

Real-time assessment of plant photosynthetic pigment contents with an artificial intelligence approach in a mobile application

Kestrlia Rega Prilianti,^{1,2} Syaiful Anam,² Tatas Hardo Panintingjati Brotosudarmo,³ Agus Suryanto²

¹Department of Informatics Engineering, Universitas Ma Chung; ²Department of Mathematics, Universitas Brawijaya;

³Ma Chung Research Centre for Photosynthetic Pigments, Universitas Ma Chung, Malang, Indonesia

Abstract

The assessment of the photosynthetic pigment contents in plants is a common procedure in agricultural studies and can describe plant conditions, such as their nutritional status, response to environmental changes, senescence, disease status and so forth. In this report, we show how the photosynthetic pigment contents in plant leaves can be predicted non-destructively and in real-time with an artificial intelligence approach. Using a convolutional neural network (CNN) model that was embedded in an Android-based mobile application, a digital image of a leaf was processed to predict the three main photosynthetic pigment contents: chlorophyll, carotenoid and anthocyanin. The data representation, low sample size handling and developmental strategies of the best CNN model are discussed in this report. Our CNN model, photosynthetic pigment prediction network (P3Net), could accurately predict the chlorophyll, carotenoid and anthocyanin contents simultaneously. The prediction error for anthocyanin was $\pm 2.93 \mu\text{g/g}$ (in the range of 0-345.45 $\mu\text{g/g}$), that for carotenoid was $\pm 2.14 \mu\text{g/g}$ (in the range of 0-211.30 $\mu\text{g/g}$) and that for chlorophyll was $\pm 5.75 \mu\text{g/g}$ (in the range of 0-892.25 $\mu\text{g/g}$). This is a promising result as a baseline for the future development of IoT smart devices in precision agriculture.

Introduction

Photosynthetic pigments are the most easily observable determinant and diagnostic tool of plant performance during different phases of development. For example, the loss of green pigment colour can also be used to monitor leaf senescence, which is correlated with plant responses to internal and external changes (Esteban *et al.*, 2015). The development of an efficient method to quantify the changes in plant pigment composition, especially based on a non-destructive approach (Croft and Chen, 2017), has become a topic of paramount interest in current agricultural studies. The spectral reflectance data from a spectrophotometer-based measurement have been intensively used for non-destructive pigment quantification of an intact leaf (Gitelson and Solovchenko, 2017, 2018). Unfortunately, spectral reflectance data cannot be easily generated and is also costly. With the recent development of computer technology, digital imaging combined with artificial intelligence has been extensively applied in the development of non-destructive plant evaluation methods. Those methods are proven to be efficient and quite accurate and can be used for real-time analysis (Inácio and Rieder, 2018).

The artificial neural network (ANN), which is a well-known method in artificial intelligence, has been used in agricultural studies to conduct plant evaluations. Using digital images as input, the ANN has been used to estimate leaf areas, identify disease symptoms, predict yields, classify varieties, and predict leaf chlorophyll contents (Gallego *et al.*, 2011). The main challenge, however, is the feature extraction task. Prior to the learning process, the most significant features of the digital image must be determined. The features will be the input of the ANN architecture. Here, human knowledge and creativity become the primary determinants to define the best features to ensure the best performance of the ANN. The convolutional neural network (CNN) method was invented to address this situation so that the feature extraction process from digital image data can be automatically performed. To improve the CNN performance, LeCun *et al.* (1998) implemented backpropagation (a supervised learning algorithm), which allowed the CNN to provide good results for the first time. In the following years, the CNN performance was further improved. By adding the rectified linear unit (ReLU) and the dropout concepts, developed AlexNet, which succeeded in classifying the Imagenet data into one of 1000 classes. Due to all of its advantages, the applications of the CNN in agricultural research have rapidly increased (Kamilaris and Prenafeta-Boldú, 2018).

The CNN input consists of the digital images, which are processed using digital image processing methods. Therefore, no feature extraction processes are needed on the original image, unlike in the case of the ANN. Those features will be extracted in the convolution layer through a constructive random process. The convolutional term in the CNN refers to the convolution technique in morphological image processing. By using a kernel, the convo-

Correspondence: Kestrlia Rega Prilianti, Department of Informatics Engineering, Universitas Ma Chung, Villa Puncak Tidar Blok N-01, Malang 65151, Indonesia.

E-mail: kestrilia.rega@machung.ac.id ; lia.rega@gmail.com

Key words: Artificial intelligence; convolutional neural network; digital image; mobile application; non-destructive method; photosynthetic pigments.

Conflict of interests: the authors declare no potential conflict of interests.

Received for publication: 13 May 2020.

Accepted for publication: 19 August 2020.

©Copyright: the Author(s), 2020

Licensee PAGEPress, Italy

Journal of Agricultural Engineering 2020; LI:1082

doi:10.4081/jae.2020.1082

This article is distributed under the terms of the Creative Commons Attribution Noncommercial License (by-nc 4.0) which permits any non-commercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

lution allows us to manipulate morphologically digital images using techniques such as edge detection, blurring, sharpening, and embossing. These improve the efficiency of the forward process in the CNN and diminish the total number of the parameters used in the network. This is also the reason why the CNN is known to be superior in performing object classifications and even object recognition on digital images. In general, CNN-based projects in agricultural studies focus mainly on the object shape during the classification or recognition task (Kamilaris and Prenafeta-Boldú, 2018). In our research, the colour was the main parameter that was considered to quantify the photosynthetic pigment contents in plants. Previously, it was shown that the CNN could properly learn plant digital images based on their colours, also in tasks such as species classification (Ghazi *et al.*, 2017), phenotyping (Ubbens and Stavness, 2017), and disease detection (Mohanty *et al.*, 2016). However, colour as the main feature and its correlation with the pigment composition in plant leaves has not been intensively studied in CNNs. Nevertheless, since Gitelson and Merzlyak (2004) showed that the intact leaf reflectance at a certain electromagnetic wavelength has a strong correlation with its photosynthetic pigment content, we hypothesized that the CNN model could also perform well in determining the nonlinear relationship between colours in leaf images and their photosynthetic pigment contents.

The main objective of our research was to develop an intelligent system for predicting plant photosynthetic pigment contents from leaf images. We focused on the implementation of the CNN method to create a model that describes the relationship between the image of a plant leaf and the contents of its main photosynthetic pigments (chlorophyll, carotenoid, and anthocyanin). For the ease of *in situ* assessment, the model was embedded into a smartphone via an android-based application. We evaluated the performance of the application related to the colour constancy problem, which in general is one of important issues to be considered in the implementation of the digital imaging method. Furthermore, we also evaluated the performance of the application across species. Herewith, we report a novel study on the use of digital images as an alternative to the spectrophotometer-based spectral reflectance data of an intact leaf for performing its photosynthetic pigment quantification.

Materials and methods

Leaf samples

In this research work, we used four species of Indonesian herbal plants, *i.e.*, *Syzigium oleana*, *Piper betle*, *Jasminum* and *Graptophyllum pictum*. The plants were selected to represent a variety of photosynthetic pigment contents in leaves. *Syzigium oleana* contained high concentrations of carotenoid and anthocyanin (Anggraini, 2017). *Graptophyllum pictum* contained high concentration of anthocyanin and chlorophyll (Rosmala *et al.*, 2016). *Jasminum* contained high concentrations of chlorophyll and carotenoid (Sabharwal *et al.*, 2013). *Piper betle* was selected to support the supplementary data because of its unique variations in its carotenoid, chlorophyll and anthocyanin contents (Preethy, 2014). To prepare the samples, each leaf was carefully selected based on its visual colour, health, and position from the terminal bud. We ensured that each leaf visual colour represented the colour diversity within the plant. The leaves must be healthy, should not look wilted or dry and have no signs of any pest attacks. The leaf samples should be taken from those located no more than five leaves below the terminal

buds. The data collection was carried out as follows: first, a digital image was taken of each leaf, and then, an organic solvent was applied to extract the necessary contents to measure the pigment concentrations using a spectrophotometer. Each detailed step will be described further in the following subsections. We measured a total of 212 leaf samples (*Syzigium oleana* = 62, *Graptophyllum pictum* = 51, *Jasminum* = 55, and *Piper betle* = 44) that were collected from several areas in Malang city, East Java, Indonesia. Those data samples were used to train the CNN model in describing the relationships between a leaf image and its photosynthetic pigment contents.

Leaf images

For the benefit of the CNN model development, the environment during the image acquisition was controlled. The leaf images were taken indoors and the illumination inside the room was stable at 1240 lux. A smartphone digital camera (13 MP, f/2.2, Sony IMX 519 CMOS sensor, VIVO Y83, 2018) was used to take the leaf images. All images were captured with the flash turned off and with the auto-mode turned on. They were saved as .jpeg files.

The CMOS sensor on most smartphone cameras captures object reflectance in the red (590-720 nm), green (480-600 nm), and blue (400-540 nm) range with the help of a Bayer filter. Millions of other colours are generated with a demosaicing algorithm. Meanwhile, chlorophyll reflects green light in the range of 500-650 nm, carotenoid reflects yellow-red light in the range of 500-760 nm and anthocyanin reflects red light around 700 nm. Therefore, the leaf reflectance quantification by the green and red filters of the camera sensor will greatly determine the accuracy of the photosynthetic pigment content quantification.

To facilitate the collection of the images of the leaf samples, a simple Android-based application was also built. The application enabled us to estimate the proper camera-object distance with a virtual frame, because the entire area of each leaf analysed must fall within the frame. This makes it possible to ensure that the camera-object distance is always stable, when pictures are taken. Moreover, a white sheet of HVS paper is needed as a calibration aid. With the help of the thresholding algorithm, the application took the colour from the white paper and then used it to correct the brightness of the image. It is assumed that the maximum intensity for each RGB component of the white paper is 255. Therefore, in so doing, one can be sure that the brightness of the images is stable for all pictures.

Laboratory analysis

Each leaf was cut into two parts of equal size. One part was prepared for the chlorophyll and carotenoid quantification, while the other part was for the anthocyanin quantification. Each part of the leaf was cut further into small pieces. To extract the pigments, 1.50 mL of organic solvent was added to the leaf materials (0.05 g) along with CaCO₃ and sodium ascorbate powder to prevent the degradation of the pigments due to oxidation. Acetone (p.a.) was used for chlorophyll and carotenoid extraction. For anthocyanin extraction, we used a mixture of MeOH:HCl:H₂O (90:1:1, v/v). The leaf materials were then ground and evenly mixed using a vortex for 1 minute, and then were incubated in ice for one minute. The procedure was repeated 3 times until complete discolouration of the leaf material. Next, the sample was centrifuged at 14,000 rpm for 2 min. The supernatant containing the pigment extract was separated from the pellet and then was kept in ice for further measurements. The absorption spectrum was measured using a UV-vis spectrophotometer (Shimadzu UV-1800 Double Beam UV/Visible Scanning, Shimadzu-Japan). The chlorophyll and carotenoid con-

centrations were calculated using Lichtenthaler's equations (Lichtenthaler, 1987). For the calculation of the anthocyanin concentration, we used the method that was published by Sims and Gamon (2002). Table 1 shows the statistical summary of calculated pigment contents for each plant species. A large standard deviation indicated that the necessary diversity of the pigment content data to train the CNN was achieved.

Model development

Each leaf image is used as input and includes three pigment content values (anthocyanin, carotenoid, and chlorophyll) as the target output. The data were divided into two parts, *i.e.*, the training set and the validation set. The data were randomly separated into these sets during the training process, with 80% for the training set and 20% for the validation set. The training set was used to build the CNN models, while the validation set was used to evaluate the model performance in predicting the new data. To speed up the training process, we applied data normalization prior to the training process (Equation 1), where z is the normalized data and x is the raw data. Therefore, z will be in the range of 0 to 1.

$$z = \frac{x - \min(x)}{[\max(x) - \min(x)]} \quad (1)$$

The basic architecture of a simple CNN consists of convolution layers (with or without max pooling) and fully-connected layers. The number of convolution layers and fully-connected layers was changed to obtain the best model. We evaluated four CNN models (Table 2). Three of them were modified from a well-known CNN architecture, *i.e.*, LeNet (LeCun *et al.*, 1998), AlexNet (Krizhevsky *et al.*, 2012) and VGGNet (Simonyan and Zisserman, 2015). The photosynthetic pigment prediction network (P3Net) was our original CNN model. Our purpose was to develop an intelligent system for a prediction task, not a classification task. Therefore, LeNet, AlexNet, and VGGNet, which were initially used for classification problems, should be modified. The modification of those three models was mainly in the output layer. In the classification problems, the output nodes represent the classes. In our pigment content prediction problem, each output node would represent the predicted anthocyanin, carotenoid, and chlorophyll contents. Especially for VGGNet, the number of hidden layers was also modified. The original sixteen hidden layers were reduced to nine hidden layers (VGG-9). We reduced them prior to overfitting occurring as a result of the low sample size data. In the presence of this deficiency, the parameters that are involved in the model cannot be too large (Truong *et al.*, 2018). However, our experiment showed that, although the complexity of the model was low, overfitting still occurred. Therefore, several methods to eliminate overfitting were implemented, including

regularization, dropout, and data augmentation.

We used the LeakyReLU as the activation function in the output nodes. The LeakyReLU formula is shown in Equation 2, where $h^{(i)}$ is the hidden unit activation, $w^{(i)}$ is the weight vector for the i^{th} hidden unit and x is the input.

$$h^{(i)} = \max(w^{(i)T}x, 0) = \begin{cases} w^{(i)T}x, & w^{(i)T}x > 0 \\ 0.7 \cdot w^{(i)T}x, & \text{else} \end{cases} \quad (2)$$

Seven gradient descent-based optimization methods were compared to train the CNN models, including the stochastic gradient descent (SGD), adaptive gradient (Adagrad), adaptive delta (Adadelta), root mean square propagation (RMSProp), adaptive momentum (Adam), adaptive max pooling (Adamax), and nesterov adaptive momentum (Nadam). For all the methods, the mean squared error (MSE) was used as the loss function. The MSE was calculated using Equation 3, where y_i was the actual pigment content, \hat{y}_i was the predicted pigment content and n was the sample size. We conducted our evaluation prior to acquiring the CNN model that best fitted the data and had the lowest complexity in order to provide a high predictive performance.

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (3)$$

Software development

We developed the CNN model using Python 2 and Keras API with the TensorFlow backend. The CNN experiment was run using Google Colaboratory on a personal computer with a 1.60 GHz Intel Core i5 processor, 8 GB of DDR 3 RAM, and the Sierra operating system. The best CNN model was saved in an HDF5 file. The file was then embedded into the Android application using Android Studio.

Performance indicators

To simplify the interpretation, the mean absolute error (MAE) was used to evaluate the performance of the CNN model to predict the actual content of the photosynthetic pigments in the leaf being analysed. It was also used to evaluate the performance of the Android application in particular conditions to justify its reliability. Equation 4 shows the MAE formula, where y_i is the actual pigment content, \hat{y}_i is the predicted pigment content and n is the sample size.

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (4)$$

Table 1. Statistical summary of the laboratory analysis.

Variety	\sum samples	Anthocyanin ($\mu\text{g/g}$)*		Carotenoid ($\mu\text{g/g}$)*		Chlorophyll ($\mu\text{g/g}$)*	
		Mean	St.Dev.	Mean	St.Dev.	Mean	St.Dev.
<i>Syzigium oleana</i>	62	59.45	73.33	31.97	32.84	124.21	147.95
<i>Piper betle</i>	44	4.58	6.61	8.94	9.92	30.48	37.23
<i>Jasminum</i>	55	0.00	0.00	44.24	39.29	306.57	171.83
<i>Graptophyllum pictum</i>	51	47.32	67.32	26.58	26.08	169.75	161.54

*Pigment content relative to the dry weight of the leaf material. St.Dev., standard deviation.

Results and discussion

Convolutional neural network model

The evaluation of the performance of the CNN models includes 3 main aspects regarding the critical success factors in developing the best model, *i.e.*, i) data representation; ii) low sample size handling; and iii) model complexity.

Data representation

In this research, colour is the main feature that the CNN model must learn. Therefore, different lighting conditions that result in colour constancy problems for the input images must be taken into consideration. In addition to applying the calibration procedure using a white sheet of paper, we also consider the use of certain types of colour models. We compared three colour models to determine the best one for overcoming colour constancy problems, including the RGB, HSV, and LAB models. The RGB colour model is the standard format for a smartphone camera. However, it is known that RGB cannot guarantee colour constancy. Indeed, the colour component (R, G, and B) values are produced by the same object. However, the reflected colour images that are captured by a camera in different lighting conditions are not identical. Unlike RGB, the HSV model allows us to separate the object colour from the surrounding environment illumination. Therefore, the HSV concept is considered to be similar to how humans see colours. We hypothesized that the application of the HSV colour model might improve the prediction accuracy. Similar to the HSV, the LAB model also separates colour components based on the illumination. The LAB model is often used as a standard in calorimetry. Other than the colour constancy problem, the issue regarding the region of interest (ROI) of the input image was also investigated. We evaluated two forms of leaf visualization. The first form visualized both the leaf and the background. The second one visualized only the leaf area. We compared those two visualizations given that the CNN model in this study was trained to recognize colours, and we assumed that the exclusion of the background might increase the

accuracy. We found that the CNN models with the HSV and LAB inputs produced smaller MAEs, when using the full leaf area with the background visualization. On the other hand, the CNN model with the RGB input provided the opposite result. Among the three-colour model, the LAB performed better predictions. It produced the smallest validation MAE (0.01253) and also did not overfit like the other two-colour models. It was confirmed that the colour model differences in the input images affected the CNN performance. In general, the HSV model provided the largest MAE, while the RGB and LAB models provided smaller, similar MAEs. Nevertheless, the RGB model was difficult to train and it always overfitted. On the contrary, the LAB model never overfitted. Given that the absence of overfitting is an indication of a successful minimization of the generalization errors of the future output, we concluded that the LAB was the best colour model to be applied to the input image of the CNN models. Moreover, the LAB is a device independent colour model, which gave us the chance to convey assorted colours crosswise over the various gadgets.

Low sample size handling

The laboratory procedures to measure the pigment contents were costly and time-consuming. This was a limit in the choice of the sample sizes. Therefore, unlike other CNN studies, the biggest challenge to train the CNN models used in this research was the low sample size data. Although many researchers argue that the CNN can only obtain superior performance, if trained with a large dataset, other researchers proved that the CNN can be trained with a low sample size data and still obtain superior performance (Keshari *et al.*, 2018). Due to this scientific evidence, we believe that we will be able to find a particular CNN model that can be trained appropriately, using our dataset. However, low sample size data resulted in severe overfitting problems in the training process of the CNN model. Overfitting is a condition where the model is too fit with the training data so that it cannot provide proper predictions for the validation data. Therefore, it was necessary to implement overfitting removal methods, such as regularization, dropout, and image augmentation (Wang and Perez, 2017). Our experiment showed that regularization and dropout were not sig-

Table 2. Convolutional neural network architectures used in the experiment.

Layer		LeNet 54×54×3	P3Net 54×54×3	Alexnet 120×120×3	VGG-9 120×120×3
Convolution	1	6 filters of size 5×5 with max pooling	32 filters of size 3×3	96 filters of size 11×11 with max pooling	64 filters of size 5×5
	2	16 filters of size 5×5 with max pooling	32 filters of size 3×3 with max pooling	256 filters of size 5×5 with max pooling	64 filters of size 3×3 with max pooling
	3	-	32 filters of size 3×3 with max pooling	384 filters of size 3×3 with max pooling	128 filters of size 3×3
	4	-	-	384 filters of size 3×3	128 filters of size 3×3 with max pooling
	5	-	-	256 filters of size 3×3	256 filters of size 3×3
	6	-	-	-	256 filters of size 3×3
	7	-	-	-	256 filters of size 3×3 with max pooling
Fully Connected	1	120 nodes, ReLU	500 nodes, ReLU	4096 nodes, ReLU	4096 nodes, ReLU
	2	85 nodes, ReLU	500 nodes, ReLU	4096 nodes, ReLU	4096 nodes, ReLU
	3	-	500 nodes, ReLU	-	-
Output		3 nodes, LeakyReLU	3 nodes, LeakyReLU	3 nodes, LeakyReLU	3 nodes, LeakyReLU

nificantly effective in removing overfitting. Conversely, hand image augmentation proved to be very effective in removing overfitting. Figure 1 depicts the MAE comparison of the training and validation processes of P3Net model with the implementation of regularization, dropout, and image augmentation. As can be seen in Figure 1D, the overfitting disappeared after the application of image augmentation. Therefore, in the end, image augmentation was used to overcome the overfitting problem.

We compared four augmentation techniques, including random rotation, zooming, feature-wise centring, and feature-wise standard deviation normalization. Random rotation and zooming did not change the values of L, A and B from the input image, because the changes that occurred were spatial changes. Feature-wise centring and feature-wise standard deviation normalization produced variations in the values of L, A and B from the input image. The centring process altered the averages of the L, A, and B values of all training data to zero. Meanwhile, standard deviation normalization divided the L, A, and B values according to the standard deviation of the entire training data. When centring and standard deviation normalization were combined, the standardized image could be formed. We found that the implementation of multiple methods was better in removing overfitting compared to implementing a single augmentation method. However, the intensity-based augmentations (feature-wise centring and feature-wise standard deviation normalization) led to more severe overfitting than in the case of spatial-based augmentation (random rotation and zooming).

Therefore, intensity-based augmentation was not applied in our next experiments. The application of the spatial-based augmentation technique (rotation and zooming) provided a 19% reduction of the MAEs and also eliminated overfitting better than the application of a single technique.

Model complexity

The four architectures that we present in this report were carefully chosen to represent the complexity of the network. The simplest is the LeNet architecture, while the others are the P3Net, the AlexNet and the VGG-9 architectures. All architectures were trained using a batch size of 100 and 1000 epochs. For each architecture, the training process was conducted 10 times. The average MAE was calculated for the performance analysis. The training MAEs of all architectures were found to be slightly better than their validation MAEs, which are reasonable and do not indicate any overfitting. It can be concluded that the implementation of the spatial-based data augmentation method worked well in preventing overfitting not only in the P3Net architecture, but also in the other three architectures. Moreover, we found that the performance of CNN was influenced by the selected optimization method. Experiments in all architectures showed that the newer variant of the gradient descent optimization method provided better MAEs. The SGD appeared to provide the largest MAEs for all CNN architectures. Nadam provided the smallest MAEs for LeNet (training MAE of 0.01759 ± 0.00216 and validation MAE of

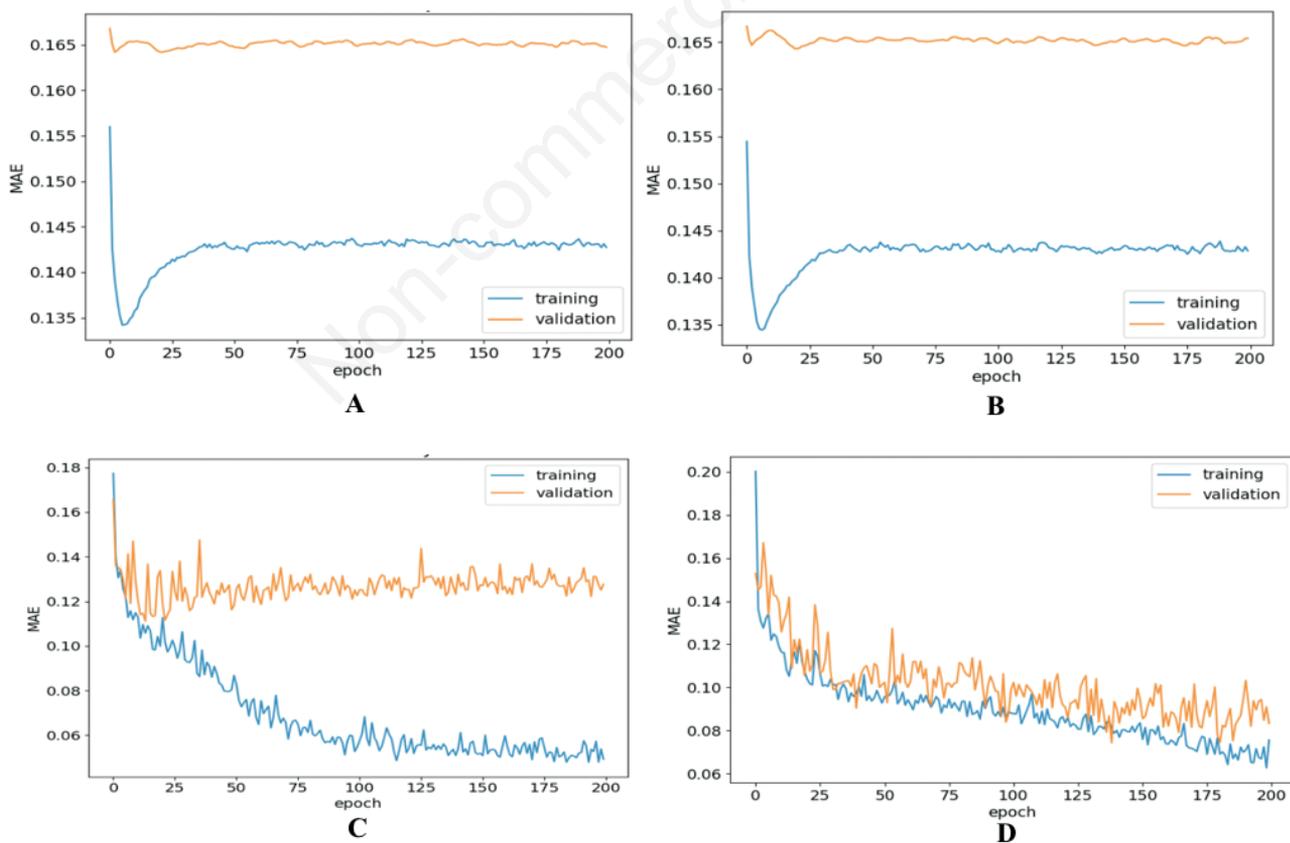


Figure 1. Comparison of the training and validation mean absolute errors (MAEs) of the P3Net using overfitting removal methods; A) L1 regularization; B) L2 regularization; C) dropout; D) image augmentation.

0.03851±0.00543), while Adamax provided the smallest MAEs for P3Net (training MAE of 0.00624±0.00071 and validation MAE of 0.01249±0.00219), AlexNet (training MAE of 0.00584±0.00117 and validation MAE of 0.01204±0.00187) and VGG-9 (training MAE of 0.00867±0.00419 and validation MAE of 0.01888±0.00443). After applying the best optimization method, we compared the performance of all four architectures with regard to the network complexity and its trained model file size (Figure 2). The CNN architecture in our experiment provided quite different results from several prior studies. Cruz *et al.* (2019) and Too *et al.* (2018) applied several CNN architectures for plant disease.

Their studies showed that by increasing the complexity of the network, the accuracy of the system increased, and vice-versa. In our experiment, an increased number of hidden layers did not necessarily result in a significant reduction in MAEs. The P3Net, which has a less complex architecture than AlexNet and VGG-9, proved to provide an equivalent MAE. This is a promising result, because the P3Net architecture ensures an easy implementation of the model in mobile device applications, since its file size (9.40 Mb) is very much smaller compared to AlexNet (442.90 Mb) and VGG-9 (335.70 Mb). However, as with most other studies, CNN architectures that are too low in complexity (LeNet) cannot promise good results. To further describe the performance of P3Net, we compared the validation MAEs for the prediction of each photosynthetic pigment. It was found that P3Net performance was not the same in predicting the content of each photosynthetic pigment. Out of the three pigments, anthocyanin provided the smallest MAE (0.00578) and the lowest standard deviation, while chlorophyll and carotenoid showed the opposite result (carotenoid's MAE =0.00899, chlorophyll's MAE =0.02344). Therefore, it seemed that anthocyanin was easier to learn for the P3Net. We observed that the behaviour was in line with the results

presented by Huang *et al.* (2014), *i.e.* the higher the anthocyanin content in leaves, the greater the prediction error of the chlorophyll content. Therefore, since 53% of our leaf samples contained anthocyanin, we obtained a large MAE for the chlorophyll content prediction. Therefore, this phenomenon can be explained as follows. In spectral based pigment prediction methods (data acquisition with spectrophotometers), the main reflectance of chlorophyll and anthocyanin was known to be in the green range. Therefore, the anthocyanin reflectance in the green range made the reflectance of the chlorophyll in the same range difficult to measure. Gitelson *et al.* (2001) experimented on several maple leaves with relatively similar chlorophyll contents, but different anthocyanin contents. The experimental results showed that a higher anthocyanin content led to lower reflectance in the green range. One of the quantitative examples that they presented is the reflectance comparison of the two maple leaves that both had chlorophyll contents of ±9.50 nmol/cm², where the first leaf contained <0.30 nmol/cm² of anthocyanin and the second leaf contained <22.90 nmol/cm² of anthocyanin. It was found that the first leaf reflectance was 23%, while the second leaf reflectance dramatically dropped to 7%. Even in dark-green leaves with high chlorophyll contents, a small amount of anthocyanin (±3 nmol cm²) can significantly decrease the reflectance in the green range. Hence, the development of spectral-based non-destructive methods (especially those using the green spectral range) for the prediction of the chlorophyll content became more difficult for anthocyanin-containing leaves. Further research will be conducted to find the right technique to reduce the anthocyanin reflectance and predict the chlorophyll content. Huang *et al.* (2014) formulated a spectral-based index, which was claimed to be able to accurately measure the chlorophyll content in leaves with various anthocyanin levels. However, its application to digital leaf images has never been studied so far.

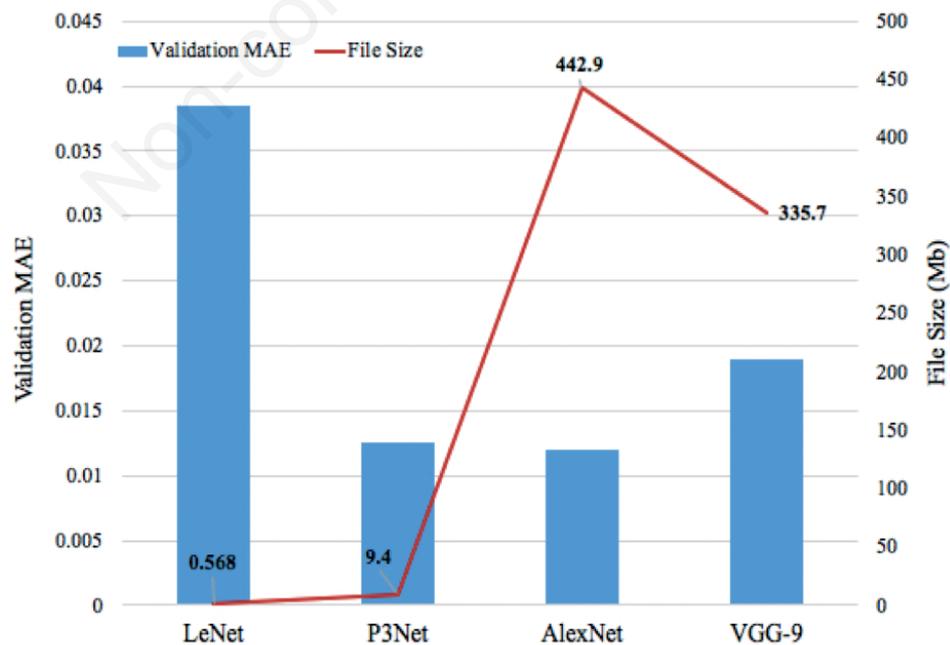


Figure 2. Validation mean absolute error (MAE) comparison for each convolutional neural network model along with its file size. The files were in the HDF5 format.

The mobile application (Leaf Piction)

Figure 3 illustrates the overall design of our plant photosynthetic pigment assessment system, which worked as follows. The user would take a photo of the plant leaf with a sheet of white paper as the background. The calibration process as described in the previous subsection was done automatically, when the camera sensor captured an object. The leaf image was sent as the input to the CNN model (P3Net) embedded in the Android application named Leaf Piction. The P3Net predicted the pigment contents and the results were displayed live on the smartphone screen. The prediction of the three photosynthetic pigment content could be seen at the bottom of the leaf image. If the user wanted to save the prediction result, he could press the save button and the data would be stored on the smartphone internal memory.

To evaluate Leaf Piction performance in relation to the colour

constancy problem, we observed changes in the content of the predicted pigments due to differences in the environment illuminance, when the leaf image was taken. We conducted a total of 128 indoor and outdoor experiments with 4 different illuminance categories, *i.e.* low (30-60 lux), medium (100-700 lux), high (1000-2500 lux), very high (3000-6000 lux). We found that the calibration procedure using a sheet of white paper significantly reduced the problem. Figure 4 shows the difference in the average MAE with and without the calibration procedure. It appears that the use of the calibration procedure can reduce the average MAE by 61.48% in the outdoor experiments and 49.55% in the indoor experiments. Outdoor assessments have a greater potential to produce less accurate data. Despite the calibration procedure, the average MAE of the outdoor assessments was still greater than the average MAE of indoor assessments. We investigated the outdoor illuminance level in

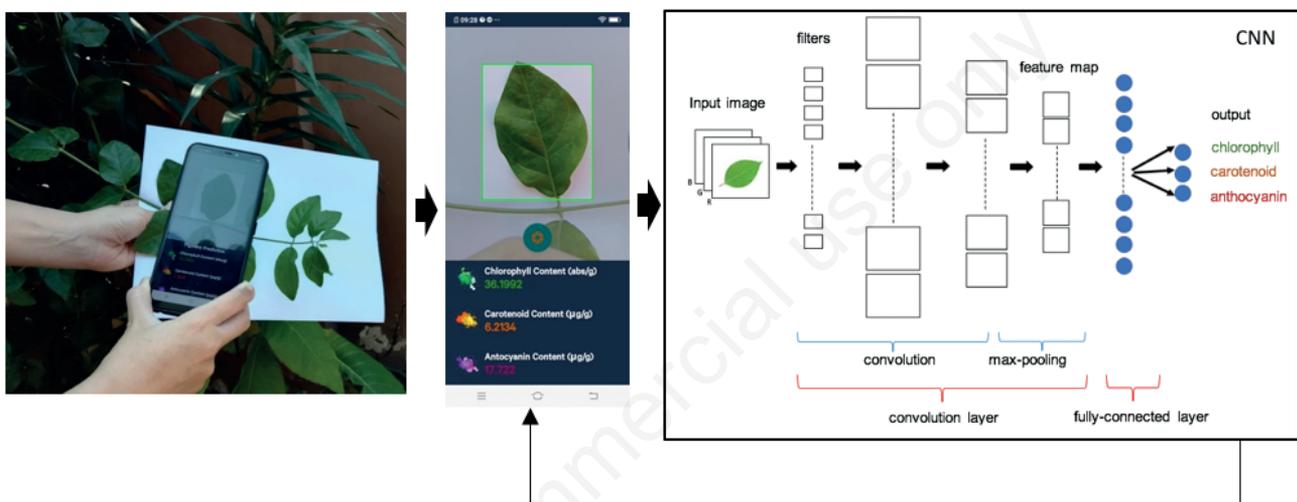


Figure 3. Real-time plant photosynthetic pigment assessment system using a mobile application. Left to right, image acquisition, Leaf Piction user interface, and the P3Net architecture for the prediction model. The whole leaf area under observation should fall within the virtual frame of the application, *i.e.* the green rectangle on the smartphone screen.

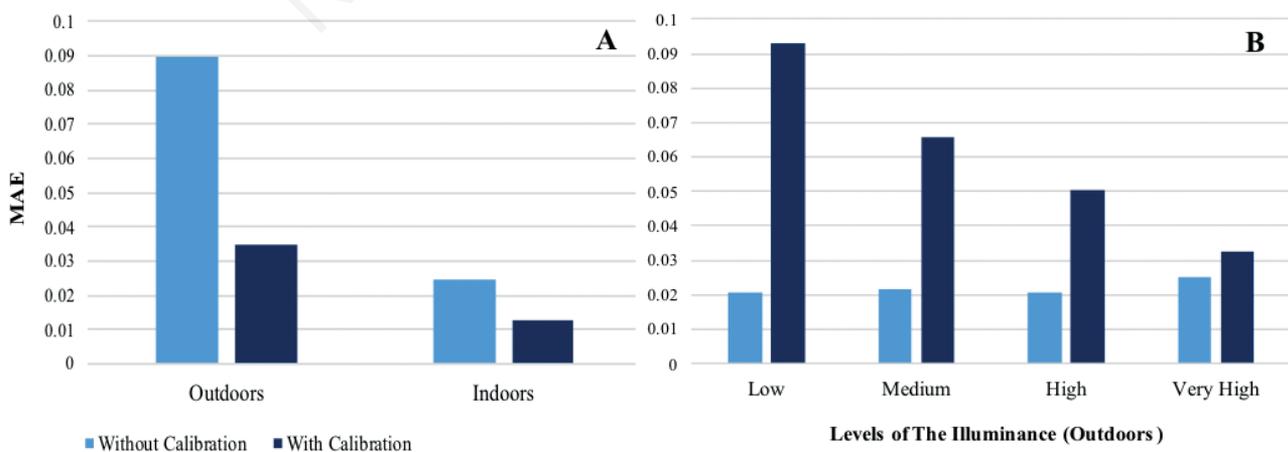


Figure 4. Performance evaluation of Leaf Piction with and without calibration procedure; A) outdoor-indoor experiment; B) experiment with 4 different levels of outdoor illuminance.

greater depth and found that without a calibration procedure the better the illuminance, the better the prediction accuracy (smaller MAE). In the condition with the calibration procedure, regardless of the illuminance level, the prediction accuracy tended to be the same (relatively equal MAE).

We also tested Leaf Piction robustness across species. We added a total of 27 leaf samples from 3 different species, *i.e.*, *Cananga odorata*, *Annona muricata*, and *Duranta erecta*. In the first experiment, we observed the behaviour of P3Net that was trained with *G. pictum*, *Jasminum*, *P. betle*, and *S. oleana* in providing predictions for the *C. odorata*, *A. muricata* and *D. erecta*. We found that P3Net no longer provided good results in the assessment of the species that were not trained before (Figure 5A). The biggest drawback was found in from the predicted value of chlorophyll. Compared to the average MAE for the other two pigments, the average MAE for chlorophyll appeared to be very large. Then, we performed a second experiment by repeating the training process for the P3Net with those 3 additional species (*Cananga odorata*, *Annona muricata*, and *Duranta erecta*). After the training process, we updated Leaf Piction with the newly trained P3Net. We found a significant decrease in the average MAE by approximately 70%. Despite the experiment was run only with a very small number of samples (9 leaves for each species), it was able to show a decreasing trend in the average MAE (Figure 5B). Therefore, we believe that by increasing the number of samples for the 3 species, its average MAE will be equivalent to the previous 4 species. Therefore, whenever there is a need for a particular species assessment, the P3Net needs to be updated by retraining it with the new data from that species. Further studies are also needed to explore the correlation between the phenotypic similarities of different species and the similarity of its photosynthetic pigment content behaviour. The results of the first experiment in Figure 5A shows that compared to *C. odorata* and *A. muricata*, the average MAE of *D. erecta* looked much smaller, even though all of them had never been trained on P3Net. Such condition may be possible because of the phenotypic similarity of *D. erecta* with one of the 4 species that was trained on P3Net.

Conclusions

In this study, we have developed P3Net, which is a CNN architecture that has the ability to predict the photosynthetic pigment contents in plant leaves using digital images. We showed that CNN was powerful for regression tasks. The P3Net performed better than the other popular CNN architectures, including LeNet, AlexNet, and VGG-9. In the presence of low sample size data, its training MAE reached 0.00624 ± 0.00071 and its validation MAE was 0.01249 ± 0.00219 (for normalized data that were scaled to 0-1). Low sample size data were the main challenge, because they resulted in severe overfitting in all the tested architectures. However, the application of spatial-based augmentation methods proved able to remove the overfitting. Therefore, we also showed that a CNN architecture could be well-trained, even when in the presence of low sample size data. We demonstrated that the best data representation was using the LAB colour model and the best ROI was using the full leaf area with a background. Moreover, since the pigment content in plants was strongly correlated with its visual colour, our experiment showed that the filters on P3Net, which were generally known to be superior in recognizing objects using shape features, were also superior in recognizing colour features. However, in order to gain a better understanding of its behaviour in greater detail, further research is needed. Finally, the implementation of P3Net on mobile applications (Leaf Piction) made it easier to use in the field. The user simply needs to use the camera on the leaf of the plant. Then, the application would display a prediction of the 3 photosynthetic pigment contents in real-time. In a denormalized form, the prediction error was $\pm 2.93 \mu\text{g/g}$ for anthocyanin (in the range of 0-345.45 $\mu\text{g/g}$), $\pm 2.14 \mu\text{g/g}$ for carotenoid (in the range of 0-211.30 $\mu\text{g/g}$) and $\pm 5.75 \mu\text{g/g}$ for chlorophyll (in the range of 0-892.25 $\mu\text{g/g}$). Leaf Piction accuracy was shown to be stable against varying illuminance, but not against out sample species (species that have not been trained on P3Net). However, the ability of Leaf Piction to predict the content of the photosynthetic pigments can be easily updated for other species of plants using an artificial intelligence-based approach.

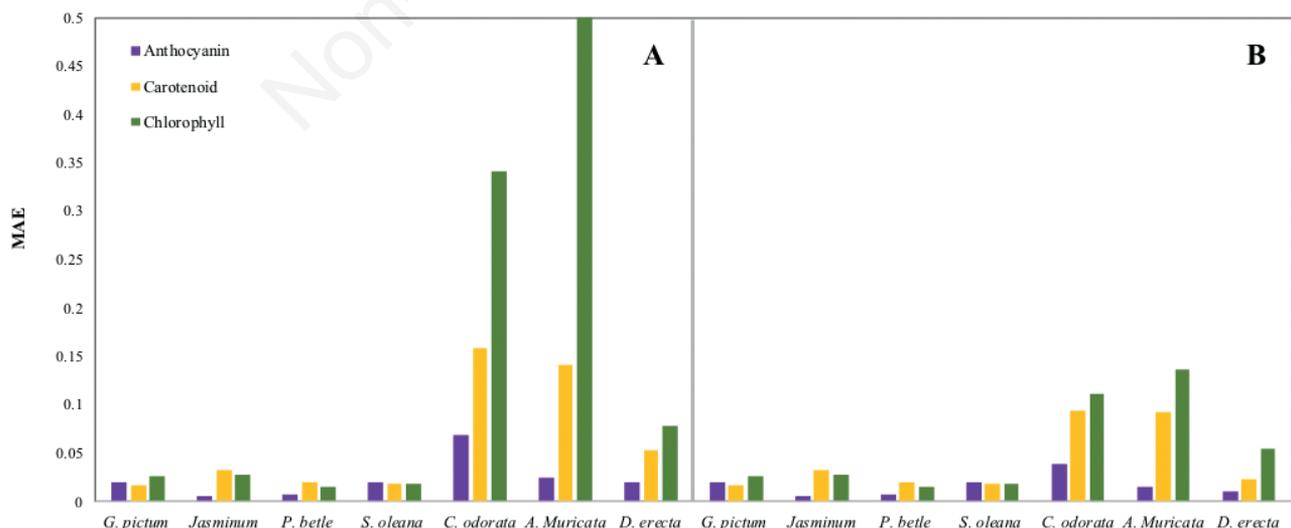


Figure 5. Performance evaluation of Leaf Piction; A) first experiment, the CNN model (P3Net) embedded in Leaf Piction was trained by *Graptophyllum pictum*, *Jasminum*, *Piper betle* and *Syzigium oleana*; B) second experiment, the P3Net was by all species, but with a limited number of samples for *Cananga odorata*, *Annona muricata*, and *Duranta erecta*.

References

- Anggraini T. 2017. Antioxidant activity of *Syzygium oleana*. Pakistan J. Nutr. 16:605-11.
- Croft H., Chen J.M. 2017. Leaf pigment content, comprehensive remote sensing. Elsevier Inc., Amsterdam, Netherlands.
- Cruz A., Ampatzidis Y., Pierro R., Materazzi A., Panattoni A., De Bellis L., Luvisi A. 2019. Detection of grapevine yellows symptoms in *Vitis vinifera* L. with artificial intelligence. Comput. Electron. Agric. 157:63-76.
- Esteban R., Barrutia O., Artetxe U., Fernández-Marín B., Hernández A., García-Plazaola J.I. 2015. Internal and external factors affecting photosynthetic pigment composition in plants: a meta-analytical approach. New Phytol. 206:268-80.
- Gallego P.P., Gago J., Landín M. 2011. Artificial neural networks technology to model and predict plant biology process. In: K. Suzuki (ed.) Artificial neural networks - methodological advances and biomedical applications. InTech., London, UK, pp 197-216.
- Ghazi M.M., Yanikoglu B., Aptoula E. 2017. Plant identification using deep neural networks via optimization of transfer learning parameters. Neurocomputing 235:228-35.
- Gitelson A.A., Merzlyak M.N. 2004. Non-destructive assessment of chlorophyll, carotenoid and anthocyanin content in higher plant leaves: principles and algorithm. In: S. Stamatiadis, J.M. Lynch, J.S. Schepers (Ed.), Remote sensing for agriculture and the environment. Greece Ella, pp 78-94.
- Gitelson A.A., Merzlyak M.N., Chivkunova O.B. 2001. Optical properties and nondestructive estimation of anthocyanin content in plant leaves. Photochem. Photobiol. 74:38-45.
- Gitelson A., Solovchenko A. 2017. Generic algorithm for estimating foliar pigment content. Geophys. Res. Lett. 44:9293-98.
- Gitelson A., Solovchenko A. 2018. Non-invasive quantification of foliar pigments: possibilities and limitations of reflectance and absorbance-based approaches. J. Photochem. Photobiol. B Biol. 17:537-44.
- Huang W.D., Lin K.H., Hsu M.H., Huang M.Y., Yang Z.W., Chao P., Yang C.M. 2014. Eliminating interference by anthocyanin in chlorophyll estimation of sweet potato (*Ipomoea batatas* L.) leaves. Bot. Stud. 55:1-10.
- Inácio D., Rieder R. 2018. Computer vision and artificial intelligence in precision agriculture for grain crops: a systematic review. Comput. Electron. Agric. 153:69-81.
- Kamilaris A., Prenafeta-Boldú F.X. 2018. Deep learning in agriculture: a survey. Comput. Electron. Agric. 147:70-90.
- Keshari R., Vatsa M., Singh R., Noore A. 2018. Learning structure and strength of CNN filters for small sample size training. pp 9349-58 in Proc. IEEE Conf. Comput. Vis. Pattern Recog., Utah, USA.
- Krizhevsky A., Sutskever I., Hinton G.E. 2012. ImageNet classification with deep convolutional neural networks. Adv. Neural Inf. Process. Syst. 25:1-9.
- LeCun Y., Bottou L., Bengio Y., Haffner P. 1998. Gradient-based learning applied to document recognition. Proc. IEEE. 86:2278-342.
- Lichtenthaler H.K. 1987. Chlorophylls and carotenoids: pigments of photosynthetic biomembranes. Methods Enzymol. 148:350-82.
- Mohanty S.P., Hughes D.P., Salathé M. 2016. Using deep learning for image-based plant disease detection. Front. Plant Sci. 7:1-10.
- Preethy. 2014. Characterization of betel vine (*Piper betle* L.) types of Malappuram District. Degree Diss., Kerala Agricultural University, India.
- Rosmala A., Khumaida N., Sukma D. 2016. Alteration of leaf anatomy of handeuleum (*Graptophyllum pictum* L. Griff) due to gamma irradiation. HAYATI J. Biosci. 23:138-42.
- Sabharwal S., Sudan S., Ranjan V. 2013. *Jasminum sambac* linn (motia): a review. Int. J. Pharm. Res. Bio-Sci. 2:108-30.
- Simonyan K., Zisserman A. 2015. Very deep convolutional networks for large-scale image recognition. pp 1-14 in: Proc. Int. Conf. Learn. Rep., San Diego, USA.
- Sims D.A., Gamon J.A. 2002. Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. Remote Sens. Environ. 81:337-54.
- Too E.C., Yujian L., Njuki S., Yingchun L. 2018. A comparative study of fine-tuning deep learning models for plant disease identification. Comput. Electron. Agric. 161:272-9.
- Truong T.D., Nguyen V.T., Tran M.T, 2018. Lightweight deep convolutional network for tiny object recognition. pp 675-82 in Proc. 7th Int. Conf. Pattern Recog. App. Methods, Madeira, Portugal.
- Ubbens J.R., Stavness I. 2017. Deep plant phenomics: a deep learning platform for complex plant phenotyping tasks. Front. Plant Sci. 8:1190.
- Wang J., Perez L. 2017. The effectiveness of data augmentation in image classification using deep learning. Available from: <https://arxiv.org/pdf/1712.04621.pdf>