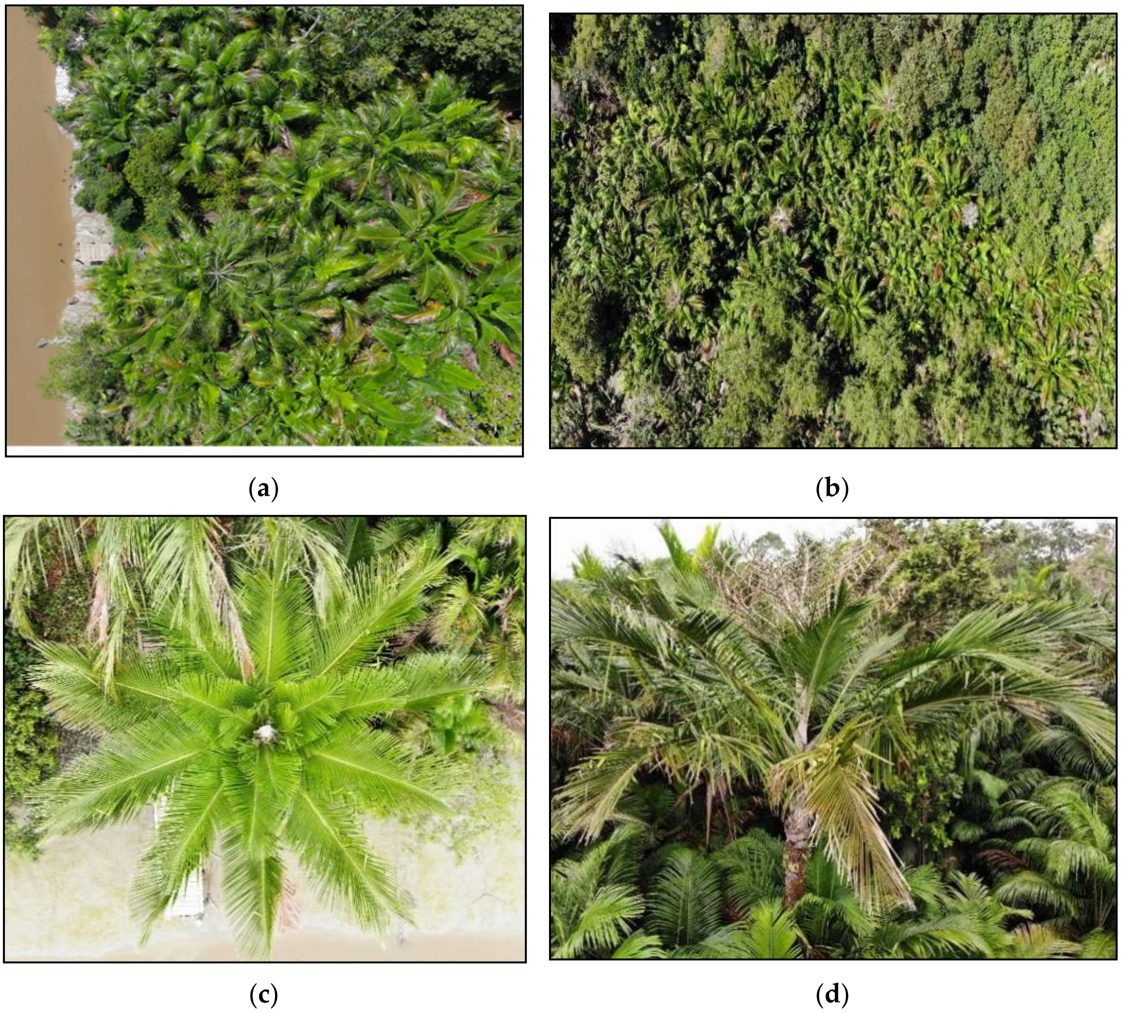


Figure 2. Location of study area: Merauke Regency. Consisting of 20 districts that cover an area of 46,791.63 km^2 with a population of 230,932 in 2020.

2.2. UAV Imagery

Unmanned aerial imagery was captured in Tambat Village, located in Merauke Regency, specifically, the Tanah Miring District. This fieldwork was performed at one of the top sago forest producers in the Merauke Regency. The sago forest images, shown in Figure 3, were captured at a height of 60 m and 100 m, longitude, and latitude of 140°36′46.8498″E–8°21′21.2544″S, respectively. The visible morphology of the sago palm as detected by the UAV is presented in Figure 3b–d. Sago palm in this fieldwork is typically natural sago forest or wild sago. It contains a palm trunk, which stores the starch. At harvest time, the trunk will be cut off and the bark will be opened, followed by further processing to extract the starch. The harvest time of these sago forest areas is commonly identified by the bloom of the flowers, as introduced in Figures 1 and 3, followed by the leaves. The dataset used in this research consists of high-resolution RGB images taken from a UAV by an Autel Robotics multi-copter. Additionally, field survey data were obtained by performing ground photography and a short interview with local sago farmers. Our study focuses on the morphology of sago rather than sago palm health classification or automatic counting; therefore, our dataset also shows other typical plants such as coconut tree and oil palm, based on their leaves, fruits, and trunks.



(**a**) (**b**)

(**c**) (**d**)

Figure 3. Sago palm captured imagery using UAV RGB and ground photography observation. (**a**,**b**) Sago palm areas in the fieldwork, and other vegetations; (**c**) Sago flowers defined by white flowers at the top center, between leaves; (**d**) Palm tree dominant features: trunk and leaves.

The methodology of this research was developed as presented in Figure 4. First, study preparation is established throughout the field survey and ground photographs around the fieldwork. We used some tools such as Google Earth engine, and a handheld global positioning system to ensure the location of the fieldwork. Then, we created the mission plan for the UAV. In the next stage, the UAV Red Green Blue (RGB) band images were downloaded and labeled. Next, the region of interest (ROI) was chosen based on each label category and class. The dataset in this study was divided into two types: (1) data trained and (2) data tested. The data trained were categorized into nine classes, namely coconut tree trunks, leaves, and fruits, as well as oil palm trunks, leaves, and fruits. The remaining classes are sago trunks, leaves, and flowers. Considering the classification process and prediction model, we applied deep learning model approaches, namely SqueezeNet, AlexNet, and ResNet-50. Hence, this dataset also involves various sizes of the imagery, blurred and yellowish images with different angles.

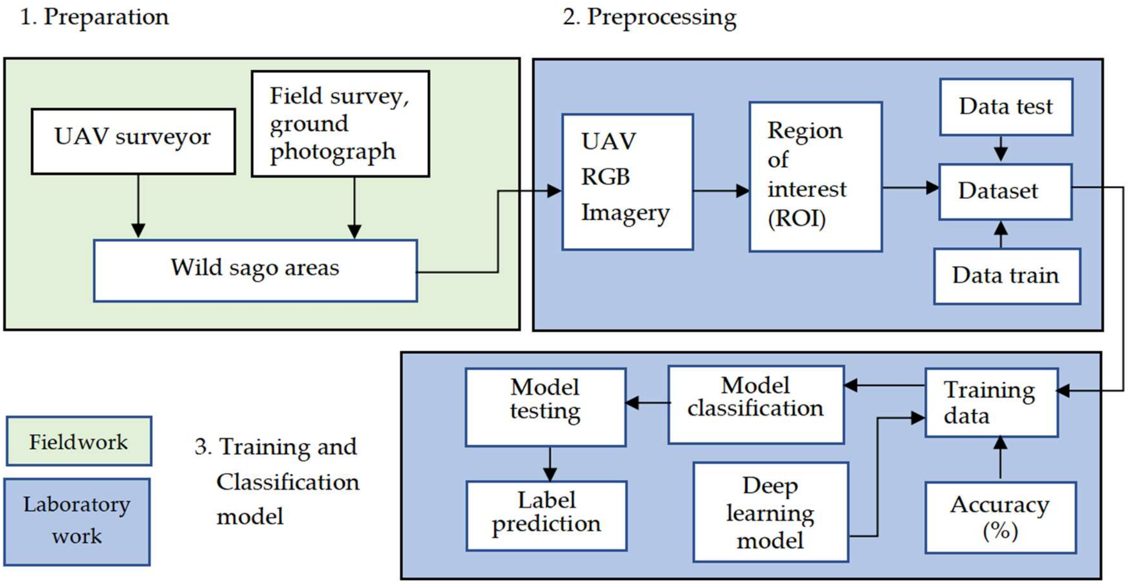


Figure 4. Study workflow. Three stages are practiced with classification and prediction based on three major features of each plant.

The classification and prediction process began after the data collection and training data were developed. Subsequently, the deep learning models were applied, including parameter optimization for instance mini batch size, initial learning rate or epoch. The earlier study has successfully combined the parameter optimization to obtain the higher performance in classification task, for example, learning rate was set up to 0.0001 with ten number of epochs [23,25]. Further, it was trained and tested using a dataset utilized from the previous stage. The accuracy of the sago palm dataset was evaluated by comparing the results of drone imagery with actual data derived from the visual interpretation, and was based on the ground survey and photographs. All of the training and testing procedures were implemented using MATLAB R2021 and deep learning scripts.

2.3. Deep Learning and Transfer Learning Models

Deep Learning (DL) models were defined throughout the pre-trained networks of several common architectures already provided in MATLAB packages for DL, such as AlexNet [26], GoogLeNet [27], ResNet-50 [28], and Inception v3 [29]. This study focused on three networks based on CNN, namely SqueezeNet, AlexNet, and ResNet-50. SqueezeNet consumes small bandwidth to export new models to the cloud or upload them directly to devices. This network is also able to deploy on embedded system or other hardware devices that have low memory [30], while AlexNet shows higher accuracy compared to other different DL, such as GoogleNet or ResNet152, in image classification on the ImageNet dataset [31], however ResNet as the backbone network shows good performance for the segmentation dataset [32]. The transfer learning (TL) strategy used in this study requires two stages: the base model, which is constructed on pre-training CNN models, and the target model [33], which is tailored to a new, specific task (Figure 5). Three pre-trained networks were used, namely SqueezeNet, ResNet, and AlexNet for the base model, and then we reconstructed the base model to our target model, with nine probability classes. These three models are trained in various datasets, such as the ImageNet Dataset, and are able to classify images into 1000 object categories or 1000 classes [34,35], such as keyboard, mouse, coffee mug, pencil. Nevertheless, TL allows a small dataset, reuse, and extraction of transfer features, and improves the training time of models [21,23,25].

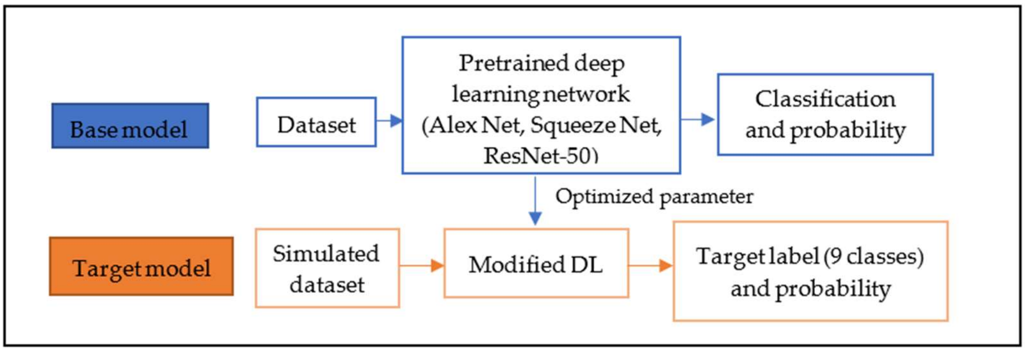

Base model
Dataset
Pretrained deep learning network (Alex Net, Squeeze Net, ResNet-50)
Classification and probability
Optimized parameter
Target model
Simulated dataset
Modified DL
Target label (9 classes) and probability

Figure 5. Transfer learning workflow in this study.

AlexNet network involves 5 convolutional layers (conv1–conv5), 3 fully connected layers (fc6, fc7, fc_new) within the ReLu layer are established after every convolution layer (Table 1). Further, the dropout layer (0.5 or 50%) avoids overfitting problems. According to the tools used in this study, the input size is 227 × 227 × 3 or 154,587 values [36], and all layers must be connected.

Table 1. AlexNet designed in this study.

Layer	Layer Name	Layer Type	Layer Details
1	Data	Image input	227 × 227 × 3 images with zero center normalization
2	Conv1	Convolution	96 11 × 11 × 3 convolutions with stride [4 4] and padding [0 0 0 0]
3	Relu1	ReLU	ReLU
4	Norm1	Cross channel normalization	Cross channel normalization with 5 channels per elemen
5	Pool1	Max pooling	3 × 3 max pooling with stride [2 2] and padding [0 0 0 0]
6	Conv2	Grouped convolution	2 groups of 128 5× 5 × 48 conv with stride [1 1] and padding [2 2 2 2]
7	Relu2	ReLU	ReLU
8	Norm2	Cross channel normalization	Cross channels normalization with 5 channels per element
9	Pool2	Max pooling	3 × 3 max pooling with stride [2 2] and padding [0 0 0 0]
10	Conv3	Convolution	384 3 × 3 × 256 convolutions with stride [1 1] and padding [1 1 1 1]
11	Relu	ReLU	ReLU
12	Conv4	Grouped convolution	2 groups of 192 3 × 3 × 192 convolutions with stride [1 1] and padding [1 1 1 1]
13	Relu4	ReLU	ReLU
14	Conv5	Grouped convolution	2 groups of 128 3 × 3 × 192 convoutions with stride [1 1] and padding [1 1 1 1]
15	Relu5	ReLU	ReLU
16	Pool5	Max pooling	3 × 3 max pooling with stride [2 2] and padding [0 0 0 0]
17	Fc6	Fully connected	4096 fully connected layer

Table 1. *Cont.*

Layer	Layer Name	Layer Type	Layer Details
18	Relu6	ReLU	ReLU
19	Drop6	Dropout	50% dropout
20	Fc7	Fully connected	4096 fully connected layer
21	Relu7	ReLU	ReLU
22	Drop7	Dropout	50% dropout
23	Fc_new	Fully connected	9 fully connected layer
24	Prob	Softmax	
25	Classoutput	Classification output	

SqueezeNet starts the network with an individual convolution layer (conv1), then a rectified linear unit (ReLU), which is a type of activation function, then the max pooling layer (Figure 6). When added to a model, max pooling reduces the dimensionality of images by decreasing the number of pixels in the output from the previous layer. Thus, the Conv+Relu layer is then extended to 8 fire modules, from fire 2 to fire 9, with a filter size combination of 1 × 1 and 3 × 3 [30]. Convolution and ReLU layer can be computationally defined as follows:

$$F_l(I) = P_{i-1}(G) \times f \tag{1}$$

where F_l describes an output feature map and l represents the lth convolution layer, while f is defined by filter size or kernel, and then P_{i-1} shows the previous layer output, and I denotes the original data image. Thus, ReLU is denoted through an equation:

$$ReLU(I) = \max(x, 0) \tag{2}$$

x is the input of activation on the lth layer, I denotes a ReLU activation output of the feature maps [37].

Another network included in this study is the Residual network (ResNet-50), the variant of the ResNet model, which has a 50-layer deep convolutional network. It contains one convolution kernel size of 7 × 7 at the beginning and ends with an average pool, a fully connected layer, and a SoftMax layer in one layer. Between these layers, there are 48 convolutional layers consisting of different kernel sizes [38]. Here, the fully connected layer′s purpose is to integrate all of the inputs from one layer connecting to every activation unit of the next layer. Thus, the residual block on the ResNet equation is as follows, where O is the output layer, I is the input layer, and $F(I_i\ W)$ is the residual map function [39].

$$O = F(I_i\ W_i\ + I) \tag{3}$$

The characteristics of each model are shown in Table 2, as follows.

Table 2. Model comparison [1].

Network Name	Depth	Image Input Size	Parameters (Millions)	Total Memory (MB)
SqueezeNet	18	227 × 227 × 3	1.24	5.20
ResNet-50	50	224 × 224 × 3	25.6	96
AlexNet	8	227 × 227 × 3	61	227

[1] Based on tool used in this study, ie., MATLAB.

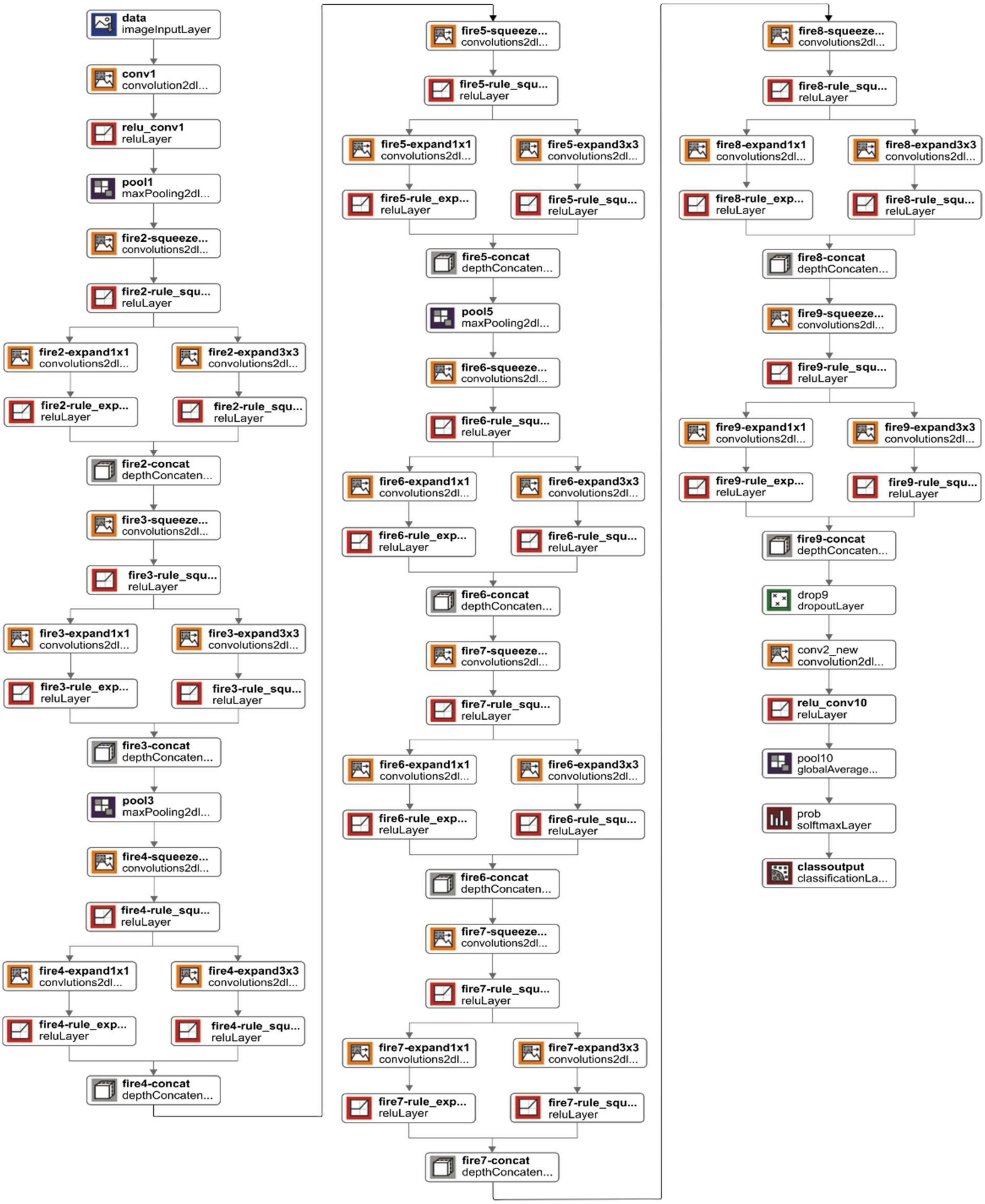
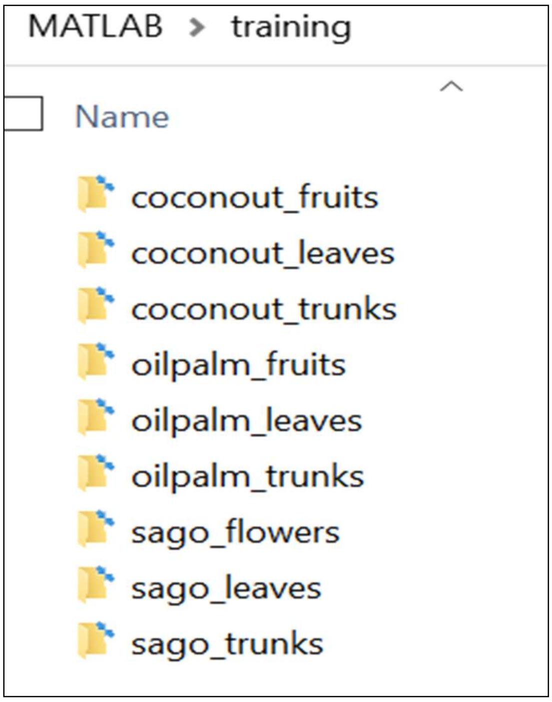


Figure 6. SqueezeNet used in this study: all layers are connected.

Once the data preparation is ready and the deep learning model has been designed, we can analyze the chosen model and optimize the parameters. If there are no errors in the model, then all of the training data can be imported as data trained (Figure 7).



MATLAB > training

Name

coconout_fruits
coconout_leaves
coconout_trunks
oilpalm_fruits
oilpalm_leaves
oilpalm_trunks
sago_flowers
sago_leaves
sago_trunks

Figure 7. Simulated data train: CF, CL, CT, OPF, OPL, OPT, SF, SL, ST, respectively.

As a result, validation accuracy appears, including the training time (elapsed time) and training cycle, such as the epoch number. Optimized parameters from deep pre-trained networks are transferred to a simulated dataset and then will be trained. These models are compared using the same number of epochs, learning rate, and batch size. At the final stage, 227 × 227 and 224 × 244 image input in the data test will be resized, then a single image as an output will be predicted and categorized (Figure 8).

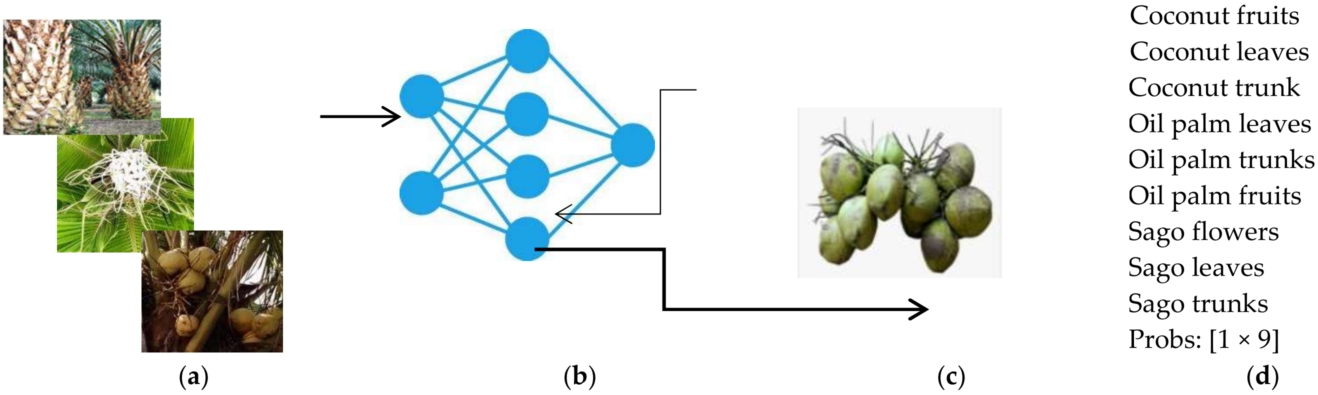


Figure 8. Examples from processing workflow. (**a**) All data trained images. (**b**) Model and optimized parameters via transfer learning. (**c**) Test image cropped (227 × 227) or (224 × 224). (**d**) Classification and probability.

2.4. Performance Evaluation

Four metrics are typically evaluated in DL and TL model evaluation, namely true positive (TP), true negative (TN), false positive (FP), and false negative (FN). In this study, TP and TN describe the correct identification of class, while FP and FN correspond to false identification of class [40]. The evaluation was investigated using an image from the validation set and their specific labels, which were not used for training. The detection ability is assessed based on the precision and the sensitivity criteria, as shown in Table 3, while the optimized parameters applied in this study are presented in Table 4.

Table 3. Evaluation criteria.

Metric	Formula	Criteria
F1-score	$\frac{2\times(Recall\ \times\ Precision)}{Recall\ +\ Precision}$	Denotes a high value, which validates the model.
Precision	$\frac{TP}{TP\ +\ FP}$	Examines the ability of the model to predict positive label.
Sensitivity (Recall)	$\frac{TP}{TP\ +\ FN}$	Defines the ability of the model to detect instances of certain classes well.
Specificity	$\frac{TN}{FP\ +\ TN}$	Defines the true negatives that are correctly identified by the model.
Accuracy	$\frac{TP\ +\ TN}{TP\ +\ FP\ +\ TN\ +\ FN}$	Examines the accurately in identifying the images to the classes.

Table 4. Optimized parameters in this study.

Parameter Name	Value
Epochs	10
Initial learning rate	0.0001
Validation frequency	9
Learning rate weight coefficient	10
Learning rate bias coefficient	10
Learning rate schedule	Constant
Momentum	0.9
L2 Regulation	0.0001
Min batch size	10

Multi-class detection can be explained as follows, for instance, in sago flowers:

1. TP, the number of actual images that are displaying sago flowers (true) and are classified as sago flowers (predicted).
2. FP, the number of actual images that are not displaying sago flowers (not true) and are classified as sago flowers (predicted).
3. FN, the number of actual images that are displaying sago flowers (true) and are classified as a different class (predicted).
4. TN, the number of actual images that are not displaying sago flowers (not true) and are classified as a different class (predicted).

Hyperparameters set in the training model of TL (Table 4) were determined from the earlier studies [22,23] by epochs, batch size, and learning rate. A very high learning rate will trigger the loss function to go up, and as a result, the accuracy of classification can be reduced. Conversely, if it is too low, it will reduce the network training speed, the correction of weight parameter correction will be slow, and it will fail to obtain a proper model accuracy. Batch size is also vital to the accuracy of models and the training process performance. Using a larger batch size will require higher GPU memory to store all of the variables (e.g., momentum), and the weights of the network also may cause overfitting; however, using a minimum batch size may lead to slow convergence of the learning algorithm. Another technique to overcome the GPU memory limitation and run large batch sizes is to split the batch into mini-batch sizes. The number of epochs defines the learning algorithm will complete passes through the entire training dataset.

An ANOVA test was employed to compare means between the accuracy (true positive) values of three models in correctly identifying the target trees' morphology. A p value less than 0.05 was designed for a statistically significant difference in all data analyses. A receiver–operating characteristics curve (ROC curve) was employed to identify the sensitivity and 1-specificity (false positive) of the three algorithms in identifying sago (flowers, leaves, and trunks) over coconut and oil palm. All data analyses were performed using the IBM SPSS version 27 (IBM Corp., Armonk, NY, USA). Additionally, an approximate cost of software measurement was estimated using function point (*FP*) analysis [41], which is described the functionality points through complexity adjustment factor (*CAF*), and unadjusted functional point (*UFP*), as follows [42]:

$$CAF = 0.65 + (0.01\ x \sum_{k=1}^{14} F_i\ x\ rating) \tag{4}$$

$$FP = UFP \times CAF \tag{5}$$

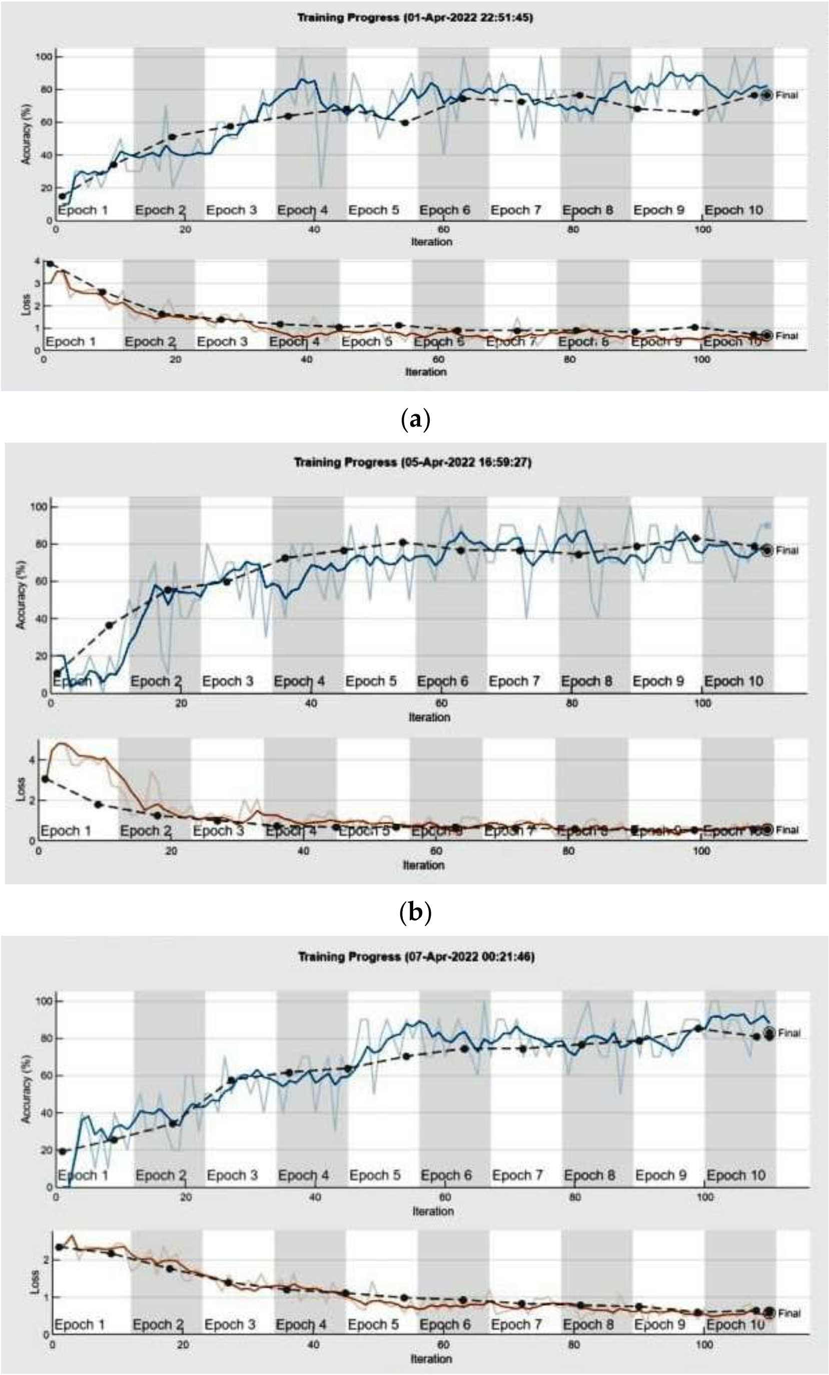
3. Results

3.1. Dataset Development

In our fieldwork, the Autel UAV flew at various altitudes of 60 m up to 100 m, as well as with different forward and sideways overlapping, during the mornings and mid-days of July 2019, August 2021, and July 2022. This stage aims to obtain various shapes as well as to enrich the dataset images, instead of counting the plants. Next, the data collection of sago palm was downloaded and cropped, then allocated according to the labels in Figure 7. The experiments were processed with an Intel Core i7 processor, with the dataset defined in RGB space being categorized in 9 classes. The data train and data test were divided separately, as presented in Figure 4, stage 2; around 70% was allocated for data training, and 30% were used for data testing and validation. This study comprised 231 total images; 68 images for testing and the rest for data training. The same dataset was used to train and test, based on the deep learning networks used. All images were pre-processed based on the three pre-trained networks, as compared in Table 2. Three pre-trained networks of deep learning were examined and then modified. Then, they were transferred to the target as transfer learning, of which the modified version is shown in Figure 5. Regarding the new task via transfer learning, as well as to obtain the aims of this study, the last layers of each model were reformed as follows: the fully connected layer, fc1000 changes to fc_new, then SoftMax layer for converting values into probabilities, and subsequently, the classification layer predictions for 1000 output size were replaced to class_output for categorizing into nine classes. Convolution1000 layer is also restored to the conv2d layer with nine num-filters. Then, the weight learning factor and bias learn rate coefficients, as presented in Table 4. Furthermore, all images were pre-assessed using resizing and normalization, for instance, rescaling, rotation, and augmentation. In addition, the datasets were all evaluated by inspecting the UAV images, visual interpretation, and ground surveys, such as photographing plants.

3.2. Training and Testing Data Performance

Considering the TL process as presented as the workflow in Figure 5, all datasets were imported into a specified workplace through MATLAB and followed by other processing, namely, training data in the modified deep learning pre-trained network. As a result, training accuracy and validation lost over ten epochs and ten min batch sizes are introduced in Figure 9. The smoothness of the accuracy and the loss of training process are described by the blue colour and orange colour, respectively. While light blue coloured dots and light orange coloured dots represent the training progress. Additionally, the validation of data trained and the loss are explained by black line coloured dots. The training progress of the three models was not quite as smooth, with accuracies of 76.60%, 76.60%, and 82.98%. However, the ResNet-50 model is more dominant when compared to the others, with the highest accuracy of 82.98%. The training loss values decreased sharply on these models in epoch 5, while the training progress increased. Subsequently, the validation accuracy and loss curves were more eased, especially in ResNet-50 and AlexNet, where the data training loss decreased during the rest of the process. Although SqueezeNet and AlexNet fluctuated after 5 epochs, AlexNet network validation was improved, while the training loss was smaller. This result demonstrates the ability of the three classifiers in recognizing the dataset.



(**a**)

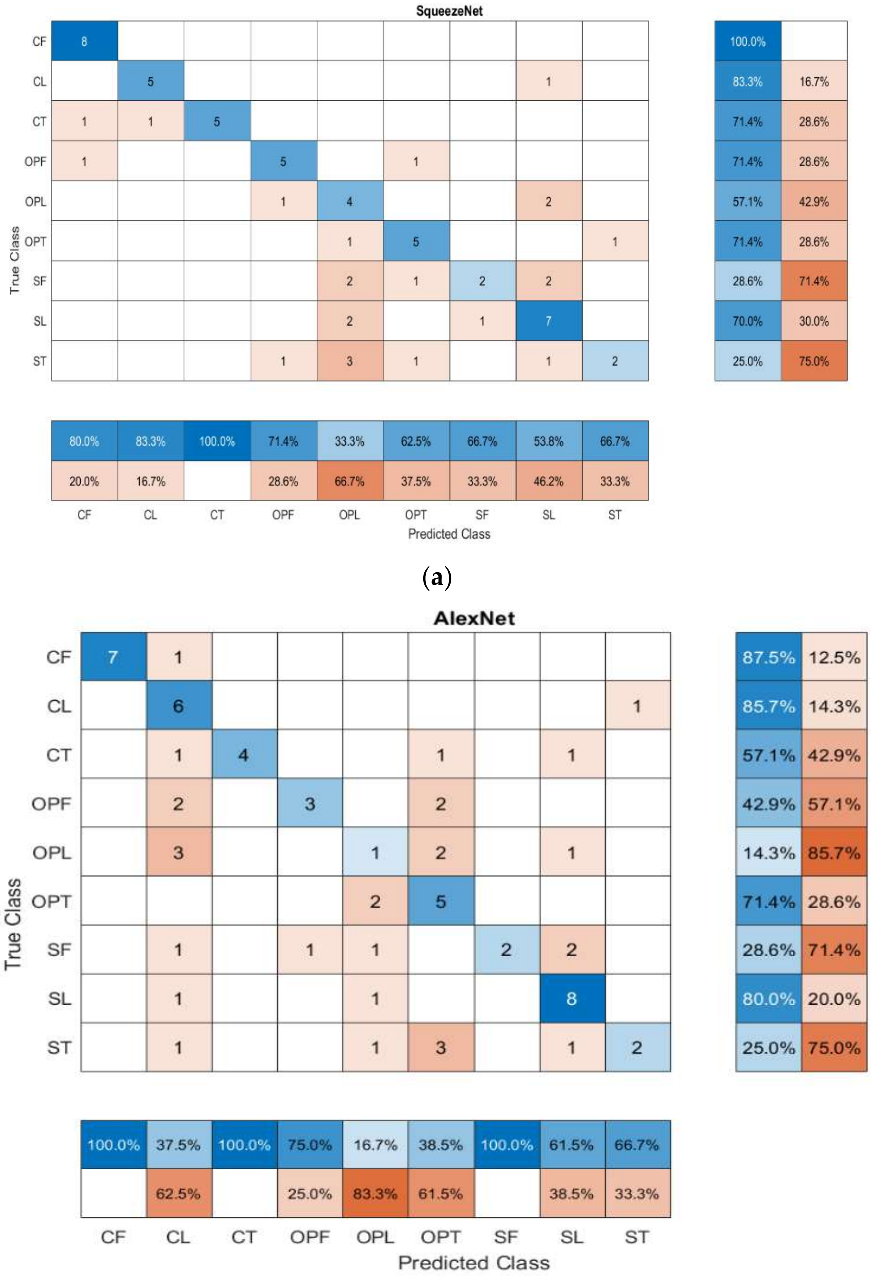
(**b**)

(**c**)

Figure 9. Training accuracy: (**a**) SqueezeNet, accuracy of 76.60%. (**b**) AlexNet, accuracy of 76.60%. (**c**) ResNet-50, the accuracy of 82.98%.

After the training process shown in Figure 9, all models were tested using the same data test, which was prepared and placed differently than the trained data. To support this testing process, we used various types of syntax that were accommodated in MATLAB2021, such as imresize, imshow, prediction, probability, and confusion matrix. The uses of imresize and imshow are basically appropriate preparation for the input test, according to the features of each model in Table 2, while the probability and categorization were generated from each model, specifically in layer name: prob, within SoftMax type (Figure 6). Next, the confusion matrix was calculated for each classification model, and the performance was visualized using the values on the confusion matrix. The confusion matrix in this study was used to describe each model's performance, consisting of the true class and the class predicted by the model. Then, the metric was calculated based on the formula shown in Table 3.

All models were able to predict the plants with 100% accuracy, such as SqueezeNet (Figure 10) for recognizing the coconut trunk (CT), AlexNet for coconut fruit (CF), coconut trunk (CT), and sago flowers (SF), while ResNet-50 recognized oil palm trunk (OPT). In the case of sago palm classification, the convolution matrix of AlexNet and ResNet-50 were superior to SqueezeNet. Despite the fact that the models were trained with a self-contained dataset and smaller datasets compared to pre-trained deep learning when utilizing 1000 images, the training accuracy rose to 82%. Meanwhile, the models obtained the expected results in the recognition of the plant's physical morphology.



SqueezeNet

True Class	CF	CL	CT	OPF	OPL	OPT	SF	SL	ST		
CF	8									100.0%	
CL		5						1		83.3%	16.7%
CT	1	1	5							71.4%	28.6%
OPF	1			5		1				71.4%	28.6%
OPL				1	4			2		57.1%	42.9%
OPT					1	5			1	71.4%	28.6%
SF					2	1	2	2		28.6%	71.4%
SL					2		1	7		70.0%	30.0%
ST				1	3	1		1	2	25.0%	75.0%
	80.0%	83.3%	100.0%	71.4%	33.3%	62.5%	66.7%	53.8%	66.7%		
	20.0%	16.7%		28.6%	66.7%	37.5%	33.3%	46.2%	33.3%		

Predicted Class

(**a**)

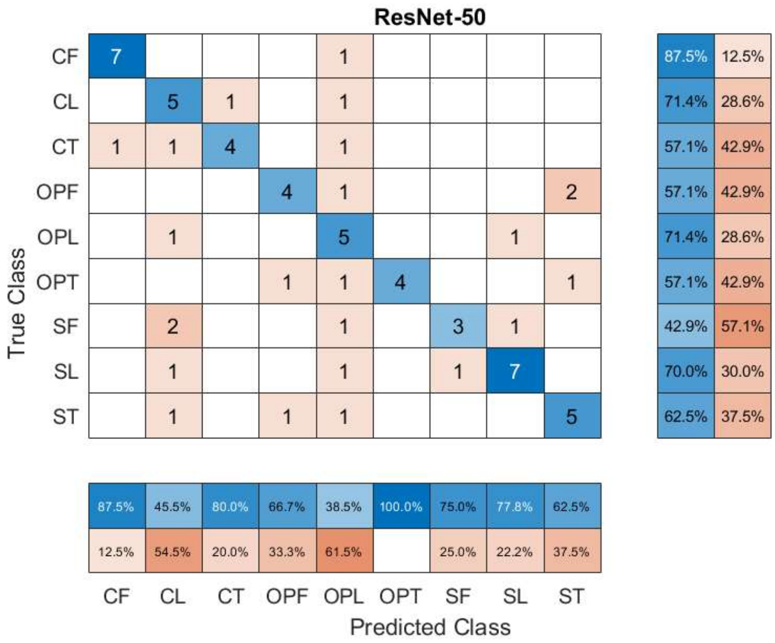
AlexNet

True Class	CF	CL	CT	OPF	OPL	OPT	SF	SL	ST		
CF	7	1								87.5%	12.5%
CL		6							1	85.7%	14.3%
CT		1	4			1		1		57.1%	42.9%
OPF		2		3		2				42.9%	57.1%
OPL		3			1	2		1		14.3%	85.7%
OPT					2	5				71.4%	28.6%
SF		1		1	1		2	2		28.6%	71.4%
SL		1			1			8		80.0%	20.0%
ST		1			1	3		1	2	25.0%	75.0%
	100.0%	37.5%	100.0%	75.0%	16.7%	38.5%	100.0%	61.5%	66.7%		
		62.5%		25.0%	83.3%	61.5%		38.5%	33.3%		

Predicted Class

(**b**)

Figure 10. *Cont.*



(**c**)

Figure 10. Confusion matrix of predictions made by TL on (**a**) SqueezeNet; (**b**) AlexNet; (**c**) ResNet-50.

3.3. Model Performance Evaluation

In terms of precision and detection of the sago palm based on leaves (SL), flowers (SF), and trunks (ST), the best performances are highlighted in bold (Table 5). Even though the AlexNet model is 100% able to detect sago flowers (SF), the sensitivity of this model or valid positive rate is only around 29%. Conversely, the ResNet-50 model is quite good as a classifier in SL and ST, with the precision value for SL and ST being 0.78 and 0.83, respectively. Precision and sensitivity should preferably be 1, which is the highest value, or close to 100% if expressed in percentage. Nevertheless, the F1 score turns out to be ideally 1 when both precision and sensitivity are increased or equal to 1. Therefore, this study examines the precision, as well as the recall or sensitivity to evaluate the performance of each model as a good classifier.

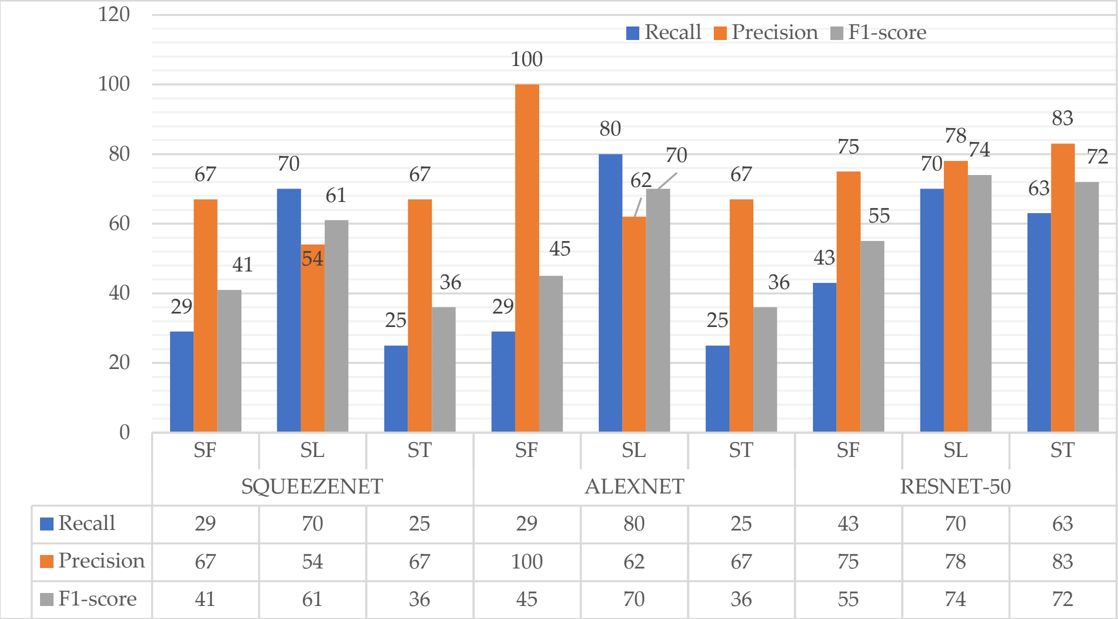
Table 5. Classification results of three networks.

Model	Training Accuracy (%)	Training Time	Image Input Size	Class	Recall (Sensitivity)	Precision	F1 Score
SqueezeNet	76.60	3 min 39 s	227 × 227	CF	1.00	0.80	0.89
				CL	0.83	0.83	0.83
				CT	0.71	1.00	0.84
				OPF	0.71	0.71	0.71
				OPL	0.57	0.33	0.42
				OPT	0.71	0.63	0.67
				SF	0.29	0.67	0.41
				SL	0.70	0.54	0.61
				ST	0.25	0.67	0.36
AlexNet	76.60	5 min 8 s	227 × 227	CF	0.88	1.00	0.94
				CL	0.86	0.38	0.53
				CT	0.57	1.00	0.73
				OPF	0.43	0.75	0.55
				OPL	0.14	0.17	0.48
				OPT	0.71	0.39	0.51
				SF	0.29	1.00	0.45
				SL	0.80	0.62	0.70
				ST	0.25	0.67	0.36

Table 5. *Cont.*

Model	Training Accuracy (%)	Training Time	Image Input Size	Class	Recall (Sensitivity)	Precision	F1 Score
ResNet-50	82.98	18 min 29 s	224 × 224	CF	0.88	0.88	0.88
				CL	0.71	0.46	0.56
				CT	0.57	0.80	0.67
				OPF	0.57	0.67	0.62
				OPL	0.71	0.39	0.50
				OPT	0.57	1.00	0.73
				SF	**0.43**	**0.75**	**0.55**
				SL	**0.70**	**0.78**	**0.74**
				ST	**0.63**	**0.83**	**0.72**

SqueezeNet performed significantly better in identifying the leaves of oil palm (OPL) than AlexNet (p = 0.046); meanwhile, no statistical significance difference was found between the tree models in accurately recognizing the target tress based on fruits, leaves, and trunks. Based on the accuracy values, the AlexNet had a more surprising performance in the detection of sago flower (SF) than the other models (Figure 11), while ResNet-50 can recognize the sago trunk (ST) and sago leaves (SL) better than other models. These results indicate that the models can distinguish sago palms from other plants used in this study. Based on this evaluation, the AlexNet and ResNet-50 can promote the preliminary detection of the sago palm.

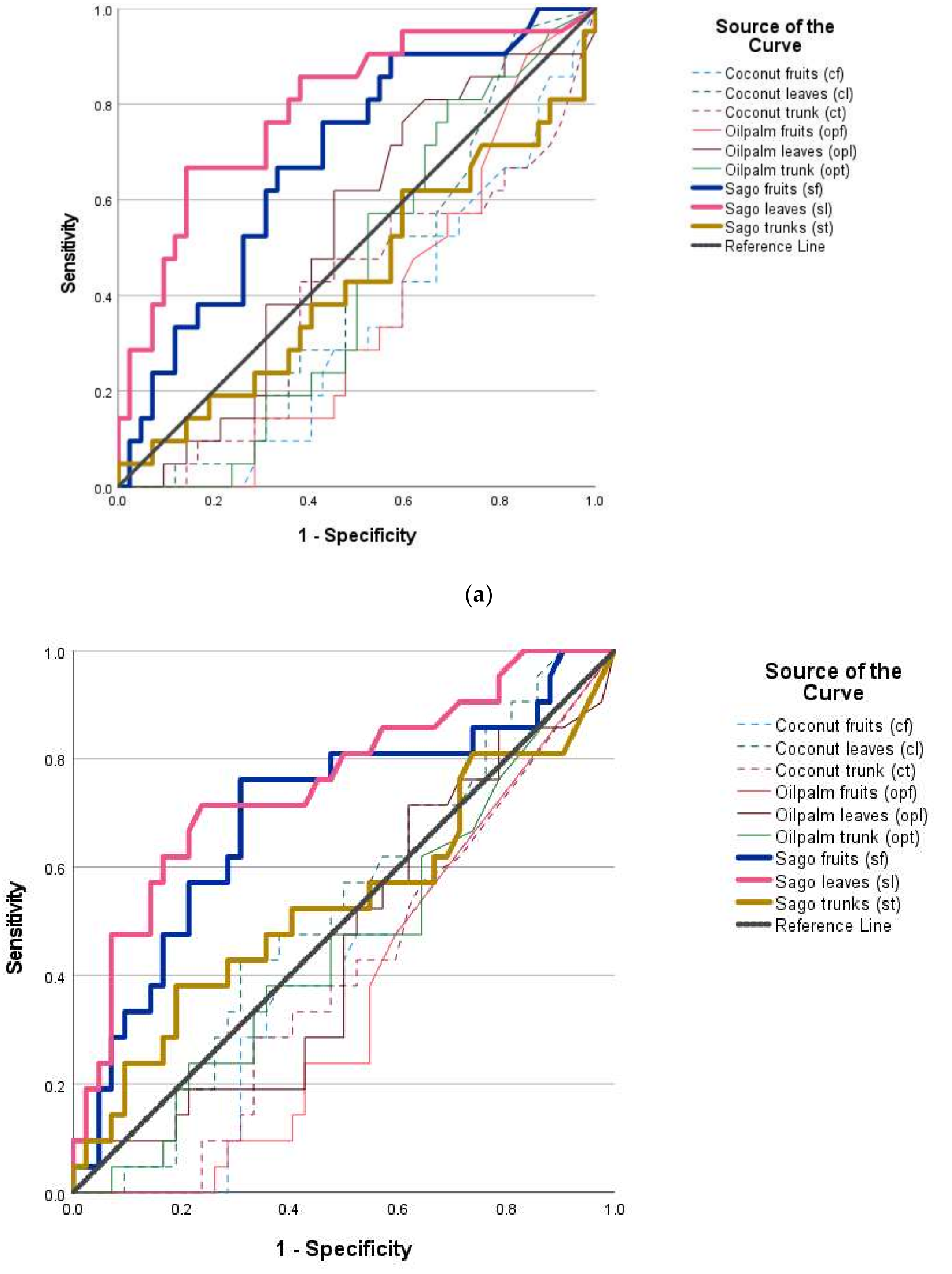


	SQUEEZENET SF	SQUEEZENET SL	SQUEEZENET ST	ALEXNET SF	ALEXNET SL	ALEXNET ST	RESNET-50 SF	RESNET-50 SL	RESNET-50 ST
Recall	29	70	25	29	80	25	43	70	63
Precision	67	54	67	100	62	67	75	78	83
F1-score	41	61	36	45	70	36	55	74	72

Figure 11. Performance of sago palm classifier in percentage (%).

ROC curves compared all experimented models on the sago palm dataset, the results showed that all algorithms were able to correctly identify sago (flowers and leaves) over coconut and oil palm (Figure 12a–c), of which ResNet50 presented the best model for predicting sago trees. SqueezeNet and ResNet 50 could also distinguish between the sago trunks over that of the coconut and palm oil, however, AlexNet was less likely to identify it (as depicted by the line under the reference values) (Figure 12a).

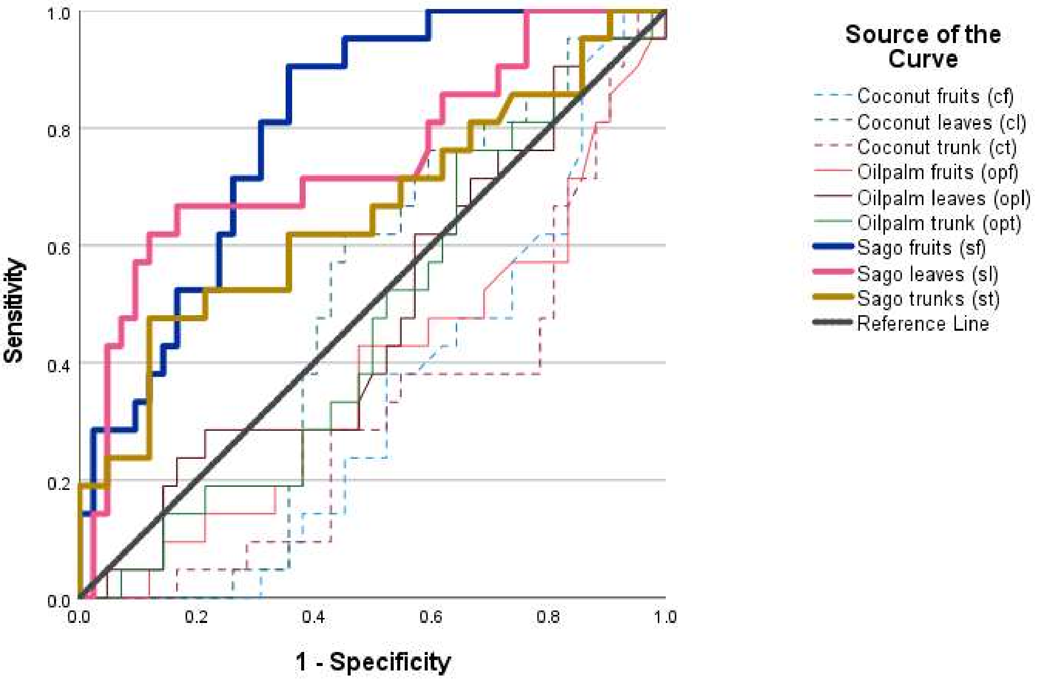
The cost analysis of model implementation according to *FP* was estimated with the result of *CAF* and *FP*, i.e., 1.21 and 1057.54 *FP*, respectively. Comparable other costs revealed similar expenses for the further deployment of the model. The only difference between them was found in the performance of three models, as presented above (Figures 11 and 12).



(**a**)

(**b**)

Figure 12. *Cont*.



(**c**)

Figure 12. (**a**) ROC curve of the AlexNet. (**b**) ROC curve of the SqueezeNet. (**c**) ROC curve of the ResNet50.

4. Discussion

The implementation of deep learning can be performed with two methods: (a) a self-designed deep learning model, and (b) transfer learning approaches. In this study, transfer learning based on three models, namely, SqueezeNet, AlexNet, and ResNet-50 were used as transfer objects to categorize and predict the three plants based on their physical morphology. Generally, the three models can detect the morphology of coconut trees well, specifically SqueezeNet, as shown by the higher precision in CF, CL, and CT—80%, 83%, 100%, respectively, when compared to sago palm or oil palm. This happens because the shape of palm oil and sago are similar, as shown in Figures 1 and 3, especially when captured by using a drone or other remote sensing technology at a specific altitude [3]. As investigated by [43], tree classification using UAV imagery and deep learning has confirmed that deep learning and transfer learning can apply to the classification of UAV imagery, however, more tree species and various study areas will improve the accuracy of the classifiers. Concerning the performance of sago classifiers, as shown in Figure 11, AlexNet can predict sago flowers (SF) at 100%, while ResNet-50 forecasts sago leaves (SL) and sago trunk (ST) at 78% and 83%, respectively. A different study of wood structure found that the testing performance of ResNet-50 as a transfer object was about 82% from 4200 images of the dataset [44,45]. Additionally, for carrot crop detection, which included 1115 standard-quality carrots and 1330 imperfect carrots using ResNet-50, it was proven that this transfer method is superior compared to the others. Even though TL can predict the class with fewer datasets, it can provide a variety of sago palm datasets that will improve the learning performances [46]. Therefore, providing more datasets with different types, angles, and shapes of the sago is recommended for further work.

Considering the availability of datasets from UAV imagery, findings have been provided by several studies, as mentioned in [40,47], for instance, weed map dataset, VOAI dataset, or other resources, such as ImageNet [48,49]; we found that these do not follow the requirements of this study, especially the dataset provided, for example, a dataset of ImageNet supports the recognition of various images such as vehicles, birds, carnivores, mammals, furniture, etc., but it is obviously not yet purposed for a sago palm dataset. However, previous studies explained how a proper dataset helps enhance learning performance; therefore, we applied transfer learning as a strategy to overtake the insufficient data, since it could train network models with a small dataset [40]. For the dataset, our study provided its own dataset captured from UAV images and labeled according to each class. The original dataset for nine classes contained 144 images, while the augmentation process obtained 19 images. The augmentation process consists of rotation, scale (Appendix B), and then 68 test images. In total, the dataset used in this study contains 231 images. Nevertheless, the existing data, for instance, UW RGB-D object dataset, provides 300 general objects in 2.5 datasets [32]. At the testing stage, the RGB images were resized based on the model (Table 2), which was also done by the earlier studies [49,50]. Transfer learning-based CNN models using UAV imagery generate one label for one image patch rather than making a class prediction for every pixel [51]. On the one hand, the presence of overlapping plants, for example, wild sago palm (e.g., lives with other vegetation, irregular shape), could be more challenging in pixel-based classification, for instance, semantic segmentation [40,47,51]. Nevertheless, providing a DL-based segmentation dataset of overlap sago combined with other models is essential for different purposes [32]. At the same time, the selected models used in this study performed detection and recognition successfully, as assessed by earlier studies [23,52]. According to the result of the metric evaluation, the ResNet-50 model outperforms, at around 90%, compared to other networks, which was also depicted by its ROC curve (Figure 12c). Nonetheless, the effects of hyperparameters of each model, such as learning rate, epochs, and minibatch sizes require consideration [53]. Consequently, fine-tuning the parameters of each model should be more noticeable, which is also described in the limitations of this current study.

Since the sago palm has become important in Indonesia, and considering that the potential area for sago in Papua province tends to be declining [5], designing a relevant application using reliable methods or algorithms needs to be considered. In the case of sago palm in our study area, the harvesting time is examined by its morphological appearance, as mentioned earlier, through the flowers. Sago palm forests commonly live together with other undistinguishable plants, but unfortunately, due to the height of the sago and the limitation of visible inspection by human or satellite images, especially in the sago area of Papua that are part of the overall ecology, sago palms are difficult to identify. After investigating other areas in Indonesia, such as South Sulawesi, which is also typically a sago or semi-cultivated forest, [3] found that the complexity of morphologic appearances, such as the similarity of typical plants, affects the results. Therefore, the result of this current study can help the local community, as well as the stakeholders to recognize the harvest time and the species properly, whether it is sago or other plants. To support this, the deployment of this current study by using appropriate fine-tuning or integrating with other frameworks to address a variety of target problems, as mentioned previously, must be considered in our further research.

5. Conclusions

This study compared the capabilities of three models for sago palm recognition based on their dominant appearances, such as leaves fruits, flowers, and trunks. Each model is transferred from pre-trained deep learning networks by substituted base layers. Likewise, the fully connected layer becomes an fc_new, SoftMax layer, and output layer; to obtain our target model, which is nine labels from nine classes, and the probabilities as well. The experiment's result, as shown in Figures 11, 12 and Table 5, ResNet-50 model was taken as a prior model for flowers, leaves, and trunks for sago palm detection. In further research, this baseline model designed is the first in its field and is expected to obtain a high accuracy, including training validation accuracy up to 90%, with less elapsed time and an improved number of epochs, which also provides more datasets of sago palms. Moreover, since the similarity of sago morphology is influenced by the current result, further work must be integrated with different environments and various sago palm datasets.

Author Contributions: Conceptualization, methodology, validation, S.M.A.L., R.C.P., F.R. and D.H.; software, S.M.A.L., F.R., formal analysis, S.M.A.L. and R.C.P.; investigation, S.M.A.L., F.R. and D.H.; resources, D.H.; writing—original draft preparation, S.M.A.L. and R.C.P.; writing—review and editing, S.M.A.L., R.C.P. and D.H.; visualization, S.M.A.L., F.R.; supervision, D.H.; project administration, D.H.; funding acquisition, D.H. All authors have read and agreed to the published version of the manuscript.

Funding: The study was financially funded by the Czech University of Life Sciences Prague through the Internal Grant Agency (IGA) of the Faculty of Engineering 2021 with Grant Number: 31130/1312/3105 Smart sago palm detection using Internet of Things (IoT) and Unmanned Aerial Vehicle (UAV) imagery.

Data Availability Statement: The data presented in this article in the form of figures and tables are a part of ongoing research being undertaken for work of the first author (S.M.A.L.). The field data including boundaries map were collected before the New Form of Papua Regency announced.

Acknowledgments: S.M.A.L. is deeply grateful to the Indonesia Endowment Fund for Education (LPDP-Indonesia) for funding and supporting her study. Additionally, thanks to Yus W from University of Musamus, for sharing his research activities in Tambat Merauke, which were greatly valued. We would like to express our deep appreciation to the head of Plantation Department (Yuli Payunglangi) of the Agency of Food and Horticulture in Merauke. The Head and Staff, also Field Instructor of the Agriculture Office of Mappi Regency of Papua Province for the assistance during the field study. We highly appreciate the local sago famers in our field work, also we are grateful to our anonymous reviewers, editor, whose comments, and suggestions improved the manuscript.

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

Appendix A

UAV and dataset information

Table A1, Figures A1–A3.

Table A1. Technical data of UAV used.

Index	Specification
Dimensions	42.4 × 35.4 × 11 cm
Battery (life and weight)	Li-Ion 7100 mAh 82 Wh; 40 min; 360 g
Video resolution	6K (5472 × 3076)
ISO range	Video-ISO 100-3200 Cr/100-6400 Manual, Photo-ISO100-3200 Car/100-12800 Manual
Camera resolution	20 Mpx; camera chip: 1′ CMOS IMX383 Sony
Maximum flight time	40 min (single charge)
Field of view	82°
Gesture control, Wi-Fi, GPS, controller control, Mobile App, homecoming, anti-collision sensors, automatic propeller stop	Provided
Speeds	72 km/h to 5 km; winds of 62–74 km/h at up to 7000 m above sea level

Figure A1. The drone used in this study: Autel Robotics EVO II Pro 6K.

Appendix B

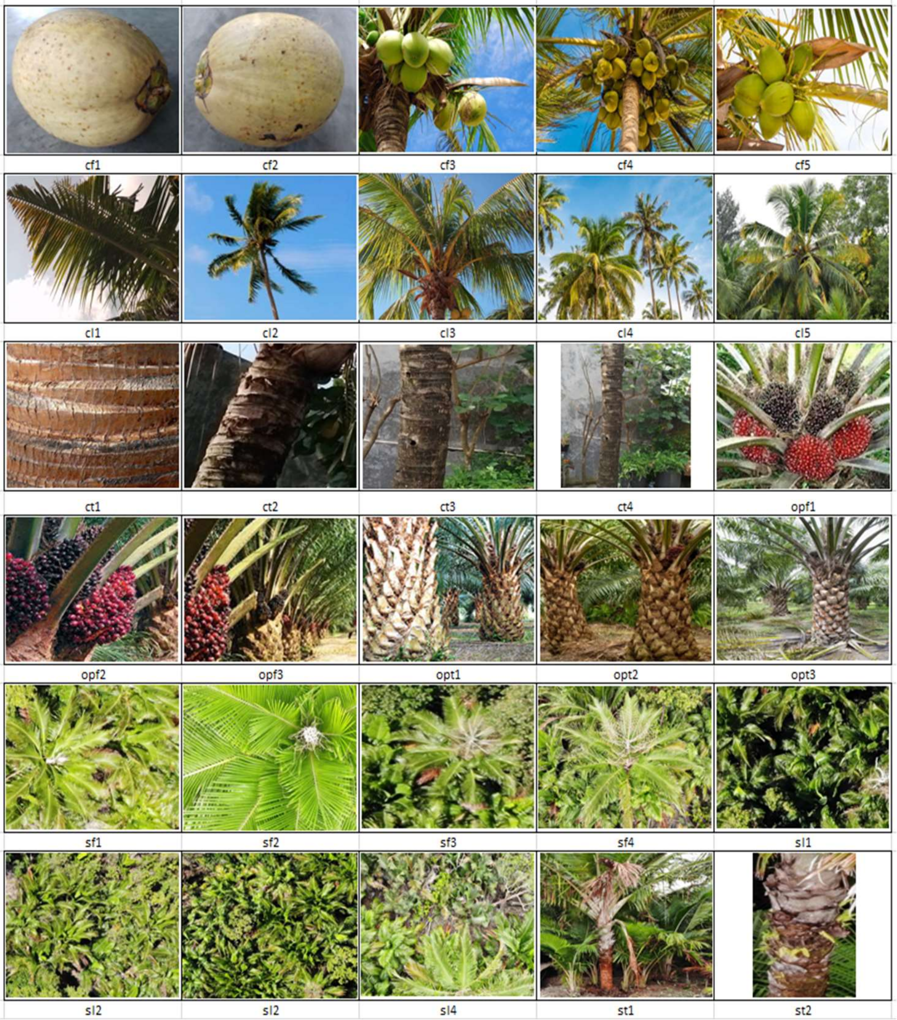


Figure A2. Sample dataset: training data.

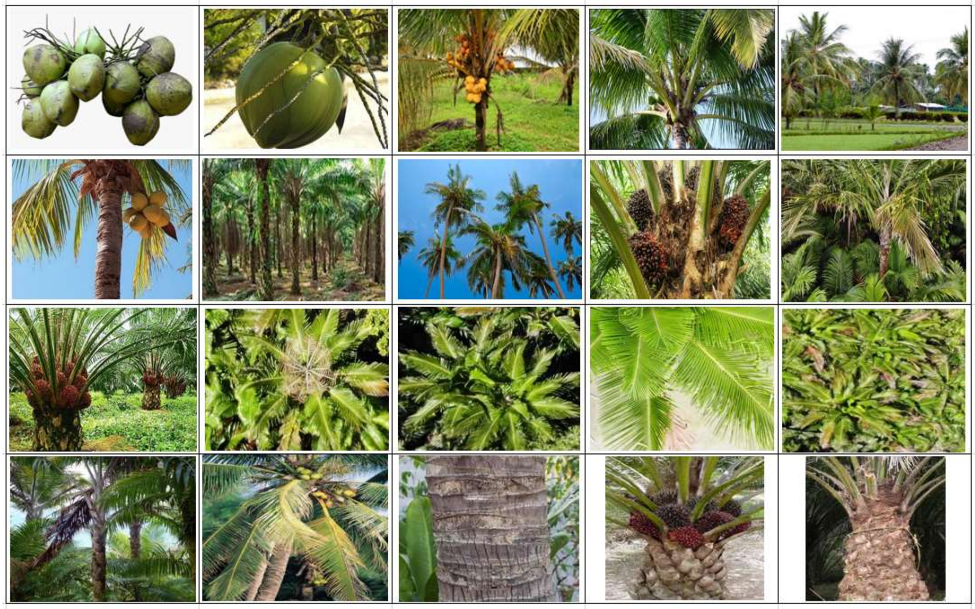


Figure A3. Sample dataset: testing data.

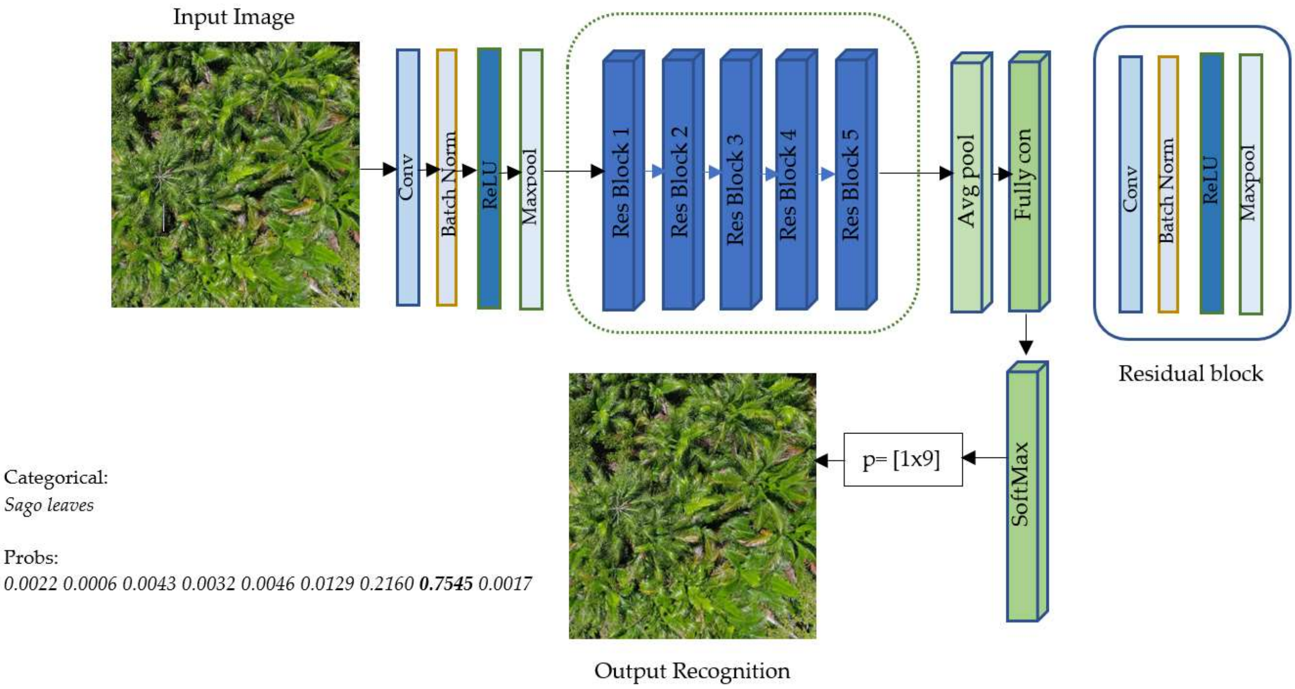


Figure A4. Model used: ResNet-50 network.

References

1. Chua, S.N.D.; Kho, E.P.; Lim, S.F.; Hussain, M.H. Sago Palm (*Metroxylon sagu*) Starch Yield, Influencing Factors and Estimation from Morphological Traits. *Adv. Mater. Process. Technol.* **2021**, 1–23. [CrossRef]
2. Ehara, H.; Toyoda, Y.; Johnson, D.V. (Eds.) *Sago Palm: Multiple Contributions to Food Security and Sustainable Livelihoods*; Springer: Singapore, 2018; ISBN 978-981-10-5268-2.
3. Hidayat, S.; Matsuoka, M.; Baja, S.; Rampisela, D.A. Object-Based Image Analysis for Sago Palm Classification: The Most Important Features from High-Resolution Satellite Imagery. *Remote Sens.* **2018**, *10*, 1319. [CrossRef]
4. Lim, L.W.K.; Chung, H.H.; Hussain, H.; Bujang, K. Sago Palm (*Metroxylon sagu* Rottb.): Now and Beyond. *Pertanika J. Trop. Agric. Sci.* **2019**, *42*, 435–451.
5. Letsoin, S.M.A.; Herak, D.; Rahmawan, F.; Purwestri, R.C. Land Cover Changes from 1990 to 2019 in Papua, Indonesia: Results of the Remote Sensing Imagery. *Sustainability* **2020**, *12*, 6623. [CrossRef]
6. Jonatan, N.J.; Ekayuliana, A.; Dhiputra, I.M.K.; Nugroho, Y.S. The Utilization of *Metroxylon sago* (Rottb.) Dregs for Low Bioethanol as Fuel Households Needs in Papua Province Indonesia. *KLS* **2017**, *3*, 150. [CrossRef]
7. Nanlohy, L.H.; Gafur, M.A. Potensi Pati Sagu Dan Pendapatan Masyarakat Di Kampung Mega Distrik Mega Kabupaten Sorong. *Median* **2020**, *12*, 21. [CrossRef]
8. Pandey, A.; Jain, K. An Intelligent System for Crop Identification and Classification from UAV Images Using Conjugated Dense Convolutional Neural Network. *Comput. Electron. Agric.* **2022**, *192*, 106543. [CrossRef]
9. Tahir, A.; Munawar, H.S.; Akram, J.; Adil, M.; Ali, S.; Kouzani, A.Z.; Mahmud, M.A.P. Automatic Target Detection from Satellite Imagery Using Machine Learning. *Sensors* **2022**, *22*, 1147. [CrossRef] [PubMed]
10. Kentsch, S.; Cabezas, M.; Tomhave, L.; Groß, J.; Burkhard, B.; Lopez Caceres, M.L.; Waki, K.; Diez, Y. Analysis of UAV-Acquired Wetland Orthomosaics Using GIS, Computer Vision, Computational Topology and Deep Learning. *Sensors* **2021**, *21*, 471. [CrossRef] [PubMed]
11. Mhango, J.K.; Grove, I.G.; Hartley, W.; Harris, E.W.; Monaghan, J.M. Applying Colour-Based Feature Extraction and Transfer Learning to Develop a High Throughput Inference System for Potato (*Solanum tuberosum* L.) Stems with Images from Unmanned Aerial Vehicles after Canopy Consolidation. *Precis. Agric.* **2022**, *23*, 643–669. [CrossRef]
12. Niu, Z.; Deng, J.; Zhang, X.; Zhang, J.; Pan, S.; Mu, H. Identifying the Branch of Kiwifruit Based on Unmanned Aerial Vehicle (UAV) Images Using Deep Learning Method. *Sensors* **2021**, *21*, 4442. [CrossRef]
13. Yang, X.; Zhang, Y.; Lv, W.; Wang, D. Image Recognition of Wind Turbine Blade Damage Based on a Deep Learning Model with Transfer Learning and an Ensemble Learning Classifier. *Renew. Energy* **2021**, *163*, 386–397. [CrossRef]
14. Srivastava, A.; Badal, T.; Saxena, P.; Vidyarthi, A.; Singh, R. UAV Surveillance for Violence Detection and Individual Identification. *Autom. Softw. Eng.* **2022**, *29*, 28. [CrossRef]
15. Liu, Y. Transfer Learning Based Multi-Layer Extreme Learning Machine for Probabilistic Wind Power Forecasting. *Appl. Energy* **2022**, *12*, 118729. [CrossRef]

16. Zhao, W.; Yamada, W.; Li, T.; Digman, M.; Runge, T. Augmenting Crop Detection for Precision Agriculture with Deep Visual Transfer Learning—A Case Study of Bale Detection. *Remote Sens.* **2020**, *13*, 23. [CrossRef]
17. 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]
18. Traore, B.B.; Kamsu-Foguem, B.; Tangara, F. Deep Convolution Neural Network for Image Recognition. *Ecol. Inform.* **2018**, *48*, 257–268. [CrossRef]
19. Altuntaş, Y.; Cömert, Z.; Kocamaz, A.F. Identification of Haploid and Diploid Maize Seeds Using Convolutional Neural Networks and a Transfer Learning Approach. *Comput. Electron. Agric.* **2019**, *163*, 104874. [CrossRef]
20. Omara, E.; Mosa, M.; Ismail, N. Emotion Analysis in Arabic Language Applying Transfer Learning. In Proceedings of the 2019 15th International Computer Engineering Conference (ICENCO), Cairo, Egypt, 29–30 December 2019; pp. 204–209.
21. Xiang, Q.; Wang, X.; Li, R.; Zhang, G.; Lai, J.; Hu, Q. Fruit Image Classification Based on MobileNetV2 with Transfer Learning Technique. In Proceedings of the 3rd International Conference on Computer Science and Application Engineering—CSAE 2019, Sanya, China, 22–24 October 2019; ACM Press: New York, NY, USA, 2019; pp. 1–7.
22. Huang, J.; Lu, X.; Chen, L.; Sun, H.; Wang, S.; Fang, G. Accurate Identification of Pine Wood Nematode Disease with a Deep Convolution Neural Network. *Remote Sens.* **2022**, *14*, 913. [CrossRef]
23. Thenmozhi, K.; Reddy, U.S. Crop Pest Classification Based on Deep Convolutional Neural Network and Transfer Learning. *Comput. Electron. Agric.* **2019**, *164*, 104906. [CrossRef]
24. Shaha, M.; Pawar, M. Transfer Learning for Image Classification. In Proceedings of the 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, 29–31 March 2018; pp. 656–660.
25. Gao, C.; Gong, Z.; Ji, X.; Dang, M.; He, Q.; Sun, H.; Guo, W. Estimation of Fusarium Head Blight Severity Based on Transfer Learning. *Agronomy* **2022**, *12*, 1876. [CrossRef]
26. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet Classification with Deep Convolutional Neural Networks. *Commun. ACM* **2017**, *60*, 84–90. [CrossRef]
27. Minowa, Y.; Kubota, Y. Identification of Broad-Leaf Trees Using Deep Learning Based on Field Photographs of Multiple Leaves. *J. For. Res.* **2022**, *27*, 246–254. [CrossRef]
28. Jin, B.; Zhang, C.; Jia, L.; Tang, Q.; Gao, L.; Zhao, G.; Qi, H. Identification of Rice Seed Varieties Based on Near-Infrared Hyperspectral Imaging Technology Combined with Deep Learning. *ACS Omega* **2022**, *7*, 4735–4749. [CrossRef] [PubMed]
29. Jahandad; Sam, S.M.; Kamardin, K.; Amir Sjarif, N.N.; Mohamed, N. Offline Signature Verification Using Deep Learning Convolutional Neural Network (CNN) Architectures GoogLeNet Inception-v1 and Inception-V3. *Procedia Comput. Sci.* **2019**, *161*, 475–483. [CrossRef]
30. Iandola, F.N.; Han, S.; Moskewicz, M.W.; Ashraf, K.; Dally, W.J.; Keutzer, K. SqueezeNet: AlexNet-Level Accuracy with 50x Fewer Parameters and <0.5 MB Model Size. *arXiv* **2016**, arXiv:1602.07360.
31. Alom, Z.; Taha, T.M.; Yakopcic, C.; Westberg, S.; Sidike, P.; Nasrin, M.S. The History Began from AlexNet: A Comprehensive Survey on Deep Learning Approaches. *arXiv* **2018**, arXiv:1803.01164.
32. Minaee, S.; Boykov, Y.Y.; Porikli, F.; Plaza, A.J.; Kehtarnavaz, N.; Terzopoulos, D. Image Segmentation Using Deep Learning: A Survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *44*, 3523–3542. [CrossRef] [PubMed]
33. Yasmeen, U.; Khan, M.A.; Tariq, U.; Khan, J.A.; Yar, M.A.E.; Hanif, C.A.; Mey, S.; Nam, Y. Citrus Diseases Recognition Using Deep Improved Genetic Algorithm. *Comput. Mater. Contin.* **2022**, *71*, 3667–3684. [CrossRef]
34. Zhang, X.; Pan, W.; Xiao, P. In-Vivo Skin Capacitive Image Classification Using AlexNet Convolution Neural Network. In Proceedings of the 2018 IEEE 3rd International Conference on Image, Vision and Computing (ICIVC), Chongqing, China, 27–29 June 2018; pp. 439–443.
35. Sun, P.; Feng, W.; Han, R.; Yan, S.; Wen, Y. Optimizing Network Performance for Distributed DNN Training on GPU Clusters: ImageNet/AlexNet Training in 1.5 Minutes. *arXiv* **2019**, arXiv:1902.06855.
36. Izzo, S.; Prezioso, E.; Giampaolo, F.; Mele, V.; Di Somma, V.; Mei, G. Classification of Urban Functional Zones through Deep Learning. *Neural Comput. Appl.* **2022**, *34*, 6973–6990. [CrossRef]
37. Muhammad, W.; Aramvith, S. Multi-Scale Inception Based Super-Resolution Using Deep Learning Approach. *Electronics* **2019**, *8*, 892. [CrossRef]
38. Sarwinda, D.; Paradisa, R.H.; Bustamam, A.; Anggia, P. Deep Learning in Image Classification Using Residual Network (ResNet) Variants for Detection of Colorectal Cancer. *Procedia Comput. Sci.* **2021**, *179*, 423–431. [CrossRef]
39. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 770–778.
40. Bouguettaya, A.; Zarzour, H.; Kechida, A.; Taberkit, A.M. Deep Learning Techniques to Classify Agricultural Crops through UAV Imagery: A Review. *Neural Comput. Appl.* **2022**, *34*, 9511–9536. [CrossRef] [PubMed]
41. Chirra, S.M.R.; Reza, H. A Survey on Software Cost Estimation Techniques. *JSEA* **2019**, *12*, 226–248. [CrossRef]
42. Hai, V.V.; Nhung, H.L.T.K.; Prokopova, Z.; Silhavy, R.; Silhavy, P. A New Approach to Calibrating Functional Complexity Weight in Software Development Effort Estimation. *Computers* **2022**, *11*, 15. [CrossRef]
43. Zhang, C.; Xia, K.; Feng, H.; Yang, Y.; Du, X. Tree Species Classification Using Deep Learning and RGB Optical Images Obtained by an Unmanned Aerial Vehicle. *J. For. Res.* **2021**, *32*, 1879–1888. [CrossRef]

44. Barmpoutis, P.; Dimitropoulos, K.; Barboutis, I.; Grammalidis, N.; Lefakis, P. Wood Species Recognition through Multidimensional Texture Analysis. *Comput. Electron. Agric.* **2018**, *144*, 241–248. [CrossRef]
45. Huang, P.; Zhao, F.; Zhu, Z.; Zhang, Y.; Li, X.; Wu, Z. Application of Variant Transfer Learning in Wood Recognition. *BioRes* **2021**, *16*, 2557–2569. [CrossRef]
46. Xie, W.; Wei, S.; Zheng, Z.; Jiang, Y.; Yang, D. Recognition of Defective Carrots Based on Deep Learning and Transfer Learning. *Food Bioprocess Technol.* **2021**, *14*, 1361–1374. [CrossRef]
47. Diez, Y.; Kentsch, S.; Fukuda, M.; Caceres, M.L.L.; Moritake, K.; Cabezas, M. Deep Learning in Forestry Using UAV-Acquired RGB Data: A Practical Review. *Remote Sens.* **2021**, *13*, 2837. [CrossRef]
48. Deng, J.; Dong, W.; Socher, R.; Li, L.-J.; Li, K.; Fei-Fei, L. ImageNet: A Large-Scale Hierarchical Image Database. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009.
49. Chakraborty, S.; Mondal, R.; Singh, P.K.; Sarkar, R.; Bhattacharjee, D. Transfer Learning with Fine Tuning for Human Action Recognition from Still Images. *Multimed. Tools Appl.* **2021**, *80*, 20547–20578. [CrossRef]
50. Han, X.; Zhong, Y.; Cao, L.; Zhang, L. Pre-Trained AlexNet Architecture with Pyramid Pooling and Supervision for High Spatial Resolution Remote Sensing Image Scene Classification. *Remote Sens.* **2017**, *9*, 848. [CrossRef]
51. Chew, R.; Rineer, J.; Beach, R.; O'Neil, M.; Ujeneza, N.; Lapidus, D.; Miano, T.; Hegarty-Craver, M.; Polly, J.; Temple, D.S. Deep Neural Networks and Transfer Learning for Food Crop Identification in UAV Images. *Drones* **2020**, *4*, 7. [CrossRef]
52. Cengil, E.; Cinar, A. Multiple Classification of Flower Images Using Transfer Learning. In Proceedings of the 2019 International Artificial Intelligence and Data Processing Symposium (IDAP), Malatya, Turkey, 21–22 September 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 1–6.
53. Feng, J.; Wang, Z.; Zha, M.; Cao, X. Flower Recognition Based on Transfer Learning and Adam Deep Learning Optimization Algorithm. In Proceedings of the 2019 International Conference on Robotics, Intelligent Control and Artificial Intelligence—RICAI 2019, Shanghai, China, 20–22 September 2019; ACM Press: New York, NY, USA, 2019; pp. 598–604.