Morphological characterization of *Liberica* coffee seedlings *via* machine vision

Anton Louise P. de Ocampo[*], Romel U. Briones, Mars G. Panganiban, Melmar D. Eje, Herbert G. Banados Jr., Myrna A. Garcia

Batangas State University The National Engineering University, Batangas, Philippines

ABSTRACT

A crux component in crop health monitoring is the morphological characterization that directly manifests the plant response to the environment and farm inputs like fertilizers and pesticides. A healthy plant presents firm leaves, well-developed flowers, fruits, and a root system. Leaf count, color, and canopy coverage can help characterize general crop health. In this study, an aerial drone is allowed to traverse an optimal path plan while positioned near an area of interest where several overhead images of crops are captured. The total collected aerial images count to 1,918 which is divided into training and testing sets with a 70:30 distribution ratio. From these aerial images, coffee seedlings are detected from the background, weeds, and other crops in the field using the VGG-16 model trained to recognize coffee seedlings. Once localized, leaf counting is performed using segmentation, while the canopy coverage estimation uses a patch-based DNN model to calculate the relative coverage concerning the overhead leaf area. Since leaf color extracted from an RGB image is very much affected by ambient light, normalization using an enhanced Triangular Greenness Index (eTGI) is implemented. The estimation results of the system reached up to 91.48% accuracy, 92.52% precision, and 93.82% recall for detecting coffee seedlings while the Mean Absolute Percentage Error (MAPE) for leaf count and canopy coverage of 11.61% and 15.67% respectively. For future work, the leaf color can be correlated to chlorophyll and percent nitrogen measurements which will require specialized instruments for validation. Estimation of chlorophyll and percent nitrogen is vital in identifying the amount and type of fertilizers to be applied.

Keywords: aerial crop monitoring, smart farming, enhanced triangular greenness index, canopy coverage

1. Introduction

The integration of emerging technologies such as machine learning, the Internet of Things (IoT), and advanced remote sensing techniques has revolutionized precision agriculture. Remote sensing plays a pivotal role in precision agriculture by providing high-resolution data for crop health monitoring and land use assessment. For instance, hyperspectral imaging has demonstrated its capability to assess crop conditions and detect diseases [1,2]. Recent studies have emphasized the role of multispectral and thermal imagery in estimating evapotranspiration and soil moisture levels, critical for irrigation management [3,4]. The advancements in satellite-based remote sensing, including sensors like Sentinel-2, enable large-scale monitoring, providing cost-effective solutions for farmers [5].

Machine learning (ML) is increasingly employed to analyze complex agricultural datasets. Techniques such as convolutional neural networks (CNNs) have shown high accuracy in crop classification and disease detection using remote sensing data [6,7]. Studies have also explored the potential of ML algorithms for yield prediction, leveraging both environmental and historical yield data [8]. Integrating ML with remote sensing can significantly enhance predictive capabilities, enabling real-time decision-making [9].

The IoT paradigm has facilitated the development of smart farming systems, where interconnected devices collect real-time data on soil, weather, and crop conditions. This data-driven approach optimizes resource utilization and improves productivity. For example, automated irrigation systems leveraging IoT sensors have demonstrated substantial water savings while maintaining crop yield [10]. IoT-based monitoring systems have also proven effective in pest management and disease control [11,12].

Advancements in agricultural technologies have significantly contributed to sustainability by reducing resource wastage and environmental degradation. Remote sensing combined with ML has been pivotal in mapping and mitigating the impacts of climate change on agriculture [13]. Furthermore, studies have highlighted the role of these technologies in promoting carbon sequestration and reducing greenhouse gas emissions through optimized farming practices [14,15].

Despite all these developments, challenges about data heterogeneity, lack of standardization, and high implementation costs arise and persist. These challenges are likely to be solved through interdisciplinary collaboration and the development of user-friendly tools. Future research should focus on the integration of blockchain for data security and the use of edge computing in real-time analysis, especially within resource-constrained environments [16,17].

To address these problems, this study aims to develop a phenotyping system for aerial images of coffee trees using machine vision. An unmanned aerial vehicle (UAV) is used for data acquisition. UAV is one of the emerging instruments used in precision agriculture, particularly data acquisition for

*Corresponding author

Email address: antonlouise.deocampo@g.batstate-u.edu.ph

plant phenotyping [18]. The phenotyping system is implemented on a personal computer or laptop to extract the morphological attributes of leaf count, level of greenness, and canopy coverage of the coffee trees. To characterize and extract the morphological attributes, the dataset gathered is fed to the machine vision algorithms. Machine learning algorithms and image processing techniques are implied and primarily associated with precision agriculture application techniques for developing innovative plant phenotyping tools [19]. Further, the study intends to use advanced methods instead of manual field measurements to improve coffee crop management and decision-making.

2. Materials and methods

This part presents the methods used by the proponents to achieve the research objectives and standards. In the first section, a description of how the data collection was performed, and some practical details about the data collection are presented. The second section presents the practical details about the detector, selected methods, and their parameters. In the final section, the evaluation methods are presented.

2.1. Data acquisition

A dataset has been created by collecting aerial images and videos of coffee trees. The researchers started to collect data in September 2021 to April 2022. The data gathering was held in the Pueblo Farm in Rosario, Batangas, which has a field area of 5000 sq. meters. Aerial images of coffee trees were captured with a single-camera drone wherein the flying height was 5 meters above the field. This UAV is equipped with 2375 mAh intelligent flight battery capacity and a camera system consisting of a sensor (1/2.3" CMOS), 3-axis image stabilizer, 4K Ultra HD: 3840×2160 24/25/30p, with a lens of 85° field of view (FOV), 35 mm format equivalent: 24 mm and an effective pixel of 12 MP. Figure 1 shows the illustration of the flight height and how the researchers captured the coffee trees using UAV.
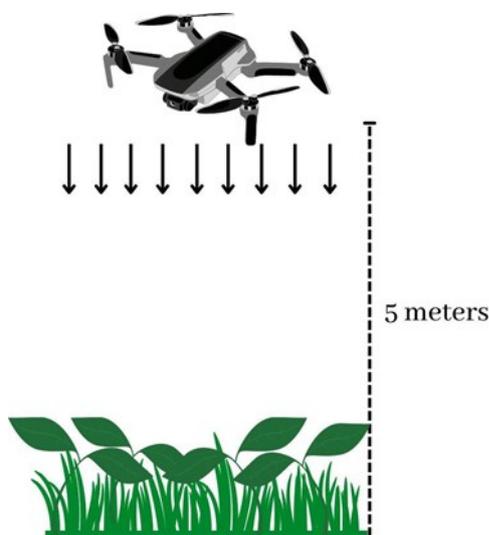

Figure 1. Flying height of the drone from the ground.

The researchers selected usable images from the photos taken using the control app for the UAV used. The research data consisted of a total of 1,918 images. To train and evaluate the model to be used in object detection, different sets for training and testing were needed. In line with this, the dataset was divided into two sets: training and testing with a 70:30 ratio, where 1343 images are used for training, and the remaining 575 are used for testing. To achieve a balanced dataset, the photos of coffee tree seedlings are matched with background noises of leaves, grasses, and the ground in a 50-50 ratio. Figure 2 shows that images with coffee trees are positive (a), and images that only contain background noises are negative (b).
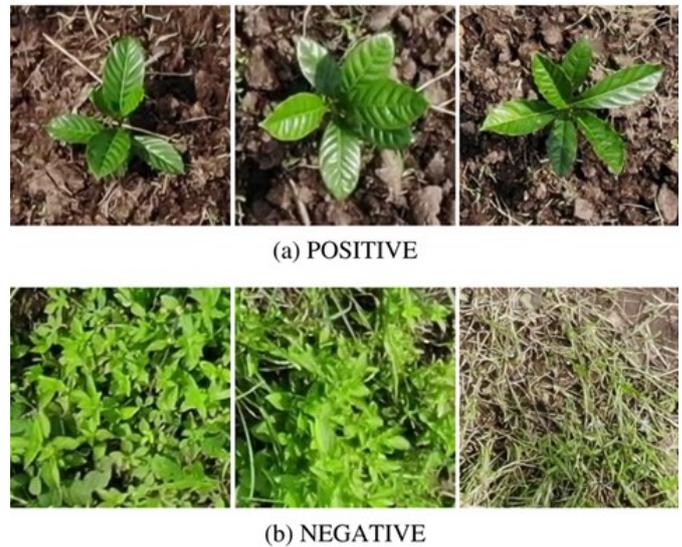

Figure 2. Sample images showing the presence of coffee trees: (a) positive and (b) negative.

2.2. Ground truthing

The researchers conducted manual field measurements on the coffee trees for ground-truthing. Five hundred crops were tagged to distinguish the crops from each other. The field measurements from these crops are used as ground-truth data and were compared to the system measurements to evaluate the accuracy of the predicted values of the system for leaf count and canopy coverage.

For the leaf count, the actual values are obtained by manually counting the number of leaves of all 500 crops in actual plants and in the photos taken.

To measure the canopy coverage of the coffee crops, the Leaf Area Index (LAI) was obtained. Strong relationships were found between canopy coverage and LAI of crop species that followed the exponential rise to a maximum form [7]. Thus, LAI can be used as a ground truth for evaluating the canopy coverage. Aerial images were used for the estimation of the canopy of the system. Thus, predicted values cannot be directly compared to the actual measurements of the leaf area. With this, the researchers used image processing methods to derive the area of leaves in terms of square centimeters per pixel by doing the following steps. First, the coffee crops were separated from the background. For the segmentation of the coffee crops, the Image Segmenter app in MATLAB® was used. Specifically, the Graph Cut feature from the Image Segmenter app was utilized. Next, the segmented image was converted into a binary image using an auto-generated function (Figure 3). The code creates the mask with the name sliderBW.
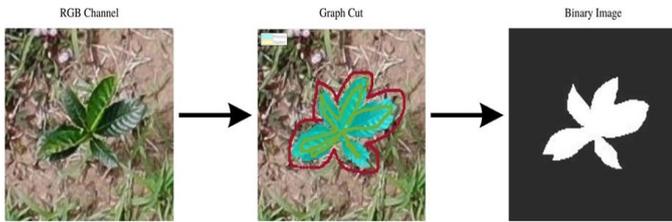
Figure 3. Workflow diagram of a masked image using RGB image.

Afterward, the researchers measured the length in pixels of a leaf in each crop. This was measured using the imdistline tool in MATLAB®. Further, the number of true pixels of the same leaf was also obtained. The actual length of the leaf was divided by the measured length in pixels. The result was then squared. This gives the squared centimeters per pixel at 5m altitude. Lastly, the result from the previous step was multiplied by the number of true pixels of the leaf. This process was repeated for every coffee crop.

2.3. Object detection

The researchers described methods to locate plants in UAV imagery. The main task for this part is to implement a deep-learning object detector to localize or detect the coffee crop in an image.

2.4. VGG-16 architecture

VGG is a convolutional neural network that employs very low convolution filters (3x3) and increases the network's depth to 16 layers. Figure 4 shows the architecture of VGG-16. The network receives images with dimensions of 224x224x3 as input. The first and second layers have 64 channels with 3x3 filter sizes and the same padding. Following a stride max pool layer, two convolution layers with 256 filter size and filter size are added. This is followed by the identical stride max pooling layer as the previous layer. Then, there are two convolution layers with 3x3 and 256 filter sizes. There are then two sets of three convolution layers and a max pool layer. Each has 512 identical filters with the same padding. This image is then sent into a convolution layer stack of two. This convolution and max pooling layer's filters are 3x3. It also uses 1x11 pixels in some layers to change the number of input channels. To prevent the spatial characteristic of the image, a 1-pixel padding, also known as identical padding, is applied after each convolution layer. By layering convolution and max-pooling layers, the researchers were able to create a 7x7x512 feature map. This output is flattened to produce a feature vector with the index (1, 25088). In addition, three fully linked layers are present: the first layer takes input from the last feature vector and generates a (1, 4096) vector, the second layer outputs a (1, 4096) vector, and the third layer outputs the channels for the input classes. Finally, the output of the third fully connected layer is passed to the softmax layer to normalize the classification vector.
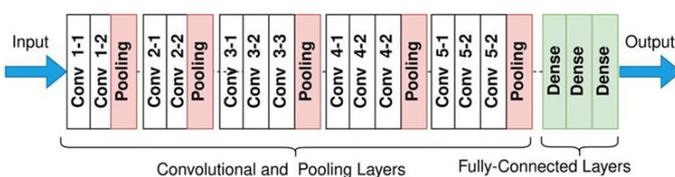

Figure 4. VGG-16 architecture.

2.5. Model training

The aerial images obtained during data gathering were enhanced to improve the image representation (Figure 5). It includes tasks conducted to manipulate digital images with the intention of quality optimization, noise reduction, or resolving lighting issues. The contour matching technique was used to tackle detection issues caused by the varying contours of the item and the complicated scene. Further, the images undergo segmentation to annotate the images pixel by pixel. This gave a more detailed model. Following the segmentation process, the features that contain the necessary information of the image input were extracted. The extracted features aided in extracting relevant data from the image input.
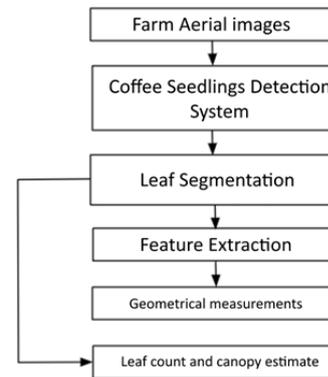

Figure 5. Flow chart of detection-based leaf counting and canopy estimation.

After finishing all the steps mentioned above, the training process was initialized. The training sets containing 1,343 images were used for model optimization. The model was trained for 25 batch sizes and 50 epochs for 5 days. Other parameters include a warm-up period of 1000 iterations, a penalty threshold 0.5, L2 regularization of 0.0005, and a learning rate 0.001. The neural network models were trained with Keras using the TensorFlow backend. The experimentation and training of the model were performed using a virtual machine with Ubuntu (64-bit) operating system. The hardware components of the PC included 5133MB Base Memory/RAM and 4 Processors.

2.6. Detection-based leaf counting

The researchers used an Application Programming Interface (API) to call out functions to the TensorFlow API model written in Python programming language.

The VGGnet model with 16 layers (VGG-16) was the network used for leaf counting. The network is taught to identify the leaves present in a coffee tree, using image labeling provided in training, and only then count them. To reduce the training time and improve the robustness, the researchers adapted a pre-trained model instead of training the entire network from scratch. The researchers re-trained the model on their data. It got the pre-trained configurations of the model from the database of leaf images from Mendeley Data, which contains more than 4500 images.

2.7. Estimation of canopy coverage

API was also used to estimate the canopy coverage. A pre-trained VGG-16 patch-based CNN model was used. This method extracts and calculates the presence of canopy regions

in an image. The VGG-16 model was re-trained on the researchers' dataset using model weights from Mendeley Data's leaf image database.

2.8. Level of greenness

The level of greenness of the coffee tree was identified using NDVI (Figure 6). NumPy Library from Python is used in the process. First, the color bands are identified after the converting the image into an NIR image. Color bands such as red(R), blue(B), and green(G) are classified. The RGB bands are plotted in a color histogram, which compromises the intensity of the colors and the wavelength of colors present in the detected crop. To get the level of greenness of the coffee tree, the green color bands are averaged using NumPy.
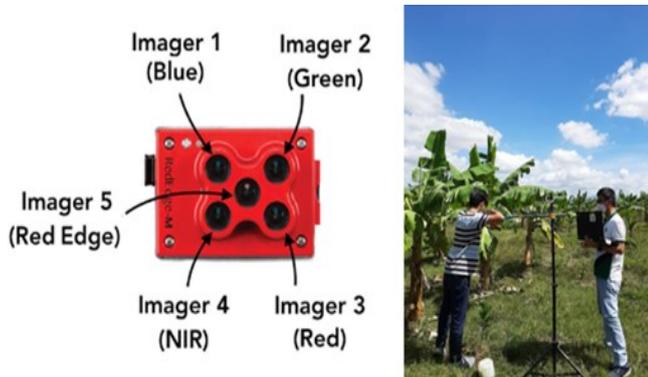

Figure 6. NDVI and RGB image acquisition.

2.9. Phenotyping system workflow

Figure 7 shows the workflow diagram for the phenotyping system. Using the --image or --i argument, the aerial images are loaded and read by the system. The coffee trees present in the input image were detected. Then, the detected crops will be converted to NIR images. To do so, the input image will be first converted into a grayscale image. Then, color channels will be created including, red(R), blue (B), and green(G) channels. Weights are set for each channel whereas R = 0.642, G = 0.532, and B = 0.44.
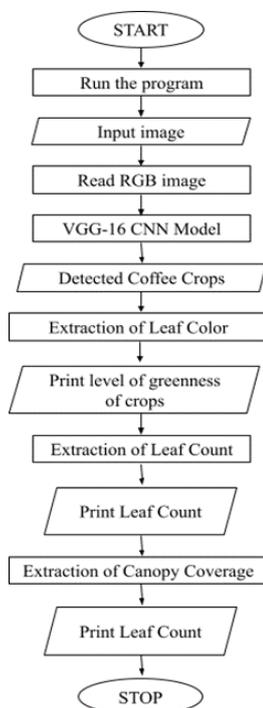

Figure 7. Workflow diagram for the phenotyping system.

The sum of the weights is used to normalize the channels. Following that, the channels will be combined to produce a NIR image. The scale ratio is 255/max to increase the fully dynamic range. The result of the conversion will then be saved to the disk with the filename "image_nir.png." Following the conversion of the RGB image to the NIR image, the system will now have the generated NIR and RGB images. The morphological attributes of the detected crops, including leaf count, level of greenness, and canopy coverage will be extracted using the processes discussed from detection-based counting to the level of greenness. The system's predicted values of the morphological attributes will be printed to the terminal.

2.10. Evaluation of the phenotyping system

In this phase, the researchers tested the performance and functionality of the system. For object detection, the model was tested on a previously unseen dataset, which is the remaining 30% of the dataset. The test data must not be used in training to avoid overfitting. The researchers used Confusion Matrix to calculate the performance accuracy of the system. The actual values (ground-truth) were compared with the predicted values of the model. The categories used for each class were derived: (1) true positive (TP) - an outcome where the model correctly predicts the object class training, (2) false positive (FP) - nothing is being detected when no object must be detected, (3) true negative (TN) - an outcome where the model incorrectly predicts the positive class. There is a detection even though no object must be detected, and (4) false negative (FN) failed to detect an object that is required to be detected. Further, the value of the predictions of the model can be determined by the following metrics:

Accuracy. It is calculated as the percentage of correctly classified predictions to all predictions.

$$Accuracy = \frac{T_p + T_n}{T_p + T_n + F_p + F_n} \qquad (1)$$

Precision. It is the ratio of correctly positive predictions to the total predicted positive predictions.

$$Precision = \frac{T_p}{T_p + F_p} \qquad (2)$$

Recall. It is the ratio of correctly positive predictions to the total predictions in the actual class.

$$Recall = \frac{T_p}{T_p + F_n} \qquad (3)$$

F1 Score. The weighted average of precision and recall is used to calculate the F1 Score.

$$F1\ Score = 2 * \frac{Precision * Recall}{Precision + Recall} \qquad (4)$$

For the leaf count and canopy coverage, the previously unseen images were fed to the system and the predicted values were recorded. The predicted values were then compared to the actual values or measurements taken by the researchers. To evaluate the accuracy of the predicted values, the following metrics were used:

Mean Absolute Percentage Error (MAPE). MAPE calculates the average of percentage errors. It is a metric for prediction accuracy. Table 1 shows the MAPE Interpretation.

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{Yi - Y\hat{}i}{Yi}\right| x100 \qquad (5)$$

Table 1. MAPE interpretation.

MAPE	Interpretation
<10	Highly Accurate Forecasting
10-20	Good Forecasting
20-50	Reasonable Forecasting
>50	Inaccurate Forecasting

Mean Square Error (MSE). The mean squared error between the predicted and actual value is measured by MSE.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Yi - Yi\hat{})^2 \qquad (6)$$

3. Results and discussion

The following section provides the features dataset collected for training the detection model, followed by the details of the ground truth measurements that validate the generated results of the system. Then, the results of model training, detection, and phenotyping are provided.

3.1. Dataset

Table 2 presents the date and time taken sequence for image acquisition, which started on September 24, 2021, when 72 images were acquired. For the next day, seven videos were gathered on November 17, 2021, which started at 3:41 in the afternoon and ended at 4:53 pm. The seven videos were processed using extracting frames in MATLAB and resulted in 1,103 images with 1 coffee seedling per image. On March 3, 2022, the third day of data gathering happened, and 168 images were acquired. Moreover, the fourth day was on March 10, 2022, which started at 11:50 in the morning and ended at 12:34 in the afternoon, 190 images were gathered. The fifth and sixth days of data gathering where on March 17 and March 26, 2022, where 130 images (5th day) and 172 images (6th day) were collected. Lastly, the last day for data gathering was April 7, 2022 which 83 images were attained. The data gathering resulted in 1,918 images, each image having 1 to 3 coffee seedlings. Images were saved as .jpg files with filenames that incorporate tray numbers (DJI_0001, DJI_0002, DJI_000,3, and so on).

Table 2. Date and time taken sequence for dataset acquisition.

Date Taken mm/dd/yy	Start Time	End Time	No. of Images
09/24/2021	2:28 PM	2:53 PM	72
11/17/2021	3:41 PM	4:53 PM	1103
03/03/2022	1:50 PM	2:50 PM	168
03/10/2022	11:50 AM	12:34 PM	190
03/17/2022	11:15 AM	1:49 PM	130
03/26/2022	12:12 PM	12:55 PM	172
04/07/2022	12:14 PM	12:44 PM	83
		Total:	1918

Aerial images were collected on different days for the phenotyping system to achieve the highest possible accuracy. Figure 8 shows the (a) overhead shots of the field, and (b) sample images for the dataset taken by UAV. Table 2 is also presents the sequence of the date and time taken for image acquisition.

Figure 8. (a) Overhead shots of the field, and (b) sample image for dataset taken by UAV.

3.2. Ground truth

Actual measurements of the morphological attributes (number of leaves, color of leaves, and canopy coverage estimation) of the coffee seedlings were obtained. Figure 9 shows the relationship between the actual measurements (taken manually in the field) and derived measurements (taken using image processing methods) that were used to generate the ground-truth of true values. The researchers achieved a linear graph between actual measurements and derived measurements, which means that these two almost had the same values or matched each others measurements. MAPE was calculated, resulting in 19.9610, which means that it was a good forecast. Thus, derived measurements were acceptable to become the true values for ground truthing in terms of canopy coverage estimation.
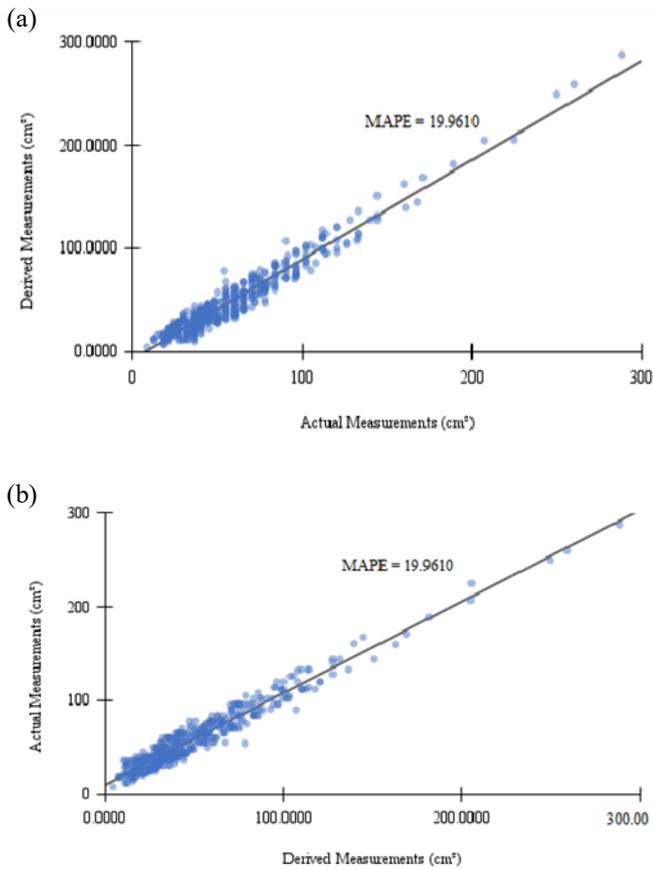
(a)

(b)

Figure 9. (a) Derived measurements (cm²) vs. actual measurements (cm²), (b) actual measurements (cm²) vs. derived measurements (cm²).

3.3. Model training

Table 3 shows the summary of the VGG-16 architecture. The dimensions of the input picture, which is the RGB image with depth 3, change to 224x224x3 as it passes through the first and second convolutional layers. It has 3x3 convolution layers with stride 1 and a maxpool layer with 2x2 stride 2 filters. The VGG-16 has 16 layers with a total of 25,088 features after flattening the last convolutional layer and 1000 nodes in the final layer because the VGG-16 primarily learned 1000-class classification tasks.

Table 4 contains all the data acquired after 20 epochs for training loss, validation loss, accuracy, and F1-score. The training progress graph for this model is shown in Figure 10, which summarizes the data in Table 3. Figure 10 (a) shows that as the epoch increases, the training loss decreases while the validation loss increases. This graph indicates that the system has attained the possible high accuracy of the model. Figure 10 (b) shows all the data acquired after 20 epochs for accuracy. In comparison, Figure 10 (c) shows all the data acquired after 20 epochs for the F1-score. The accuracy of the model in the training set is 96.21%, and the F1-Score is 91.21% after 20 epochs.

Table 3. VGG-16 model summary

Layer (type)	Output Shape	Param #
input 1 (Input Layer)	[(None, 224, 224, 3)]	0
conv1 1 (Conv2D)	(None, 224, 224, 64)	1792
conv1_2 (Conv2D)	(None, 224, 224, 64)	36928
pool1 1 (MaxPooling2D)	(None, 112, 112, 64)	0
conv2 1 (Conv2D)	(None, 112, 112, 128)	73856
conv2_2 (Conv2D)	(None, 112, 112, 128)	147584
pool2 1 (MaxPooling2D)	(None, 56, 56, 128)	0
conv3 1 (Conv2D)	(None, 56, 56, 256)	295168
conv3_2 (Conv2D)	(None, 56, 56, 256)	590080
conv3 3 (Conv2D)	(None, 56, 56, 256)	590080
pool3 1 (MaxPooling2D)	(None, 28, 28, 256)	0
conv4_1 (Conv2D)	(None, 28, 28, 512)	1180160
conv4 2 (Conv2D)	(None, 28, 28, 512)	2359808
conv4 3 (Conv2D)	(None, 28, 28, 512)	2359808
pool4 1 (MaxPooling2D)	(None, 14, 14, 512)	0
conv5 1 (Conv2D)	(None, 14, 14, 512)	2359808
conv5 2 (Conv2D)	(None, 14, 14, 512)	2359808
conv5_3 (Conv2D)	(None, 14, 14, 512)	2359808
pool5 1 (MaxPooling2D)	(None, 7, 7, 512)	0
flatten (Flatten)	(None, 25088)	0
fc 1 (Dense)	(None, 4096)	27336989
fc 2 (Dense)	(None, 4096)	8784825
predictions (Dense)	(None, 1000)	1474944
Total params		52,311,446
Trainable Params		52,311,446
Non-trainable Params		0

Table 4. Accuracy of the model training.

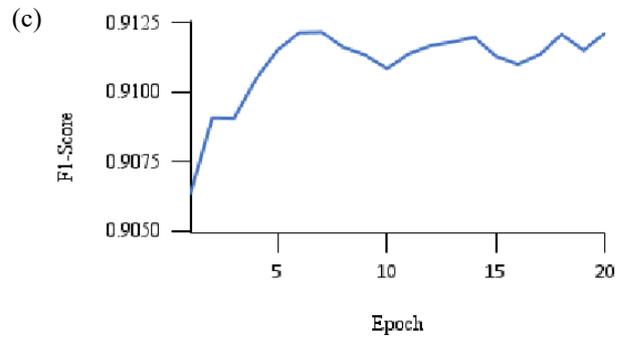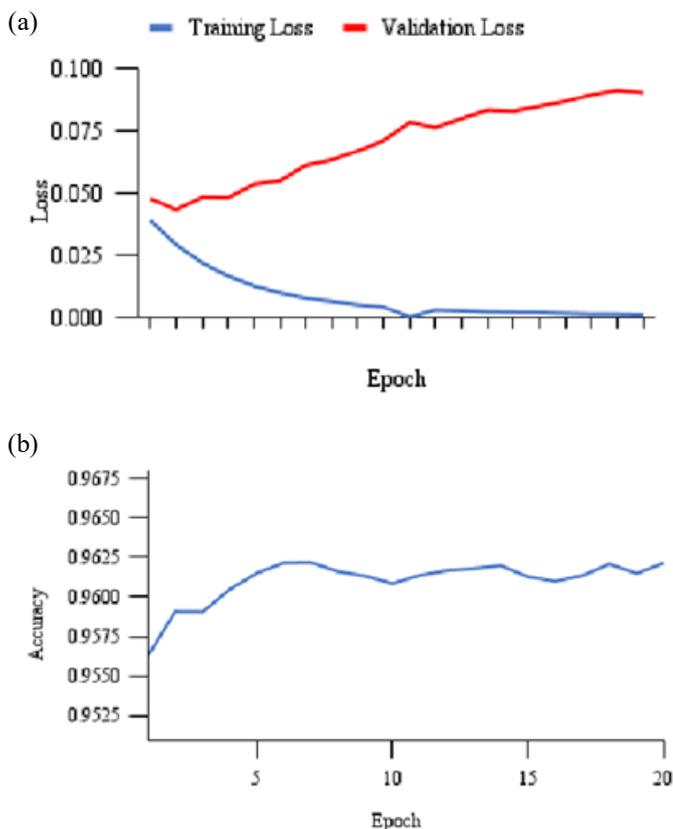| Epoch	Training Loss	Validation Loss	Accuracy	F1-Score
1	0.0391	0.0479	0.9563	0.9603
2	0.0296	0.0434	0.9591	0.9091
3	0.0221	0.0485	0.9591	0.9091
4	0.0169	0.0483	0.9605	0.9105
5	0.0130	0.0537	0.9615	0.9115
6	0.0103	0.0551	0.9621	0.9121
7	0.0082	0.0612	0.9622	0.9122
8	0.0069	0.0633	0.9616	0.9116
9	0.0055	0.0670	0.9613	0.9113
10	0.0046	0.0712	0.9609	0.9109
11	0.0004	0.0783	0.9614	0.9114
12	0.0033	0.0764	0.9617	0.9117
13	0.0030	0.0799	0.9618	0.9118
14	0.0025	0.0832	0.9620	0.9120
15	0.0023	0.0828	0.9613	0.9113
16	0.0022	0.0848	0.9610	0.9110
17	0.0019	0.0867	0.9614	0.9114
18	0.0016	0.0894	0.9621	0.9121
19	0.0016	0.0912	0.9615	0.9115
20	0.0014	0.0903	0.9621	0.1921

(a)

(b)

(c)

Figure 10. (a) Training and validation loss, (b) accuracy, and (c) f1-score.

3.4. Detection-based Leaf Counting

In Figure 11, the TensorFlow API is accessed via the web. The identified images of the coffee trees were uploaded to the web server and loaded to the re-trained VGG-16 model. The model processed the image and counted the number of leaves on the coffee tree. The predicted values were then sent to the API web server again, then back to the phenotyping system to display the results of the leaf count in the output terminal. Table 5 presents the tally of the predicted values of leaf count acquired from 500 test sets for the extraction of the morphological features fed on the phenotyping system.
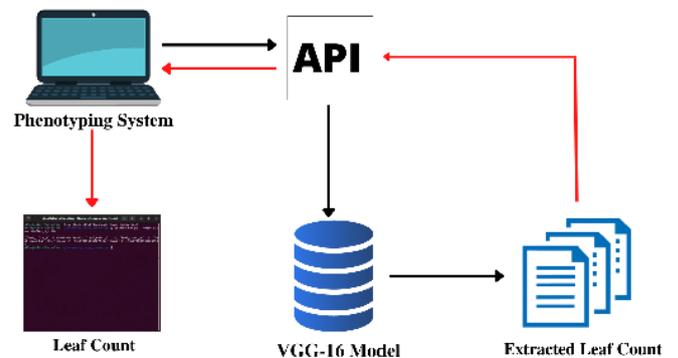

Figure 11. Extraction of leaf count

Table 5. Leaf count distribution table of the 500 test sets.

No. of Leaves	Frequency
1 - 4	1
5 - 8	10
9 - 12	179
13 - 16	163
17 - 20	122
21+	3
Total	**478**

3.5. Estimation of canopy coverage

Figure 12 also used API, but a different pre-trained model was used to estimate canopy coverage. As with leaf counting, the TensorFlow API is accessed via the web. The identified images were saved to a web server before being fed into the re-trained VGG-16 patch-based CNN model. The image was processed by the model, which calculated the

coffee tree's canopy coverage. The predicted values were then sent back to the API web server before being sent back to the phenotyping system to display the canopy coverage results in the output terminal. Table 6 summarizes the predicted canopy coverage values obtained from the phenotyping system.
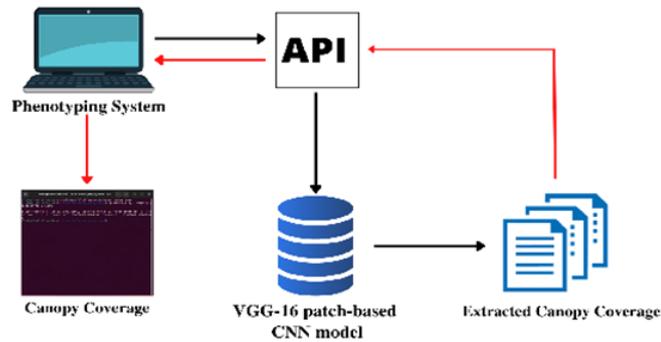

Figure 12. Extraction of canopy coverage.

Table 6. Canopy coverage distribution table of the 500 test sets.

Canopy Coverage (cm²)	Frequency
1 - 20	64
21 - 40	169
41 - 60	110
61 - 80	64
81 - 100	42
101+	29
Total	**478**

3.6. Level of greenness

In Figure 13, the RGB image was converted to an NIR image. After that, the color bands (RGB) were identified, producing the NDVI image. The RGB bands were then plotted in a color histogram, which consisted of the intensity and wavelength of colors detected on crops. Through this, the leaf color was extracted by averaging the green color bands from the color histogram. The proponents used a scale of 1 to 10 to represent the level of greenness of the coffee tree. Table 7 represents the leaf color intensity scaling from 0 to 10 and also the frequency of each. As the result of the graph provided, most coffee crops have a leaf color intensity 10.
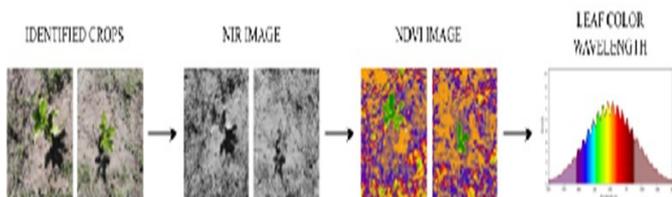

Figure 13. Extraction of leaf color intensity.

Table 7. Leaf color intensity distribution table of the 500 test sets.

Leaf Color Intensity	Frequency
0	7
1	7
2	4
3	7
4	5
5	3
6	6
7	4
8	9
9	1
10	447
Total	**500**

3.7. Phenotyping system program execution

The researchers developed a phenotyping system of extracting morphological attributes from aerial images of coffee trees in Python. To run the program, a command terminal was utilized and the run.py script was responsible for loading and saving images. Figure 14 presents the phenotyping system, demonstrating the object detection and extraction of the morphological attributes of coffee tree aerial images. Figure 14(a) shows the object detection of coffee trees present in the input image.

First, the image is loaded into the command terminal. The aerial image was read, and then the coffee crops were detected through the VGG-16 model.

In Figure 14(b), extraction of morphological attributes took place. From the crops recognized by object detection, morphological attributes were extracted. For the leaf color, identified crops were pre-processed by the system to produce grayscale and contour-matching images. The RGB image is then converted to an NIR image, yielding the NDVI image. The green color bands were then averaged resulting in the extracted level of greenness of the coffee tree. For the canopy coverage and the leaf count, the system used an API and pre-trained models. VGG-16 patch-based CNN model was used for canopy coverage, while VGGnet model with 16 layers was used for leaf count. Identified crops of coffee tree images were processed through these models to extract the morphological attributes. As a result, the predicted values of the estimation of canopy coverage, number of leaves, and level of greenness were displayed in the output terminal of the system.
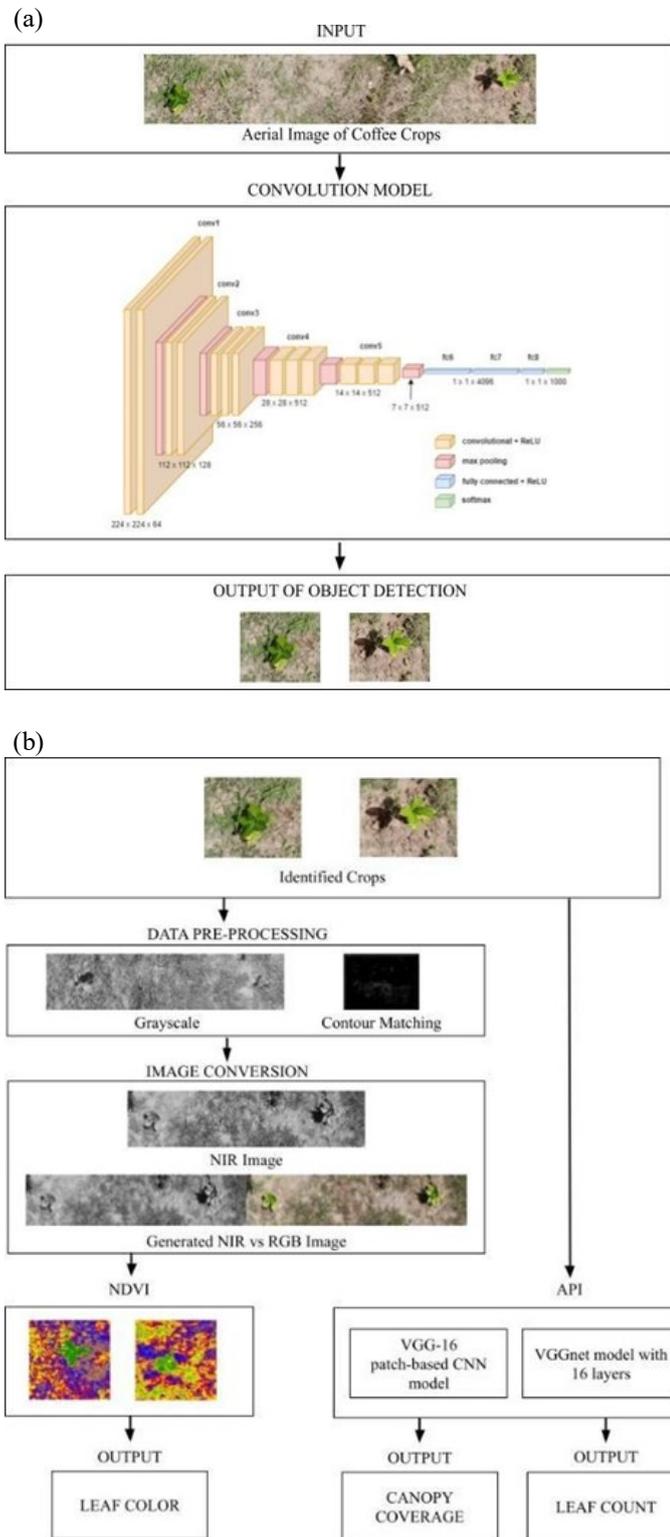
(a)

(b)

Figure 14. Phenotyping system showing the (a) object detection of coffee trees and (b) extraction of morphological attributes.

3.8. Testing of the phenotyping system

About 500 previously unseen coffee crops were fed to the system for testing. Out of the 500 coffee crops, the following were the results of the two sample test images of the phenotyping system. Table 8 demonstrates the corresponding output for each input of the sample test images. The filename of the sample test image was entered into a common terminal as the input of the phenotyping system. For each input, there are two sets of outputs. The first set of outputs was the results for the leaf count, level of greenness, and canopy coverage for each detected crop. These results were displayed in the

terminal. Meanwhile, the second set of output was the images generated by the phenotyping system, which was saved to disk.

Table 8. Results of the two sample test images of the phenotyping system. Results of the two sample test images of the phenotyping system.

3.9. Evaluation of the performance of machine vision in phenotyping

Table 9 shows the values obtained for the evaluation metrics of object detection through the confusion matrix. The accuracy acquired was 0.91478, precision was 0.925208, recall was 0.9382, and F1-score was 0.93166, all high values indicating that the model used in object detection was good and accurate.

Table 9. Evaluation results for the object detection.

Accuracy	Precision	Recall	F1-Score
0.91478	0.925208	0.9382	0.93166

Figure 15 shows that the predicted and true values for the leaf count overlap, indicating that the values obtained were relatively close. Based on the MAPE interpretation metrics, the calculated MAPE for the leaf count was 11.6093, indicating that it is a great prediction.
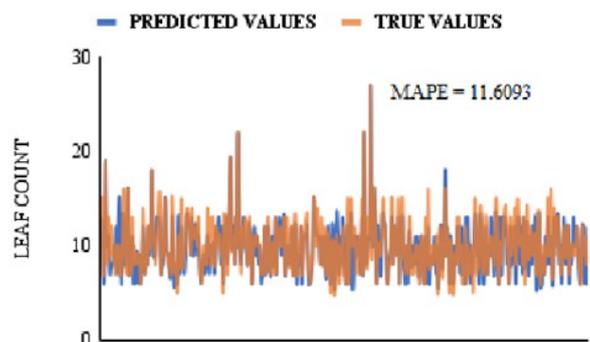

Figure 15. Predicted values vs true values.

Figure 16 shows the canopy coverage between the true and predicted values of the coffee crops. It demonstrates that the majority of the values obtained manually and through the system have a high linear correlation.

Figure 15. Predicted values vs true values.

Figure 16 shows the canopy coverage between the true and predicted values of the coffee crops. It demonstrates that the majority of the values obtained manually and through the system have a high linear correlation.
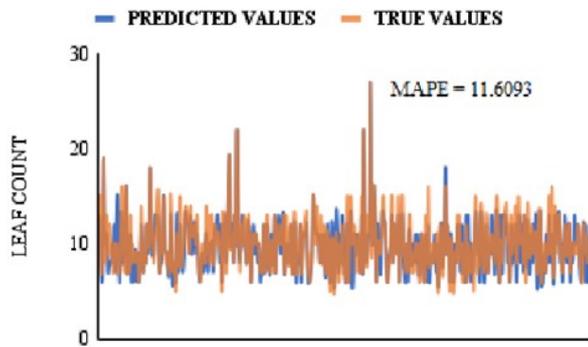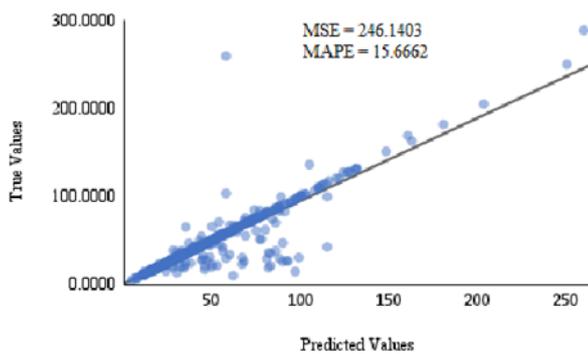

Figure 16. Data evaluation of canopy coverage of coffee crops (predicted values vs true values).

The MSE and MAPE between the true and predicted canopy coverage values were also calculated; the MSE and MAPE are 246.1403 and 15.6662, respectively. These values are low errors indicating good prediction accuracy.

However, MSE, or Mean Squared Error, is not a good indicator of accuracy for the system since it is affected by the standard error, is based on the square of error, and is scale-dependent, whereas MAPE is not since it can be computed with respect to data that are guaranteed to be strictly positive. That is why MAPE is a more precise metric.

4. Conclusions

This study focuses on developing a phenotyping system for aerial images of coffee trees using machine vision. A single camera drone was used for data gathering, and the researchers were able to establish a total number of 1,918 aerial images dataset of coffee trees. It was divided into two sets: training and testing, with a 70:30 ratio of 1,343 images for training and 575 for testing. Using a Python-based phenotyping system, the researchers were able to extract the morphological characteristics of leaf count, leaf color intensity, and canopy coverage of the coffee trees. VGG-16 was used for object detection and extraction of leaf count, NDVI was used to determine the level of greenness, and VGG-16 patch-based CNN model was used to estimate canopy coverage.

The evaluation results show that the model for object detection has a high accuracy rate with a value of 0.91478 for

accuracy, 0.925208 for precision, 0.9382 for recall, and 0.93166 for f1-score. The MAPE for leaf count was 11.6093%, and 15.6662% for canopy coverage, indicating that the phenotyping system has good forecasting accuracy in extracting the morphological characteristics of coffee trees.

The researchers would like to make the following recommendations for future researchers. These suggestions may help improve and enhance the phenotyping system for coffee trees. The researchers suggested getting the leaf color of coffee crops based on their nitrogen level content. With this, coffee crops can be determined if they are healthy or not. Other algorithms can be used by future researchers in getting the morphological attributes of coffee crops to possibly get a higher accuracy for the phenotyping system. Future researchers can make a Graphical User Interface (GUI) for the phenotyping system for coffee trees which will provide a convenient and intuitive way to operate the phenotyping system for coffee trees.

References

[1] Polat N, Memduhoğlu A, Kaya Y. Triangular greenness index analysis for monitoring fungal disease in pine trees: a uav-based approach. Bartın Orman Fak Derg [Internet]. 2024 Feb 15. Available from: https://doi.org/10.24011/barofd.1352729

[2] Concepcion RS, Lauguico SC, Tobias RR, Dadios EP, Bandala AA, Sybingco E. Estimation of photosynthetic growth signature at the canopy scale using new genetic algorithm-modified visible band triangular greenness index. In: 2020 international conference on advanced robotics and intelligent systems (ARIS) [Internet]; 2020 Aug 19-21; Taipei, Taiwan. [place unknown]: IEEE; 2020. Available from: https://doi.org/10.1109/aris50834.2020.9205787

[3] Lemes EM, Coelho L, Andrade SL, Oliveira AD, Marques MG, Nascimento FM, Cunha JP. Triangular greenness index to evaluate the effects of dicamba in soybean. AgriEngineering [Internet]. 2022 Aug 23;4 (3):758-69. Available from: https://doi.org/10.3390/agriengineering4030049

[4] De Ocampo AL, Bandala AA, Dadios EP. Estimation of triangular greenness index for unknown peakwavelength sensitivity of cmos-acquired crop images. In: 2019 IEEE 11th international conference on humanoid, nanotechnology, information technology, communication and control, environment, and management (HNICEM) [Internet]; 2019 Nov 29-Dec 1; Laoag, Philippines. [place unknown]: IEEE; 2019. Available from: https://doi.org/10.1109/hnicem48295.2019.9072796

[5] Patil SM, Choudhary S, Kholová J, Anbazhagan K, Parnandi Y, Gattu P, Mallayee S, Ksvv P, Kumar VP, Rajalakshmi P, Chandramouli M, Adinarayana J. UAV-based digital field phenotyping for crop nitrogen estimation using RGB imagery. In: 2023 IEEE IAS global conference on emerging technologies (globconet) [Internet]; 2023 May 19-21; London, United Kingdom. [place unknown]: IEEE; 2023. Available from: https://doi.org/10.1109/globconet56651.2023.10150110

[6] Lee K, Han X. A study on leveraging unmanned aerial vehicle collaborative driving and aerial photography

systems to improve the accuracy of crop phenotyping. Remote Sens [Internet]. 2023 Aug 7;15(15):3903. Available from: https://doi.org/10.3390/rs15153903

[7] Zhang Y, Yang Y, Zhang Q, Duan R, Liu J, Qin Y, Wang X. Toward multi-stage phenotyping of soybean with multimodal UAV sensor data: a comparison of machine learning approaches for leaf area index estimation. Remote Sens [Internet]. 2022 Dec 20;15(1):7. Available from: https://doi.org/10.3390/rs15010007

[8] Nguyen C, Sagan V, Bhadra S, Moose S. UAV multisensory data fusion and multi-task deep learning for high-throughput maize phenotyping. Sensors [Internet]. 2023 Feb 6;23(4):1827. Available from: https://doi.org/10.3390/s23041827

[9] Zahra S, Ruiz H, Jung J, Adams T. UAV-Based phenotyping: a non-destructive approach to studying wheat growth patterns for crop improvement and breeding programs. Remote Sens [Internet]. 2024 Oct 5;16(19):3710. Available from: https://doi.org/10.3390/rs16193710

[10] Zhu W, Sun Z, Huang Y, Yang T, Li J, Zhu K, Zhang J, Yang B, Shao C, Peng J, Li S, Hu H, Liao X. Optimization of multi-source UAV RS agro-monitoring schemes designed for field-scale crop phenotyping. Precis Agric [Internet]. 2021 May 3. Available from: https://doi.org/10.1007/s11119-021-09811-0

[11] Mndela Y, Ndou N, Nyamugama A. Irrigation scheduling for small-scale crops based on crop water content patterns derived from UAV multispectral imagery. Sustainability [Internet]. 2023 Aug 6;15 (15):12034. Available from: https://doi.org/10.3390/su151512034

[12] Wang H, Singh K, Poudel H, Ravichandran P, Natarajan M, Eisenreich B. Estimation of crop height and digital biomass from uav-based multispectral imagery. In: 2023 13th workshop on hyperspectral imaging and signal processing: evolution in remote sensing (WHISPERS) [Internet]; 2023 Oct 31-Nov 2; Athens, Greece. [place unknown]: IEEE; 2023. Available from: https://doi.org/10.1109/whispers61460.2023.10431021

[13] Xing M, Yang J, Song Y, Shang J, Zhou X, Wang J. Improved leaf area index retrieval using 3-D point clouds from UAV imagery. IEEE Geosci Remote Sens Lett [Internet]. 2024:1. Available from: https://doi.org/10.1109/lgrs.2024.3434685

[14] Gano B, Dembele JS, Ndour A, Luquet D, Beurier G, Diouf D, Audebert A. Using UAV borne, multi-spectral imaging for the field phenotyping of shoot biomass, leaf area index and height of west african sorghum varieties under two contrasted water conditions. Agronomy [Internet]. 2021 Apr 27;11(5):850. Available from: https://doi.org/10.3390/agronomy11050850

[15] Balota M, Sarkar S, Bennett RS, Burow MD. Phenotyping peanut drought stress with aerial remote-sensing and crop index data. Agriculture [Internet]. 2024 Apr 2;14(4):565. Available from: https://doi.org/10.3390/agriculture14040565

[16] Holman FH, Riche AB, Castle M, Wooster MJ, Hawkesford MJ. Radiometric calibration of 'commercial off the shelf' cameras for uav-based high-resolution temporal crop phenotyping of reflectance and NDVI. Remote Sens [Internet]. 2019 Jul 11;11(14):1657. Available from: https://doi.org/10.3390/rs11141657

[17] Chen Q, Zheng B, Chenu K, Hu P, Chapman SC. Unsupervised plot-scale LAI phenotyping via uav-based imaging, modelling, and machine learning. Plant Phenomics [Internet]. 2022 Jul 4;2022:1-19. Available from: https://doi.org/10.34133/2022/9768253

[18] Xu R, Li C, Bernardes S. Development and testing of a uav-based multi-sensor system for plant phenotyping and precision agriculture. Remote Sens [Internet]. 2021 Sep 4;13(17):3517. Available from: https://doi.org/10.3390/rs13173517

[19] Jin X, Li Z, Atzberger C. Editorial for the special issue "estimation of crop phenotyping traits using unmanned ground vehicle and unmanned aerial vehicle imagery". Remote Sens [Internet]. 2020 Mar 13;12(6):940. Available from: https://doi.org/10.3390/rs12060940