AGRIVITA Journal of Agricultural Science. 2021. 43(3): 589-610

AGRIVITA
Journal of Agricultural Science
www.agrivita.ub.ac.id

Non-destructive *in Situ* Measurement of Aquaponic Lettuce Leaf Photosynthetic Pigments and Nutrient Concentration Using Hybrid Genetic Programming

Ronnie S. Concepcion II[1*], Elmer P. Dadios[1] and Joel Cuello[2]

[1] De La Salle University, Manila, Philippines
[2] University of Arizona, Tucson, USA

ARTICLE INFO

Keywords:
Computer vision
Leaf nutrient level prediction
Leaf pigment prediction
Lettuce
Machine learning

Article History:
Received: March 6, 2021
Accepted: August 16, 2021

*) Corresponding author:
E-mail: ronnie_concepcionii@dlsu.edu.ph

ABSTRACT

Phytopigment and nutrient concentration determination normally rely on laboratory chemical analysis. However, non-destructive and onsite measurements are necessary for intelligent closed environment agricultural systems. In this study, the impact of photosynthetic light treatments on aquaponic lettuce leaf canopy (*Lactuca sativa var. Altima*) was evaluated using UV-Vis spectrophotometry (300-800 nm), fourier transform infrared spectroscopy (4000-500 per cm), and the integrated computer vision and computational intelligence. Hybrid decision tree and multigene symbolic regression genetic programming (DT-MSRGP) exhibited the highest predictive accuracies of 80.9%, 89.9%, 83.5%, 85.5%, 81.3%, and 83.4% for chlorophylls a and b, β-carotene, anthocyanin, lutein, and vitamin C concentrations present in lettuce leaf canopy based on spectro-textural-morphological signatures. An increase in β-carotene and anthocyanin concentrations verified that these molecular pigments act as a natural sunscreen to protect lettuce from light stress and an increase in chlorophylls a and b ratio in the white light treatment corresponds to reduced emphasis on photon energy absorbance in chloroplast photosystem II. Red-blue light induces chlorophyll b concentration while white light promotes all other pigments and vitamin C. It was confirmed that the use of the DT-MSRGP model is essential as the concentration of phytopigment and nutrients significantly change during the head development and harvest stages.

INTRODUCTION

Lettuce (*Lactuca sativa* L.) leaf is a strong natural indicator of the apparent health and quality of its whole food-making system (Sublett, Casey Barickman, & Sams, 2018). Its leaf anatomy consists of a waxy layer of cuticle, epidermis, palisade mesophyll, spongy mesophyll, vascular tissue, and the stoma. Chloroplast cells exist in palisade and spongy mesophyll layers. The phytopigments or biochromes have different concentrations of chlorophyll, carotene, anthocyanin, and lutein molecules resulting in differing light reflectances on the outside architecture of the leaf when light hits the leaf surface. Chlorophyll, being the primary light-harvesting pigment in the chloroplast photosystems, reflects a 550 nm wavelength of light that makes it easily visible by human eyes (Li, Scales, Blankenship, Willows, & Chen, 2012; Ritchie, 2006; 2008). The higher concentration of chlorophyll a (α-Chlorophyll, $C_{55}H_{72}MgN_4O_5$) and chlorophyll b (β-Chlorophyll, $C_{55}H_{70}MgN_4O_6$) can be observed in the water-plastoquinone oxidoreductase (photosystem II, PSII) that lies on the inner layer of the thylakoid membrane of leaves (Concepcion et al., 2020). Additional chlorophyll a molecule excitation in the 660 nm wavelength enables PSII. The only accessory pigment residing in the plastocyanin-ferredoxin oxidoreductase (photosystem I, PSI) layer is the group of carotenes. Beta-carotene (β-carotene, $C_{40}H_{56}$) gives the

ISSN: 0126-0537 Accredited First Grade by Ministry of Research, Technology and Higher Education of The Republic of Indonesia, Decree No: 30/E/KPT/2018

Cite this as: Concepcion II, R. S., Dadios, E. P., & Cuello, J. (2021). Non-destructive *in situ* measurement of aquaponic lettuce leaf photosynthetic pigments and nutrient concentrations using hybrid genetic programming. *AGRIVITA Journal of Agricultural Science, 43*(3), 589–610. https://doi.org/10.17503/agrivita.v43i3.2961

yellow-orange and red-orange colors to leaves. On the other hand, anthocyanin ($C_{15}H_{11}O$) and lutein ($C_{40}H_{56}O_2$, ε-carotene) coexist with chlorophylls in PSII as accessory pigments. Anthocyanin is a flavonoid that gives the red to purple color on the leaf. Lutein is a xanthophyll related to beta-carotene that absorbs blue light and appears in yellow in low concentration. Moreover, not only pigments are present on leaf tissue rather it also includes phytonutrients. Vitamin C (ascorbic acid, $C_6H_8O_6$) is a water-soluble organic compound synthesized by plants; however, it is very sensitive to the induced increase in heat and light intensity. Even green light (480 to 560 nm) induces photosynthesis (Liu, Fu, Wang, & Liu, 2016). Thus, both plant nutrients and pigments are highly responsive to light treatments, and their consequent growth and quality are reflected by its canopy traits.

There is a study that conducted the combined effect of adjusting red and blue light intensity and nutrient solution control to the growth and quality of lettuce confirming that intense illumination and low nitrate in hydroponic tank results in high vitamin C yield in lettuce leaves (Song et al., 2020). Customary effective techniques in measuring phytopigments are through colorimetry (Henderson, 2015; Song et al., 2020), hyperspectral reflectance imaging (Pacumbaba & Beyl, 2011; Steidle Neto et al., 2016), chlorophyll fluorescence imaging (Hägele et al., 2016), artificial neural network (ANN) with RGB inputs (Odabas, Simsek, Lee, & İseri, 2017), and hybrid computer vision and paper chromatography (Concepcion et al., 2020). Chlorophyll normalized vegetation index (CNDVI) through the Miniature Leaf Spectrometer CI-710 is a good metric to measure the total chlorophyll concentration (Alsiņa, Dūma, Dubova, Šenberga, & Daģis, 2016). For phytochemical measurement, Coomassie brilliant blue G-250 dye method, UV spectrophotometry (Gholami, Rahemi, Kholdebarin, & Rastegar, 2012; Kapur et al., 2012), high-performance liquid chromatography (HPLC) (Tarrago-Trani, Phillips, & Cotty, 2012; Thrane et al., 2015; Zeng, 2013), and 2, 6-dichlorophenol indophenol titration (Song et al., 2020) are the conventional laboratory methods. Tillman's method and multiproduct calibration technique were also used to estimate Vitamin C concentration (Santos, Lima, Març o, & Valderrama, 2016). Likewise, iodine concentration was correlated to determine Vitamin C (Olotu et al., 2020). The pH differential method (Lee, Durst, & Wrolstad, 2005) and RGB color information (Yang et al., 2016) are also considered in estimating anthocyanin concentration. Hence, the challenges in the destructive determination of plant leaf chemical and pigment concentrations through laboratory techniques are the major fallback in real-time monitoring in precision agriculture. It delivers very accurate measurements but is not absolutely practical for repeated checking.

Despite the above mentioned advances in plant nutrient and pigment bioassay, destructive means such as vibration spectrophotometry and the use of reagents are still the common practice that deliberately stretches time during testing. Acquisition of such expensive state-of-the-art machines is also considered a major problem for startup agricultural industry especially in the Philippines which is a developing country. The need for reducing the cost of determining plant nutrients and pigments for phenotyping and farm management is still an open research area. Other spectral spaces have not been utilized in vegetation segmentation and feature extractions for predicting leaf nutrients and pigment concentrations. As of this writing, the interaction of light treatments with loose leaf lettuce nutrient and pigment concentrations is not yet comprehensively explored and detailed out.

The objective of this study is to develop a novel *in situ* non-destructive approach in determining the nutrient and pigment concentrations of aquaponic lettuce (*Lactuca sativa var. Altima*) leaf canopy through integrated computer vision and computational intelligence. The spectro-textural-morphological signatures of leaf canopy were extracted through computer vision processing. The suitability of generalized processing regression, support vector machine, regression tree, adaptive neuro-fuzzy inference system, and multigene genetic symbolic regression programming was explored in predicting the corresponding leaf nutrient and pigment concentrations. It was contrasted with the image-based deep transfer learning of MobileNetV2, ResNet101, and InceptionV3. This study focused only on analyzing lettuce crops that are on the head development and harvest stages due to a very small amount of nutrients and pigments that can be extracted from vegetative lettuces. The developed approach and computational model enable the low-income farmers, agriculturists, and scientists to predict vitamin C, chlorophyll a, chlorophyll b, β-carotene, anthocyanin, and lutein concentrations on lettuce leaf canopy by using only a consumer-grade camera and laptop computer embedded with the predictive computational intelligence model. The developed model in this study is part of the 10-module vision-phenotype-lettuce (VIPHLET) collective model used for adaptive nutrient management in a lettuce smart farm.

MATERIALS AND METHODS

The lettuce leaf canopy sample is the input to this model (Fig. 1). As the predictive machine learning model requires a lettuce image as input, lettuce was individually set to the canopy imaging system. It is followed by the application of computer vision that handles the spectro-textural-morphological feature extraction and selection. This study explored both the feature-based machine learning and deep image transfer networks in predicting lettuce leaf nutrient and pigment concentrations based on the ground truth generated using UV-Vis spectrophotometric approach (Fig. 1). Fourier transform infrared spectroscopy was used to provide additional organic chemical information regarding the composition of the lettuce pigments. Throughout this study, MATLAB R2020a is the sole computer vision and computational intelligence platform used.

Plant Material and Cultivation Condition

Loose-leaf lettuce (*Lactuca sativa var. Altima*) is the chosen cultivar because it is the main crop being cultivated in Morong, Rizal, Philippines (14.5181° N and 121.2390° E). The cultivation experiment was performed last September to October 2020 following the 6-week crop life cycle for lettuce. All seeds undergo two weeks of cold stratification in chiller temperature. The aquaponic cultivation setup is composed of two subsystems namely aquaculture and hydroponic subsystems. The aquaculture subsystem is the source of nutrients for crop intake through the cyclic effluents of 20 red tilapia (*Oreochromis niloticus* Linnaeus). An aerator was placed inside the aquaculture tank to provide enough dissolved oxygen to the water. The hydroponic subsystem design is nutrient film technique (NFT) using 3-inch food-grade polyvinyl chloride (PVC) pipes as the main water channel with 6 inches hole-to-hole separation for lettuce crops. Each hole has a 1.5-inch diameter that holds the plastic cup with 1 square inch rockwool. The water nitrate, phosphate, and potassium concentrations were maintained at 124.70 mg/l, 65.23 mg/l, and 168.74 mg/l from germination to harvesting stage (Concepcion II & Dadios, 2021) by adding fractions of simple nutrient addition program (SNAP) solution (Institute of Plant Breeding, University of the Philippines Los Baños) and freshwater through tempering approach. Two artificial light treatments were established, namely red-blue (RB) and white (W) lights, to induce differences in the growth and quality of lettuce (Table 1). There are 60 lettuces grown in red and blue light treatment, and 60 lettuces are exposed to white light. To ensure that all plots will have growing lettuce for the experiment, four lettuce seeds (Ramgo International Corporation, Philippines) were sown on each rockwool. A total of 480 seeds was used in the whole experiment. These artificial photosynthetic light sources utilized 60 pieces of red-blue LED strips and two T8 full spectra LED lamps, respectively. An Arduino ESP32 microcontroller equipped with a real-time clock (RTC) module activates the lighting system for 16 hours per day (6:00 a.m. to 10:00 p.m.) and enables dark hours for 8 hours (10:00 p.m. to 6:00 a.m.). One 5 V DC brushless exhaust fan was installed on each layer on the hydroponic bed to circulate properly the air and maintain a balance of ambient temperature.

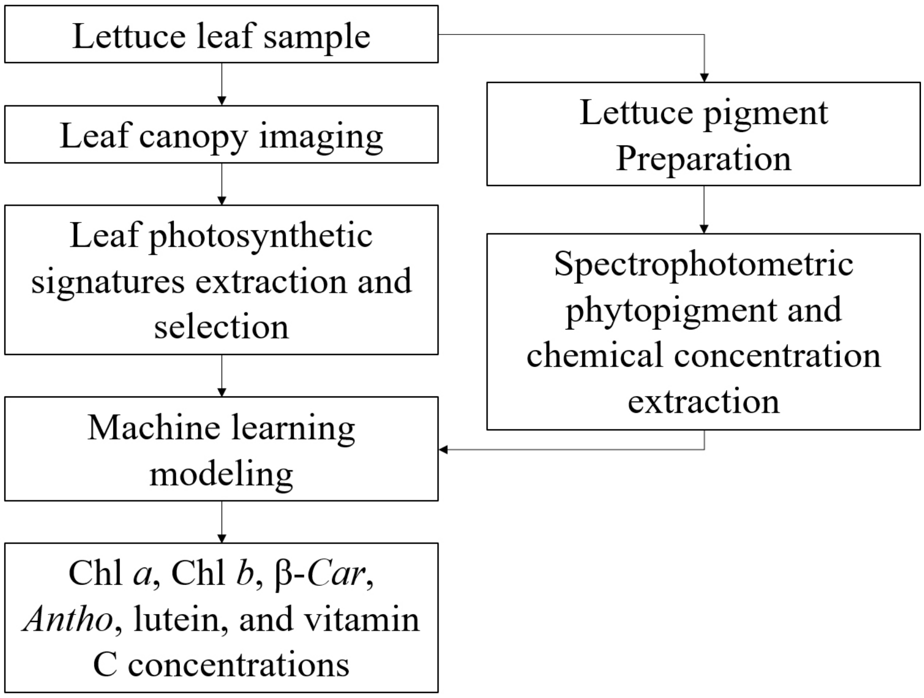

Fig. 1. Overall system architecture of the development of *in situ* measurement of *Lactuca sativa* leaf phytopigments and nutrients as impacted by white and red-blue light treatments using integrated computer vision and computational intelligence

Table 1. Artificial photosynthetic light treatment used in cultivating green loose-leaf lettuces from germination to harvest stage

Light treatment	Number of planted lettuces	Cultivation weeks	Total captured images
Red and blue spectrum	30	5*	30
Red and blue spectrum	30	6^+	30
White spectrum	30	5*	30
White spectrum	30	6^+	30

Remarks: * Means head development stage and $^+$ means harvest stage

Lettuce Pigment Preparation and UV-Vis Spectrophotometric Assay of Phytopigments and Nutrient Concentration

Leaf tissue samples were peeled from the abaxial surface far from the midrib area. Combination of light and dark-pigmented leaf tissues totaling 0.5 g was collected from the topmost leaf canopy for each lettuce. This avoids the local selection resulting to bias leaf extract. 30 lettuces from each light treatment in the head development and harvest stages were utilized in this experimental phase. The leaf tissues were furtherly cut into small pieces using alcohol-cleansed scissors to break the thylakoid membrane and allows 75% acetone (C_3H_6O) to easily penetrate the palisade and spongy mesophyll layers where the chloroplast cells reside. The ratio of leaf tissue and solvent extract is 0.5 g: 15 ml. Using a mini-spatula, leaf tissues were set to the bottom of the borosilicate test tubes with lids and the solvent is added into it. The test tubes were submerged to 65°C thermostatic bath incubation for 30 minutes of incubation. Lettuce leaf extracts were cooled down at room temperature before processing in a spectrophotometer.

Using 1 cm wide quartz cuvette, it was filled with 3.5 ml of leaf extract and set to 300 to 800 nm wavelength sweeping for absorbance reading using Shimadzu UV-Vis 1900 spectrophotometer. This step is necessary to determine the peak absorbance value of the lettuce leaf extract that will verify the presence of readable pigment concentration. The concentration (mg/g) of included leaf pigments namely chlorophyll a, chlorophyll b, β-carotene, anthocyanin, and lutein were spectrophotometrically determined using optical densities (OD) of 480 nm, 495 nm, 663 nm, 645 nm, and 700 nm according to the procedure described by Ben Ghnaya, Charles, Hourmant, Ben Hamida, & Branchard (2007) and Gholami, Rahemi, Kholdebarin, & Rastegar (2012) and the adapted pigment equations (1-4). Anthocyanin pigment concentration (mg/g) that is based on cyanidin-3-glucoside extraction was adapted from the procedure done by Lee, Durst, & Wrolstad (2005) and Steidle Neto et al. (2016) which is the pH-differential method that dilutes 0.1% of hydrochloric acid (HCL) to every 1 ml of leaf extract. Potassium chloride (KCl) and sodium acetate ($C_2H_3NaO_2$) buffers for pH 1.0 and 4.5, respectively, were diluted to leaf extracts before 30 minutes of centrifugation, then, samples undergo spectrophotometric measurement in the 520 nm and 700 nm wavelengths (5). The concentration of vitamin C (mg/g) is determined based on the procedure of Kapur et al. (2012) that homogenizes the leaf extract sample with 50 g of metaphosphoric acid (HPO_3) and 40 ml of the glacial acetic acid solution, spectrophotometric reading at 521 nm, and by adapting the chemical equation (6). $V_{solvent}$ is the volume of acetone solvent, $m_{freshleaf}$ is the mass of fresh leaf tissue, MW is the molecular weight of cyanidin (449.2 g/mol), DF is the dilution factor, and ε is the molecular extinction coefficient of 26,900 l/mol/cm. The chemical assays of extracted leaf pigments were repeated three times.

$$\mathbf{Chl}\ \boldsymbol{a} = \left[\frac{(12.7\ OD_{663} - 2.69\ OD_{645}) * V_{solvent}}{m_{freshleaf}}\right] \tag{1}$$

$$\mathbf{Chl}\ \boldsymbol{b} = \left[\frac{(22.9\ OD_{645} - 4.86\ OD_{663}) * V_{solvent}}{m_{freshleaf}}\right] \tag{2}$$

$$\boldsymbol{\beta}\ \mathbf{Car} = \left[\frac{(17.16\ OD_{495} - 3.96\ OD_{480}) * V_{solvent}}{m_{freshleaf}}\right] \tag{3}$$

$$\mathbf{Lutein} = \left[\frac{(11.5\ OD_{480} - 20.61\ OD_{495}) * V_{solvent}}{m_{freshleaf}}\right] \tag{4}$$

$$\mathbf{Anth} = \left[\frac{\left(\frac{\left((OD_{520} - OD_{700})_{pH1.0} - (OD_{520} - OD_{700})_{pH4.5}\right) * MW * DF}{\varepsilon}\right) * V_{solvent}}{m_{freshleaf}}\right] \tag{5}$$

$$\mathbf{Vit.\ C} = \left[\frac{(26.9542\ OD_{521} - 0.2183) * V_{solvent}}{m_{freshleaf}}\right] \tag{6}$$

Fourier Transform Infrared Spectroscopic Chemical Composition Determination

To obtain the organic chemical information present in the extracted leaf extracts of lettuce in the head development and harvest stages of red and blue and white light treatments, Fourier transform infrared spectroscopy (FTIR) was performed. The leaf extracts before adding the reagents were separately air-dried on a ceramic container. The residues were scraped from the container and mixed with potassium bromide (KBr) salt in a 1:10 (residue to KBr) ratio. The solid mixture was homogenized using mortar and pestle. The translucent disc of KBr pellet was generated using the 5000 lb/in^2 press. This pelleting technique is more appropriate than testing the liquid samples in FTIR because the leaf samples have a minute amount of water that can disrupt the readings. Then, the KBr pellet with leaf extract residues was transferred to the Thermo Scientific Nicolet 470 FTIR spectroscope for transmittance readings in the 4000 to 500 per cm wavenumber with 0.09 per cm resolution.

Leaf Canopy Imaging System

An IP Logitech Camera 1080p was installed at the top of a 7 in x 7 in x 10 in rectangular transparent acrylic platform in a downward (negative x-axis) direction in capturing the whole leaf canopy (Fig. 2). The camera is wired to the Acer Nitro 5 laptop with core i7 8th generation processor for camera control using MATLAB 2020a. Captured images were set to 3000 x 3000 pixels (aspect ratio of 1:1). Textured red cloth was used as the non-vegetative background of lettuce to increase the highlight of green (vegetative) and red (non-vegetative) pixels (Fig. 2). The imaging system is exposed to ambient white light during capturing. All lettuce of light treatments undergoes image capturing before it is utilized for destructive testing in spectrophotometry.

Leaf Photosynthetic Signature Extraction and Selection

Vegetation segmentation is employed using CIELab thresholding. It starts with transforming the RGB digital images to the scientific CIELab color space. It is proceeded by defining the channel thresholds as follows: 1.344 to 100 for lightness (L), -30.754 to 8.114 for the green-red component (a*), and -9.1 to 51.157 for blue-yellow component (b*). Pixel values outside these thresholds are considered as non-vegetative and removed from the image. From its binarized form, small blobs less than 150 pixels were also removed to ensure that only the region of interest (ROI) is retained in the resulting image. Lastly, the image is masked into its original RGB image resulting in a segmented lettuce leaf. Three groups of features were extracted from the pre-processed images namely spectral, textural, and morphological. Spectral features include RGB, HSV, L*a*b*, and YCbCr color components of the segmented lettuce leaf. Texture features include the combined five Haralick and gray level co-occurrence matrix (GLCM) components of contrast, correlation, energy, entropy, and homogeneity that were extracted from the gray and binarized transformed images. For the morphological features, only the leaf area was included. Overall, there are 18 spectro-textural-morphological features to be used as predictors in the predictive models. To develop the least computational cost computational intelligence model, feature selection through a decision tree was employed. The 18-feature predictors were set as input to the regression tree (fitrtree) model for chlorophyll a, chlorophyll b, β-carotene, anthocyanin, lutein, and vitamin C, which means for every phytonutrient and pigment there are different selected most significant predictors. To further alleviate the error in the predictor, outlier removal using the Hampel identifier algorithm was employed. A Hampel filters the input predictors in detecting and removing outliers from that dataset. For each row of combined 18 predictors, the Hampel algorithm computes the median and standard deviation of the 10 samples from each side, and if the original sample is more than thrice the computed standard deviation, it will be replaced by the computed median.

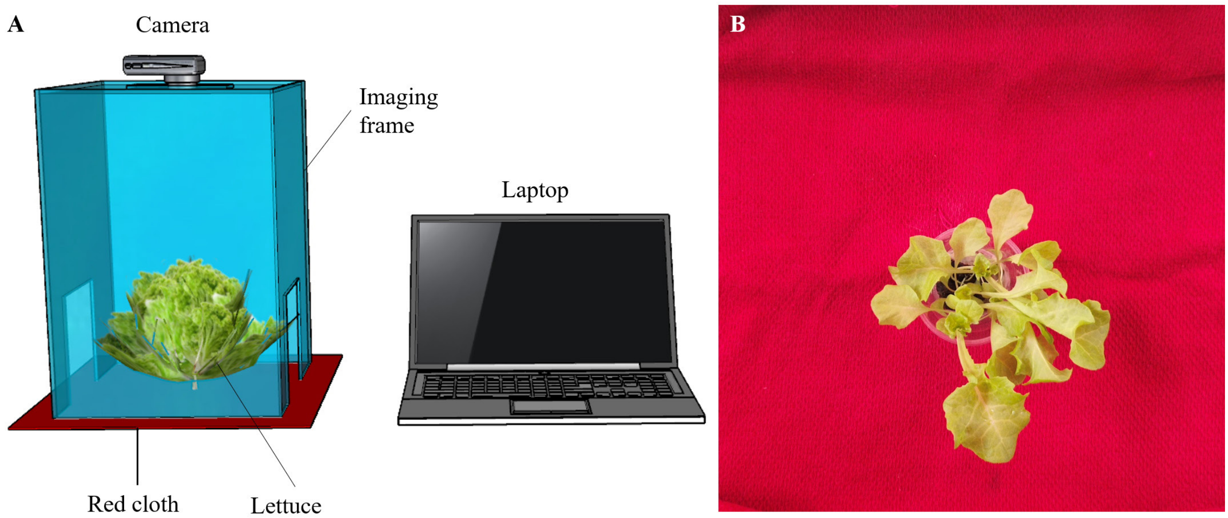

Fig. 2. (A) Image capturing platform composed of a camera, imaging frame, red cloth, and laptop, and (B) sample captured image for lettuce with red background

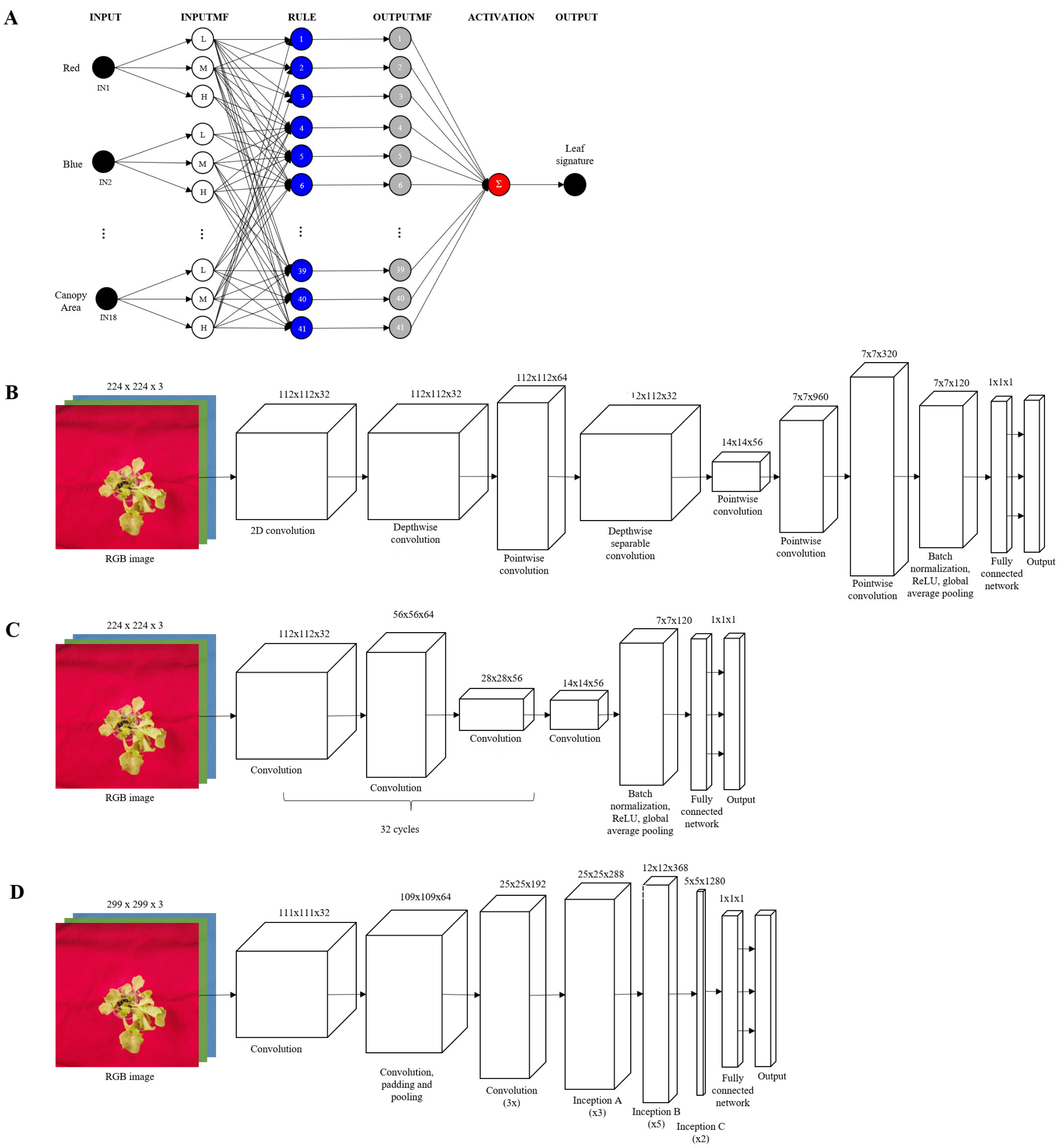

Fig. 3. Model architecture for (A) 18-predictor ANFIS, (B) MobileNetV2, (C) ResNet101, and (D) InceptionV3

Development of Feature-based Machine Learning Prediction Model

Feature-based machine learning is a subset of computational intelligence that trains a predictive model, whether it is for classification or regression, using multiple categorized numerical data as its features (Lauguico, Concepcion II, Alejandrino, Tobias, & Dadios, 2020; Loresco & Dadios, 2018). The features tested in the study were generalized processing regression (GPR), regression-based support vector machine (RSVM), regression tree (RTree), adaptive neuro-fuzzy inference system (ANFIS), and multigene symbolic regression genetic programming (MSRGP). Only the best configuration of each model was discussed here. These models were constructed and optimized twice. The first cycle is by using the 18-feature predictor set as input to the models and the second cycle uses the selected or most significant features only. The GPR model was configured by using squared exponential kernel function, constant basis function with beta of 204.6906 and sigma of 70.0387, exact predict method, an exact fit method for an active set size of 60, and random active set method. The RSVM model was configured by using a bias of $8.0197e^3$, sequential minimal optimization (SMO) as the solver, box constraint of 0.073044, kernel scale of 0.0066897, and epsilon value of 22.469. The RTree model was configured using 11 minimum leaf size. The ANFIS model was configured using 0.25 cluster influence for the whole predictor set, error goal of 0, initial step size of 0.01, step size decrease rate of 0.5, step size increase rate of 1.5, and epoch number of 100. The sample ANFIS architecture using 18-feature vectors generated 41 rules in predicting a single output parameter (Fig. 3A). The hyper parameters of the above mentioned feature-based machine learning models were optimized using Bayesian algorithm. Among these 5 feature-based machine learning models, only the MRSGP is symbolic which generates the mathematical equation of the model. The MSRGP model was configured using population size of 100, generation size of 100, tournament size of 50, pareto rate of 0.2, elite fraction of 0.1, maximum genes of 10, maximum tree depth of 5, and mathematical function nodes of {'times', 'minus', 'plus', 'sqrt', 'square', 'sin', 'cos', 'add3', 'mult3', 'log', 'cube', 'neg', 'abs'}. Stratified sampling was employed by dividing the entire data set to 56% for training, 24% for validation, and 20% for testing. There is no the same data row was used for each modeling phase to ensure no over fitting will happen.

Development of Image-based Deep Transfer Network Prediction Model

In contrast with the feature-based machine learning models, deep-transfer networks are convolutional neural network (CNN) rooted predictive models directly utilizing images as inputs (Figs. 3B-3D. In this study, the raw RGB images were inputted into the models. MobileNetV2, ResNet101, and InceptionV3 were specially chosen because their average accuracies and disk memory space consumption characteristics lie in the region of optimal balance. Before the features were actually extracted from the images, input images are required to be resized (224 x 224 x 3 pixels for each color component in MobileNetV2 and ResNet101, and 299 x 299 x 3 pixels for InceptionV3). Data augmentation through horizontal reflection, and x and y-axis translation within -30-to-30-pixel range was employed. These pre-trained deep learning models were configured using stochastic gradient descent with momentum (SGDM) as training algorithm, minimum batch size of 20, maximum epochs of 7, initial learn rate of 0.0001, piecewise learn rate schedule, learn rate dop factor of 0.1, learn rate drop period of 20, every-epoch shuffling, and validation frequency of 3. Stratified sampling was employed by dividing the entire image set to 56% for training, 24% for validation, and 20% for testing.

Model Evaluation Metrics and Statistical Analysis

The developed feature-based and image-based predictive models were evaluated using root mean square error (RMSE), R^2, and mean absolute error (MAE) by comparing the ground truth and predicted values from training, validation, and testing stages. Inference time was also included to determine if reducing input features can increase the computational cost in terms of time. The biological data were statistically processed through Pearson correlation analysis with $\alpha \leq 0.05$ consideration.

RESULTS AND DISCUSSION

Absorbance Characteristics of Extracted Leaf Pigment

The extracted lettuce leaf pigments that underwent ultraviolet to near-infrared wavelength sweeping provide an active absorbance curve from 400 nm to 670 nm region (Fig. 4). However, there is a considerable plateau with very weak absorbance strength in the anthocyanin region from 440 nm to 650 nm. It means that low visible red to violet reflectance can be observed in the abaxial leaf surface samples. Anthocyanin concentration is found more on the lettuces cultivated in the red-blue light treatment and head development lettuce in white light cultivation. This agrees with the fact that only green lettuce variety was used. Absorbance through appeared at 711.5 nm approaching the 800 nm endpoint. There is a significant 0.965 absorbance crest at 663 nm for harvest stage lettuce cultivated in white light (Fig. 4). This corresponds to chlorophyll a molecules that are actively present in the sample as supported by Shah, Buraidah, Teo, Careem, & Arof (2016) that verifying chlorophyll a pigment has peak absorbances across 430 nm and 663 nm light. Moreover, there is a significant 0.78 absorbance crest at 415.5 nm that corresponds to the presence of chlorophyll b molecules in the leaf samples as supported by Shah, Buraidah, Teo, Careem, & Arof (2016) that verifying chlorophyll b pigment has peak absorbances across 400 nm and 650 nm light. At 645 nm, lettuce cultivated at white light has the highest absorbance of 0.334. All extracted lettuce pigments exhibit a distinct concentration of chlorophylls a and b, hence, chlorophyll b as accessory pigment for photosystems I and II transfers the absorbed light energy to chlorophyll *a* through resonance. Since chlorophyll-a is active at the reaction center of the antenna array of the plant leaf system and chlorophyll b is the one that regulates the antenna size, both are considered responsible for energy transfer in anoxygenic photosynthesis. For beta-carotene and lutein, wavelength absorbances at 480 nm and 495 nm crested on lettuces cultivated using red and blue lights (Fig. 4). For anthocyanin, absorbance at 520 nm and 700 nm are most significant on head development lettuce leaf using red-blue light and harvest lettuce leaf using white light respectively. For vitamin C, absorbance at 521 nm is most significant at head development lettuce in the red-blue light treatment.

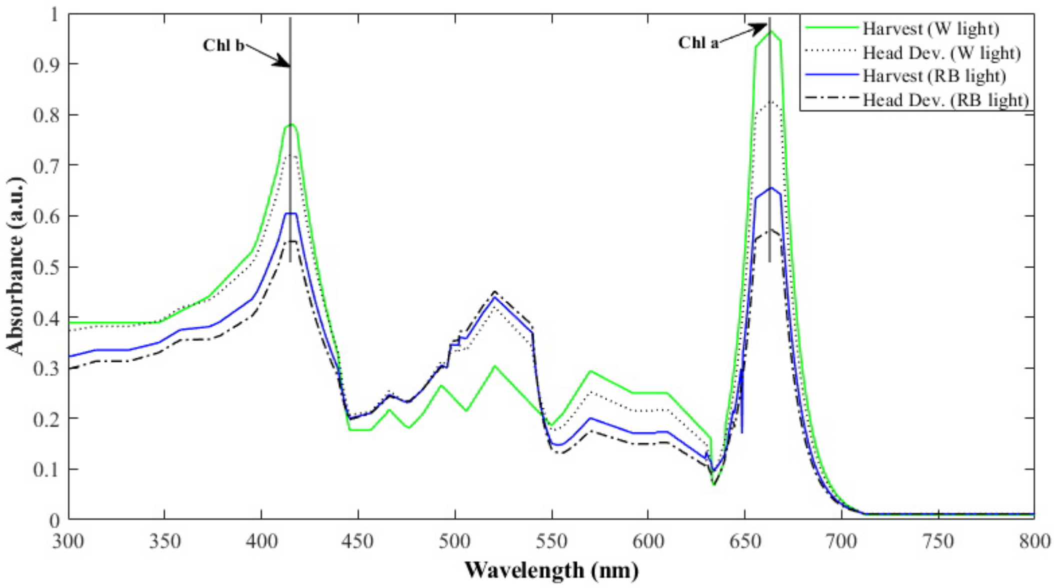

Fig. 4. Absorbance curve of lettuce leaf cultivated in white and red-blue light treatments showing characteristics in 415.5 nm, 480 nm, 495 nm, 520 nm, 521 nm, 645 nm, 663 nm, and 700 nm wavelength. Primary pigments, Chl *a* and Chl *b*, dominant in all leaf extracts are highlighted

Lettuce Leaf Pigment Organic Composition

The organic molecular information of head development and harvest lettuce cultivated in white and red and blue light treatments were verified in the infrared transmittance spectrum from 8000 to 500 wavenumber (Fig. 5). For the head development (WAG 5) green lettuce cultivated in T8 white LED spectrum at 3950 lumens, 3496.87 per cm is attributed to strong and broad O-H stretching with intermolecular bonding. There are one medium C-H bending and one medium C-H bending that appeared in the functional and fingerprint regions respectively. The first one is at 3010.36 per cm and the other is at 1401.09 per cm that is for gem dimethyl alkane molecules vibration. Medium C=C stretching is evident at 1640.85 per cm for disubstituted alkene vibration. There is strong C-O stretching at 1310.77 per cm, 1150.87 per cm, and 1097.27 per cm for the vibration of alkyl aryl ether, a tertiary alcohol, and primary alcohol compounds respectively. A strong C-Br stretching is evident at 745.17 per cm for halo compound vibration. The transmittance of harvest lettuce in white light treatment is significantly the strongest and head development lettuce in red and blue light treatment is the weakest. However, the infrared transmittance characteristic curve for lettuces cultivated in red and blue light treatment is substantially the same as their white light treated lettuce of the same life stage but with diminished strength. All extract leaf pigments contain methyl and aldehyde groups. Based on the generated FTIR spectrum (Fig. 5), the pigments are purely cellulose-based having different concentrations of chlorophylls and other phytopigments which is important in analyzing the light impact on crops. White light induces more chlorophyll on leaf chloroplast photosystems resulting in better crop yield compared to red and blue light treatments.

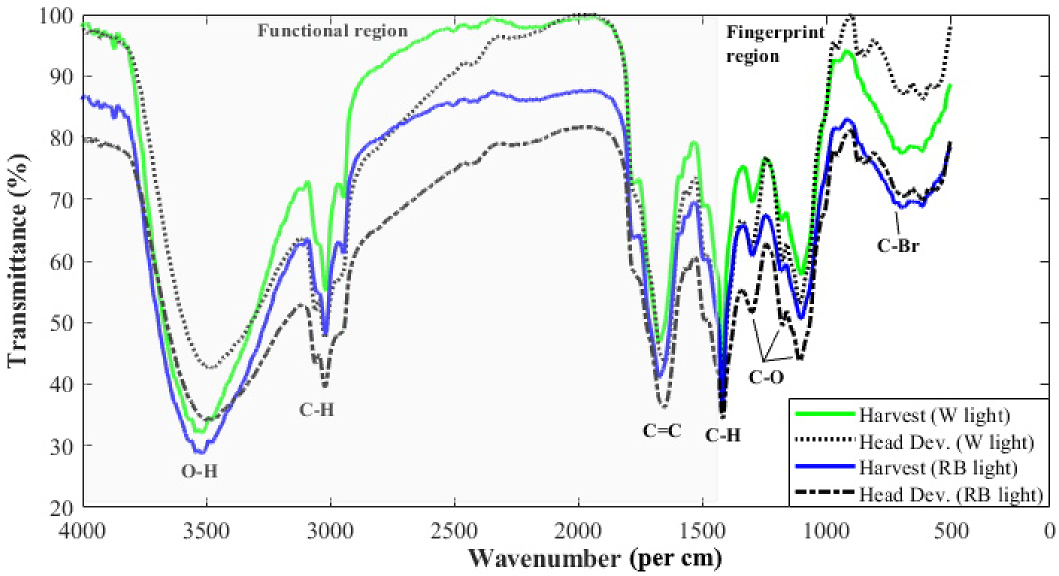

Fig. 5. FTIR spectra of head development and harvest lettuce stages that are exposed in white and red and blue light treatments

Significant Leaf Signatures

By applying a decision tree for selecting highly significant predictors for chlorophyll-a prediction, value, lightness, blue-yellow component, and luma were removed from the list as they were not logically essential in estimating chlorophyll-a (Fig. 6A). Green, blue, value, luma, blue-difference, energy, and canopy area were considered as non-essential predictors in estimating chlorophyll b (Fig. 6A). Green, hue, value, lightness, luma, homogeneity, and canopy area were considered as non-essential predictors in estimating beta-carotene (Fig. 6A). Blue, hue, value, lightness, luma, contrast, energy, homogeneity, and canopy area were considered non-essential predictors in estimating anthocyanin (Fig. 6A). Red, green, value, lightness, luma, blue-difference, and canopy area were considered non-essential predictors in estimating lutein (Fig. 6A). Green, value, lightness, luma, contrast, energy, and homogeneity were considered non-essential predictors in estimating vitamin C (Fig. 6A). Saturation, green-red, red-difference, correlation, and entropy are strongly considered as highly significant and impactful predictors in estimating phytopigments and nutrient concentration of lettuce leaf. On the other hand, value and luma components are completely not selected, and lightness was selected only for estimating chlorophyll b. It appears that brightness in lettuce leaf images is totally not essential and can be eliminated in the list of predictors. Overall, there are only 14, 11, 11, 9, 11, and 11 considerable independent combined predictors from the original 18 features that can sensitively estimate chlorophyll a, chlorophyll b, β-carotene, anthocyanin, lutein, and vitamin C concentrations.

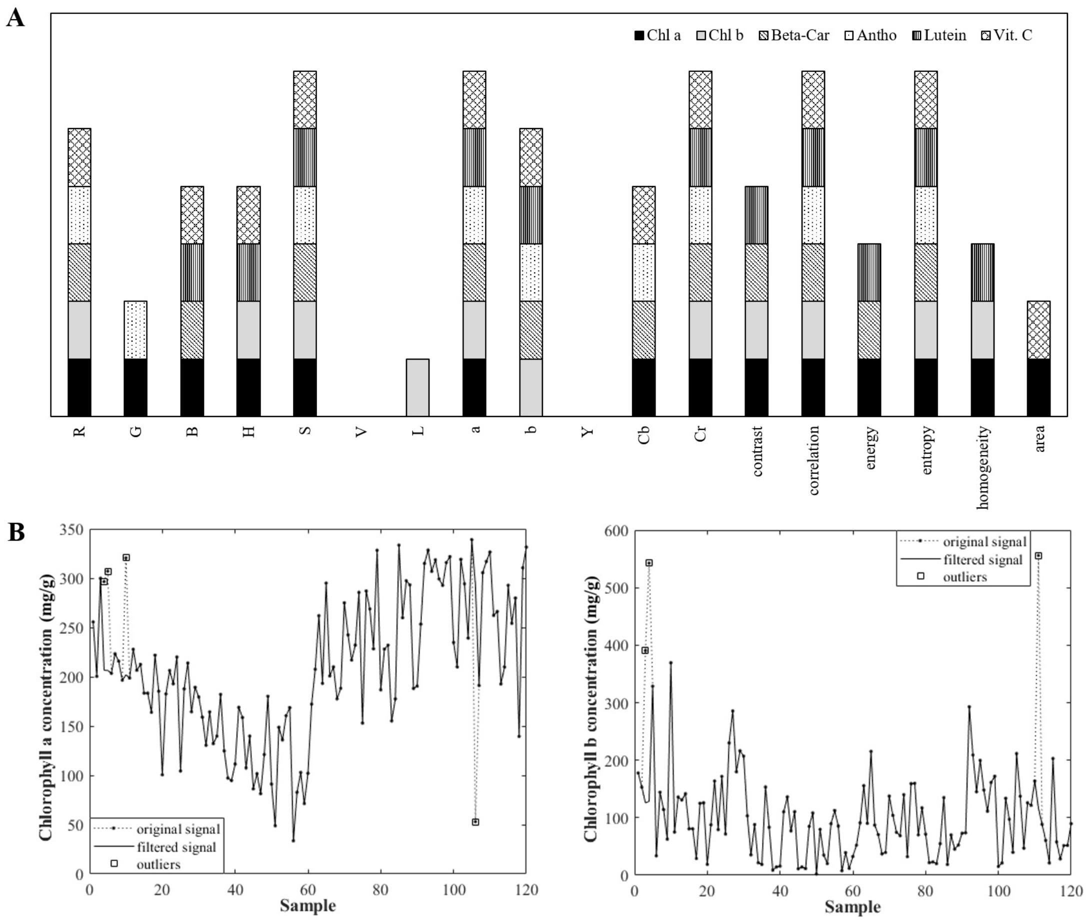

Fig. 6. (A) Distribution of extracted and selected features in predicting chlorophylls *a* and *b*, beta-carotene, anthocyanin, and vitamin C concentrations; (B) sample of Hampel outlier filtering for chlorophyll prediction indicating outliers with square markers

Correlation analysis confirmed that entropy and leaf canopy area are almost perfectly related to each other as they have a strong positive correlation for both head development and harvest stage lettuce in both light treatments (Table 2). Both green-red and blue-yellow components have a strong negative correlation with vitamin C at $\alpha \leq 0.05$ for harvest stage lettuce cultivated in red-blue light and white light respectively. Red and blue-yellow components have a strong positive correlation with beta-carotene and vitamin C for harvest stage lettuce cultivated in red-blue light and a weak positive correlation for head development lettuce treated with white light. Red-difference and correlation components have the same characterization of weak negative correlation to chlorophyll a and lutein concentrations to those head development lettuces grown in both light treatments. These complex correlations to the most significant predictors imply that correct combinations of predictors are necessary to generate acceptable concentration estimations.

Predicted Phytopigments and Nutrient Concentration Using Feature-Based Machine Learning

The generated multigene symbolic regression genetic programming models (Table 3) were explicitly configured for predicting chlorophylls a and b, β-carotene, anthocyanin, lutein, and vitamin C concentrations of lettuce plant leaf canopy in an onsite smart farm situation. The use of feature-based machine learning models namely generalized processing regression, regression support vector machine, regression tree, adaptive neuro-fuzzy inference system, and multigene genetic programming, was successfully employed in predicting phytopigments and nutrient concentrations with the instrumentally measured values based on (1-6) as the target values (Fig. 7). In this discussion, the model is subscripted with the number of features used in that specific model. For example, $ANFIS_{18}$ is the predictive ANFIS model utilizing the original 18 features, and $MSRGP_{14}$ (hybrid DT and MSRGP) is the predictive MSRGP model utilizing the decision tree-selected features. Each hybrid DT and predictive model uses different features (Fig. 6A). For chlorophyll a concentration prediction, $MSRGP_{14}$ bested out other computational intelligence models with an accuracy of 80.9%, RMSE of 58.691, and MAE of 44.656 (Fig. 7A, Table 4). Chlorophyll a concentrations between 100 to 30 mg/g have the fittest prediction. $RTree_{14}$ exhibited the lowest RMSE that is 2.808% lower than $MSRGP_{14}$. In the light of inference time, both $MSRGP_{18}$ and $MSRGP_{14}$ have the fastest prediction duration of 10.20 seconds which is 93.138% faster than $RSVM_{14}$. ANFIS variants exhibited data overfitting as expressed by 100% accuracy from training to validation, and a sudden decrease to 43.8% in the testing phase. For chlorophyll b concentration prediction in lettuce leaf canopy, $MSRGP_{18}$ bested out other computational intelligence models with an accuracy of 89.9%, RMSE of 32.063, and MAE of 23.443 (Fig. 7B, Table 4). Chlorophyll b concentrations between 70 mg/g and 170 mg/g have the most accurate predictions. The DT-MSRGP for chlorophyll b concentration prediction is 7.6% behind in the accuracy of $MSRGP_{18}$, however, both MSRGP variants exhibited the same inference time that is 90.373% faster than the slowest $RSVM_{18}$. $RSVM_{11}$ and $ANFIS_{11}$ yielded the worst accuracy and highest RMSE. For β-carotene concentration prediction in lettuce leaf canopy, $MSRGP_{11}$ bested out other computational intelligence models with an accuracy of 83.5%, RMSE of 52.343, and MAE of 40.467 (Fig. 7C, Table 4). β-carotene concentrations between 100 mg/g and 250 mg/g have the lowest prediction error. $RTree_{11}$ was observed to have acceptable accuracies in the training and validation stages. Despite the model complexity of $MSRGP_{18}$, it exhibited the worst accuracy of 2.2% only and RMSE that is higher to $MSRGP_{11}$ by a factor of 1.782. For anthocyanin concentration prediction in lettuce leaf canopy, $MSRGP_{9}$ bested out other computational intelligence models with an accuracy of 85.5%, RMSE of 0.002, and MAE of 0.001 (Fig. 7D, Table 4). Anthocyanin level above 8 mg/g was observed to have a larger prediction error. $MSRGP_{18}$ has the highest RMSE and $ANFIS_{9}$ has the weakest sensitivity. For lutein concentration prediction in lettuce leaf canopy, $MSRGP_{11}$ bested out other computational intelligence models with an accuracy of 81.3%, RMSE of 56.269, and MAE of 44.317 (Fig. 7E, Table 4). Lutein concentrations in the range of 50 mg/g and 210 mg/g have the fittest predictions. The worst RMSE is exhibited by $MSRGP_{18}$ with a factor of 490.929 higher than the $MSRGP_{11}$. For vitamin C concentration prediction in lettuce leaf canopy, $MSRGP_{11}$ bested out other computational intelligence models with an accuracy of 83.4%, RMSE of 93.494, and MAE of 76.355 (Fig. 7F, Table 4). Vitamin C concentrations from 230 mg/g to 330 mg/g and 630 mg/g to 680 mg/g have the fittest predictions. Among the phytopigments and nutrient concentrations, chlorophyll a and β-carotene concentrations are the slowest and fastest to be predicted with 45.464 seconds and 19.445 seconds respectively. DT-ANFIS models in predicting phytopigments and nutrient concentrations were observed to consistently yield poor sensitivity. On the other hand, DT-MSRGP completely exhibited the least computational cost and the most accurate in predicting selected phytopigment and nutrient concentrations. DT-MSRGP also has the advantage of easy embedding to microcontrollers such as Arduino for real-time and *in situ* evaluation of lettuce leaf canopy given that its model is symbolic (Table 3) which requires no further libraries to be downloaded for it to process. Hence, by comparing the reference and predicted concentrations (Fig. 7), it can be implied that the DT-MSRGP model is highly acceptable for smart farm and precision agriculture application by just using cheap consumer-grade camera.

Table 2. Pearson correlation analysis of lettuce leaf signatures after Hampel outlier filtering

Light treatment	Cultivation stage	N	S Chl b	a* Vit. C	Cr Chl a	Correlation Lutein	Entropy Area	R β-Car	b* Vit. C
Red and blue	Head development	30	0.276	0.147	-0.036	-0.207	0.994	-0.200	-0.201
Red and blue	Harvest	25	-0.224	-0.593	-0.311	0.160	0.999	0.564	0.590
White	Head development	22	0.023	-0.209	-0.162	-0.074	0.986	0.192	0.170
White	Harvest	30	0.132	0.512	0.155	-0.202	0.988	-0.287	-0.560

Remarks: a*: red-green component, b*: yellow-blue component

Table 3. Formulated multigene genetic programming predictive models in estimating chlorophyll *a*, chlorophyll b, β-carotene, anthocyanin, lutein, and vitamin C concentrations (mg/g) with lettuce leaf spectro-textural-morphological features as predictors

Output parameter (mg/g)	Predictive model
Chlorophyll a	$76.2 \tanh(\lvert energy + \cos(G) + \cos(contrast)\rvert) – 34.8 \sin(R + G + area) – 9410 \cos(\sin(R)) – 135 \log(\log(B)) – 4800 \cos(2 \cos(\sin(R))) – 7.81e^5\ contrast + 19.9 \sin(\sin(\sin(area))) + 36.9 \lvert contrast + \cos(B) + \cos(energy)\rvert + 51.9 \cos(\sin(R))^{0.5} \sin(2R + G) + 5.33e^4\ contrast^{0.5} + 6780$
Chlorophyll b	$35.7 \sin(\sin(\sin(Y^2))) – 58 \sin(R^{1.5}b) – 103 \sin(\sin(\sin(\sin(B)))) – 23 \sin(\cos(YCb^{0.5} \sin(R))) + 36.9 \cos(RYCb^{0.5}) + 44.7 \sin(Rb \sin(R)) + 20.6 \cos(YCb^{0.5}(R + Y)^{0.5}) + 17.5 \sin((Cb – area)^2) – 24.6 \sin(R) +500 \sin(\sin(B))^9 + 87.1$
β-carotene	$45.1 \sin(\cos(entropy^4)) – 32.8 \cos(entropy^3) – 891 \cos(B^2\ entropy^2) – 3.69e^4\ energy – 431 \exp(-B^2\ entropy^2) + 1030 \sin(\cos(B^2\ entropy^2)) + 5.38e^4 \tanh(energy)^2 + 174\ energy\ entropy^2 + 0.306 B^2 + 486 \exp(\tanh(entropy)^2) \exp(-B^2\ entropy^2) \cos(B^2\ entropy^2) + 5760$
Anthocyanin	$0.00878\ Cb – 0.00456\ S – 2.79e^{-4}\ G – 0.0308\ entropy – 0.00456 \sin(\cos(G)) + 0.00255 \sin(S – Cb + ab) – 0.00229 \cos(Ga) + 0.0264 \tanh(entropy + 3.91) + 0.00123 \sin(\cos(G) + ab + Cb^2) + 0.734 \sin(S) + 0.0264 \exp(-CbCr) + 0.00338 \lvert\sin(Cb\ entropy)\rvert^2 + 0.00345 R^{0.5} (\cos(G) + R^{0.5})^{0.5} + 0.00439 R^{0.5} – 1.155$
Lutein	$26.1 \cos(2 \lvert b\rvert (Cr + 2\ entropy)) – 11.8 \sin(\lvert b\rvert (b + contrast + entropy)) – 1510 S – 119 \cos(\sin(\lvert b\rvert (Cr + energy + entropy))) – 1510 \log(Cr) – 68.5 (\cos(b) + S\ correlation^2)^6 – 1.28 e^7 (contrast^{0.5} + S\ contrast\ correlation)^3 + 6.95 \log(Cr) \log(entropy)^4 + 57.9\ energy^{12} \sin(\lvert b\rvert (Cr + energy + entropy)) + 3.4e^4\ contrast^{0.5} – 158\ energy^{24} + 7070$
Vitamin C	$8150 \cos(\sin(\cos(\sin(entropy)))) – 59.6 B + 59.6 \cos(a^2) + 48.3 \cos(343 \sin(entropy)(a + b)) + 42.2 \cos(B – \cos(2a)) + 53.2 \cos(686a \sin(entropy)) + 47.3 \sin(\sin(Cb\ area – B^2H)) + 42.3 \sin(\sin(\sin(b)) – B^3) + 516 \cos(a) – 1.07e^5 S^2 + 0.316 aB \sin(\sin(H)(B + Cr + 2 \cos(B) + \cos(a)) – 5490$

Remarks: All independent parameters used in the predictive models are the exact parameters generated in feature extraction and selection stage. R is the red reflectance; G is the green reflectance; B is the blue reflectance; H is hue; S is saturation; a is the green-red component; b is the blue-yellow component; Cb is the blue-difference chroma; Cr is the red-difference chroma; contrast, correlation, entropy, energy, and entropy are the texture features; and area is the morphological 2D surface of leaf canopy

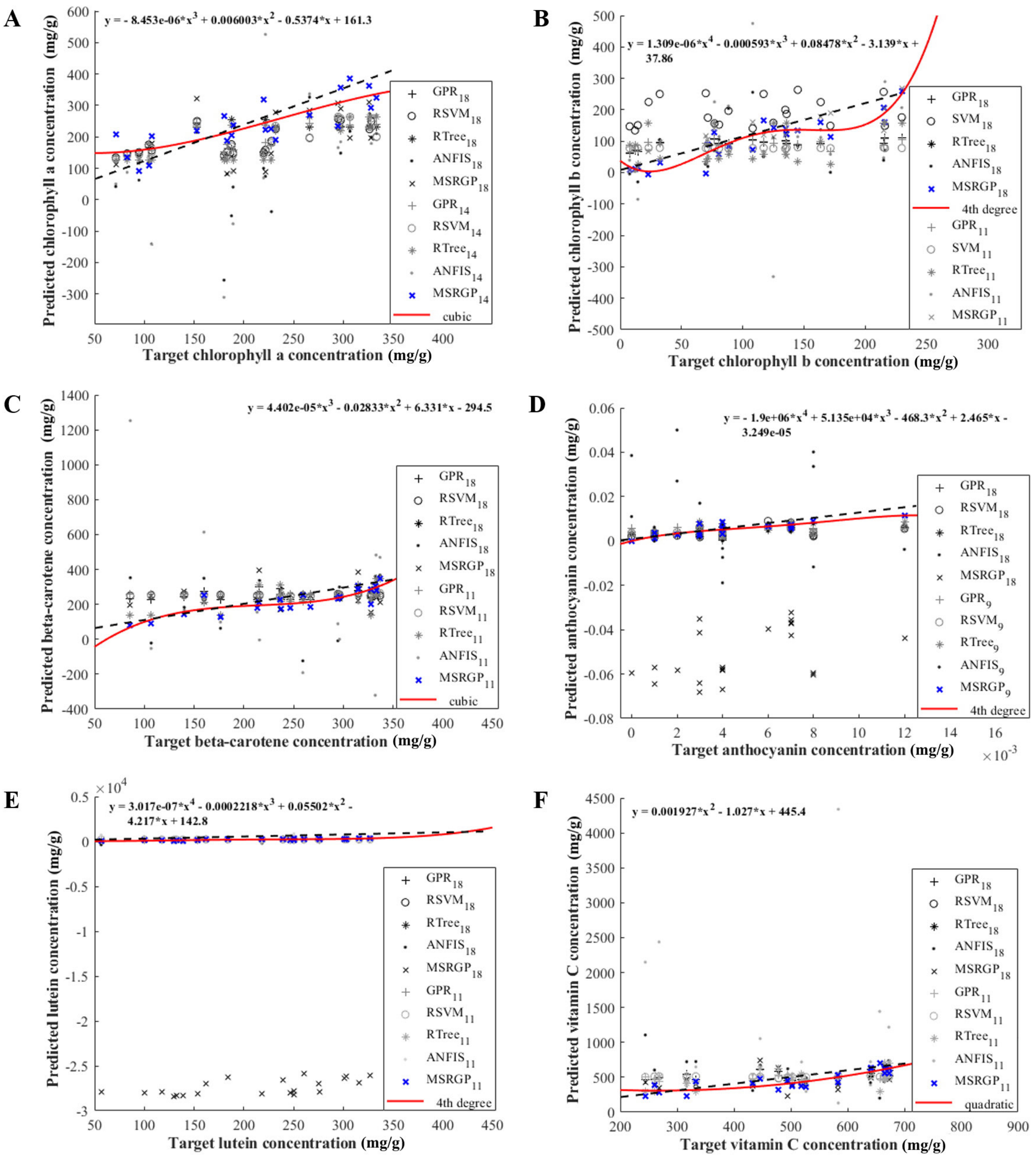

Fig. 7. Prediction curve of the developed feature-based machine learning models in predicting (A-B) chlorophylls a and b, (C) beta-carotene, (D) anthocyanin, (E) lutein, and (F) vitamin C concentrations in lettuce leaf canopy with spectro-textural-morphological leaf signatures as inputs

Table 4. Performance evaluation of developed feature-based machine learning models and image-based deep transfer networks in predicting chlorophylls a and b, β-carotene, anthocyanin, lutein, and vitamin C concentrations of individual lettuce leaf canopy

Model	No. of features	Training			Validation			Testing			
		RMSE	R^2	MAE	RMSE	R^2	MAE	RMSE	R^2	MAE	Inference time (s)
For chlorophyll *a* concentration prediction											
GPR	18	41.415	0.849	34.419	47.218	0.661	37.228	61.782	0.765	55.251	36.273
GPR	14	40.967	0.853	34.090	47.079	0.664	37.164	61.734	0.765	55.256	22.844
RSVM	18	64.053	0.461	49.157	56.335	0.469	40.280	64.613	0.703	57.833	105.952
RSVM	14	64.146	0.460	49.259	56.662	0.474	39.585	62.684	0.726	56.212	148.654
RTree	18	26.827	0.928	21.366	49.033	0.653	41.466	66.414	0.687	60.116	36.273
RTree	14	26.827	0.928	21.366	44.382	0.725	38.200	57.043	0.769	50.679	23.986
ANFIS	18	0.002	1.000	0.002	0.002	1.000	0.002	158.887	0.518	122.176	36.273
ANFIS	14	0.002	1.000	0.002	0.002	1.000	0.002	167.469	0.438	119.139	23.986
MSRGP	18	34.722	0.877	27.351	36.351	0.841	27.726	78.763	0.568	65.198	10.200
MSRGP	14	108.722	0.856	98.480	121.249	0.302	101.317	58.691	0.809	44.656	10.200
MobileNetV2	3.50E+06	43.808	0.746	34.718	61.331	0.647	48.606	87.616	0.498	69.437	462.00
ResNet101	4.46E+07	57.266	0.055	46.110	80.173	0.047	64.554	114.533	0.036	92.220	1327.00
InceptionV3	2.39E+07	42.942	0.233	38.057	60.119	0.202	53.280	85.885	0.156	76.115	2047.00
For chlorophyll *b* concentration prediction											
GPR	18	54.308	0.722	46.527	60.092	0.769	49.175	63.130	0.500	52.283	61.503
GPR	11	56.164	0.688	47.883	62.314	0.738	50.788	65.136	0.405	53.648	22.844
RSVM	18	31.697	0.870	23.775	126.973	0.170	109.759	111.822	0.149	97.299	105.952
RSVM	11	65.392	0.002	52.017	74.240	0.258	55.373	72.733	0.186	58.695	20.305
RTree	18	129.464	0.533	116.072	72.141	0.012	57.732	70.718	0.363	61.118	29.882
RTree	11	31.697	0.870	23.775	32.886	0.887	23.270	70.718	0.363	61.118	23.986
ANFIS	18	0.002	1.000	0.001	0.002	1.000	0.001	82.215	0.401	61.293	29.882
ANFIS	11	0.002	1.000	0.001	0.002	1.000	0.001	147.368	0.240	95.715	23.986
MSRGP	18	35.597	0.833	27.828	41.237	0.811	32.990	32.063	0.899	23.443	10.200
MSRGP	11	47.016	0.706	39.166	44.832	0.773	37.655	39.720	0.823	30.350	10.200
MobileNetV2	3.50E+06	55.151	0.044	30.897	77.212	0.038	43.256	110.303	0.029	61.794	185.000
ResNet101	4.46E+07	57.898	0.033	32.749	81.057	0.028	45.848	115.795	0.022	65.498	507.000
InceptionV3	2.39E+07	57.744	0.140	38.202	80.841	0.121	53.483	115.488	0.093	76.405	572.000
For β-carotene concentration prediction											
GPR	18	48.431	0.624	39.641	57.504	0.598	47.285	71.462	0.354	61.160	17.248
GPR	11	50.990	0.565	41.657	59.980	0.541	48.382	76.897	0.421	63.186	18.772
RSVM	18	57.014	0.349	45.414	68.459	0.215	54.789	74.955	0.263	60.893	26.311
RSVM	11	57.619	0.328	46.811	67.882	0.275	54.868	72.521	0.593	59.139	18.851
RTree	18	25.464	0.908	19.642	44.199	0.793	34.080	64.111	0.603	47.579	29.882
RTree	11	25.464	0.908	19.642	22.460	0.956	18.206	64.111	0.603	47.579	16.752
ANFIS	18	0.003	1.000	0.002	0.003	1.000	0.002	148.566	0.164	107.574	29.882
ANFIS	11	0.003	1.000	0.002	0.003	1.000	0.002	344.331	0.318	215.380	16.752
MSRGP	18	47.958	0.706	38.949	49.948	0.736	39.737	93.295	0.022	78.062	10.000
MSRGP	11	54.381	0.833	47.274	54.258	0.868	46.893	52.343	0.835	40.467	10.000
MobileNetV2	3.50E+06	30.338	0.476	24.279	42.473	0.412	33.991	60.675	0.317	48.559	635.000
ResNet101	4.46E+07	36.029	0.492	29.021	50.441	0.426	40.629	72.059	0.328	58.041	1254.000
InceptionV3	2.39E+07	44.623	0.421	31.388	62.473	0.365	43.943	89.247	0.281	62.775	522.000

Table 4. *continued*

Model	No. of features	Training			Validation			Testing			
		RMSE	R^2	MAE	RMSE	R^2	MAE	RMSE	R^2	MAE	Inference time (s)
For anthocyanin concentration prediction											
GPR	18	0.001	0.991	0.000	0.000	0.999	0.000	0.003	0.319	0.002	21.076
GPR	9	0.001	0.986	0.001	0.000	0.998	0.000	0.003	0.286	0.002	21.415
RSVM	18	0.003	0.453	0.002	0.003	0.392	0.003	0.003	0.502	0.002	103.781
RSVM	9	0.003	0.485	0.002	0.003	0.419	0.002	0.002	0.574	0.002	35.077
RTree	18	0.001	0.872	0.001	0.003	0.578	0.002	0.003	0.496	0.002	18.716
RTree	9	0.001	0.889	0.001	0.001	0.897	0.001	0.002	0.575	0.002	10.547
ANFIS	18	0.000	1.000	0.000	0.000	1.000	0.000	0.016	0.214	0.010	18.716
ANFIS	9	0.000	1.000	0.000	0.000	1.000	0.000	0.011	0.018	0.006	10.547
MSRGP	18	0.055	0.593	0.054	0.054	0.609	0.053	0.057	0.474	0.056	10.200
MSRGP	9	0.001	0.880	0.001	0.001	0.936	0.001	0.002	0.855	0.001	10.200
MobileNetV2	3.50E+06	0.277	0.019	0.231	0.388	0.017	0.323	0.555	0.013	0.462	244.000
ResNet101	4.46E+07	0.263	0.417	0.202	0.368	0.361	0.283	0.526	0.278	0.404	677.000
InceptionV3	2.39E+07	0.159	0.244	0.131	0.223	0.212	0.184	0.318	0.163	0.263	401.000
For lutein concentration prediction											
GPR	18	38.391	0.693	31.070	37.154	0.714	30.250	65.117	0.573	58.475	16.333
GPR	11	39.058	0.683	31.286	37.944	0.690	31.333	65.119	0.584	58.494	29.964
RSVM	18	49.423	0.388	40.681	42.685	0.602	34.879	71.411	0.641	63.662	38.845
RSVM	11	49.603	0.363	41.143	49.152	0.357	41.404	70.806	0.656	63.514	59.044
RTree	18	21.616	0.912	14.743	30.784	0.805	26.404	66.595	0.580	51.461	33.380
RTree	11	21.616	0.912	14.743	27.998	0.839	21.972	67.419	0.568	52.477	24.472
ANFIS	18	0.003	1.000	0.003	0.003	1.000	0.003	155.826	0.603	101.609	33.380
ANFIS	11	0.003	1.000	0.003	0.003	1.000	0.003	214.706	0.048	127.732	24.472
MSRGP	18	$27.4e^3$	0.386	$27.3e^3$	$27.4e^3$	0.385	$27.3e^3$	$27.6e^3$	0.603	$27.6e^3$	10.200
MSRGP	11	120.033	0.787	109.813	128.304	0.700	121.550	56.269	0.813	44.317	10.200
MobileNetV2	3.50E+06	39.798	0.230	30.650	55.717	0.199	42.909	79.596	0.153	61.299	237.000
ResNet101	4.46E+07	63.271	0.367	45.907	88.579	0.318	64.270	126.541	0.244	91.814	481.000
InceptionV3	2.39E+07	59.564	0.059	40.112	83.390	0.051	56.156	119.128	0.040	80.223	547.000
For vitamin C concentration prediction											
GPR	18	98.182	0.618	81.093	116.081	0.583	96.096	135.378	0.351	119.131	23.974
GPR	11	96.661	0.633	79.853	18.737	0.997	15.174	133.811	0.422	118.738	24.239
RSVM	18	115.294	0.348	92.995	137.522	0.199	111.983	142.830	0.182	117.010	25.094
RSVM	11	115.326	0.349	92.899	137.501	0.204	111.950	142.650	0.195	116.948	40.584
RTree	18	50.188	0.912	39.062	122.284	0.488	95.480	122.606	0.596	85.239	23.637
RTree	11	50.188	0.912	39.062	124.984	0.451	100.359	122.606	0.596	85.239	26.552
ANFIS	18	0.003	1.000	0.003	0.003	1.000	0.002	281.271	0.296	201.509	23.637
ANFIS	11	0.003	1.000	0.003	0.003	1.000	0.002	1075.409	0.111	575.835	26.552
MSRGP	18	88.554	0.852	70.823	87.038	0.878	69.371	175.639	0.223	153.733	10.200
MSRGP	11	407.683	0.347	231.015	277.790	0.410	176.305	93.494	0.834	76.355	10.200
MobileNetV2	3.50E+06	69.151	0.415	54.402	96.811	0.360	76.162	138.302	0.277	108.803	201.000
ResNet101	4.46E+07	$2.76e^2$	0.225	$2.68e^2$	$3.87e^2$	0.195	$3.75e^2$	184.107	0.150	150.231	287.000
InceptionV3	2.39E+07	74.924	0.052	58.765	104.894	0.045	82.271	149.848	0.035	117.530	314.000

Predicted Phytopigments and Nutrient Concentration Using Image-Based Deep Transfer Network

Instead of using the extracted features (Fig. 6A), the developed deep transfer networks namely MobileNetV2, ResNet101, and InceptionV3, utilized raw lettuce leaf images (Fig. 3B, 3C, and 3D). MobileNetV2 exhibited the highest accuracy among the selected deep learning models in predicting chlorophyll-*a* and vitamin C concentrations in green lettuce leaf canopy with 49.8% and 27.7% respectively (Fig. 8A and 8F, Table 4). ResNet101 exhibited the highest 32.8%, 27.8%, and 24.4% accuracies in predicting β-carotene, anthocyanin, and lutein concentrations, respectively (Fig. 8C, 8D, and 8F). InceptionV3 has the highest accuracy in predicting chlorophyll b concentration with 9.3% (Fig. 8B). The same with the feature-based machine learning models, predicting chlorophyll-*a* concentration requires longer inference time using deep transfer networks.

Chlorophyll b, on the other hand, is the fastest to evaluate. Moreover, MobileNetV2 having the least features of $3.5x10^6$ has an average of 327.333 seconds inference time making it the fastest among the three deep transfer networks. However, when the yielded model evaluation metrics of deep transfer networks are compared with feature-based machine learning models, it can be deduced that the latter significantly performed better. Even the worst-performing DT-ANFIS machine learning model has higher RMSE and accuracy in each concentration prediction.

After several explorations in estimating phytopigment and nutrient concentrations of lettuce leaf canopy through integrated computer vision and computational intelligence, this study successfully developed a hybrid DT-MSRGP surpassing the sensitivities of GPR, RSVM, RTree, ANFIS, MobileNetV2, ResNet101, and InceptionV3 in evaluating spectro-textural-morphological leaf signatures and raw images. The trend of quantifying phytopigment and nutrient concentrations of leaves is through destructive colorimetry (Song et al., 2020), and spectrophotometry and chromatography (Kapur et al., 2012; Tarrago-Trani, Phillips, & Cotty, 2012). The most related techniques in this study are partial least square regression and quadratic regression using RGB and HSI color spaces (Table 5). The developed DT-MSRGP model completely bested these techniques by Steidle Neto et al. (2016) and Yang et al. (2016) in predicting chlorophylls, beta-carotene, anthocyanin, lutein, and vitamin C concentrations. Though Steidle Neto et al. (2016) has a 9% advantage in the prediction accuracy of anthocyanin, their laboratory technique used 99.5% acetone instead of 75%, and the concentration convention used is mg/kg instead of mg/g. Likewise, the partial least regression was applied to the derived color of leaf extract and not to the raw leaf itself. Thus, the hybrid DT-MSRGP model has the advantage for accurate on-site evaluation of lettuce leaf canopy using an inexpensive digital camera.

Impacts of Lighting to Phytopigments and Nutrients

Plant leaves are effective indicators of the current state of its environment. If the environment is outside the thresholds of suitable cultivation, leaf pigment, and nutrient concentrations manifest it. Interestingly for *Lactuca sativa var. Altima*, white light treatment has a great deal with pigment and nutrient concentrations (Fig. 9). It was observed that chlorophyll b is highly concentrated (average of 111.262 mg/g for WAG 5 and 6) when lettuce is cultivated in the red-blue light spectrum. However, the white light spectrum has induced considerably the chlorophyll a, β-carotene, anthocyanin, lutein, and vitamin C concentrations for WAG 5 and 6 with the average values of 251.418 mg/g, 277.807 mg/g, 0.0067 mg/g, 252.276 mg/g, and 558.950 mg/g, respectively. The highest chlorophyll a, β-carotene, anthocyanin, lutein, and vitamin C concentrations were found in white light-cultivated head development lettuces followed by red-blue light-treated harvest stage lettuces. Most concentrated chlorophyll b was found in red-blue light-cultivated harvest stage lettuces followed by white light-treated head development lettuces. For red-blue light treatment, there is 12.457 mg/g, 15.382 mg/g, 7.583 mg/g, 0.0002, 11.052 mg/g, and 15.784 mg/g increase per day in the concentration of chlorophyll a, chlorophyll b, β-carotene, anthocyanin, lutein, and vitamin C from head development to harvest stage. Conversely, there is a subsequent decline of 5.497 mg/g, 6.419 mg/g, 7.656 mg/g, 0.003 mg/g, 8.288 mg/g, and 15.787 mg/g per day in the concentration of chlorophyll a, chlorophyll b, β-carotene, anthocyanin, lutein, and vitamin C. Hence, it is still conclusive that white light treatment is better than red-blue treatment especially for head development and harvest stage lettuce despite the diminished concentration during WAG 6. It could be an agricultural practice consideration that green loose-leaf lettuce should be harvested during the fifth week after germination to avoid the natural decline of nutrient and pigment concentrations provided that it is cultivated using white light. By using red-blue light, lettuce should be cultivated on the sixth week after germination. In red-blue light treatment for lettuce, there is 164.522 ± 43.598 mg/g, 111.262 ± 53.838 mg/g, 253.275 ± 26.541 mg/g, 0.0031 ± 0.0008 mg/g, 204.515 ± 38.682 mg/g, and 507.574 ± 66.245 mg/g concentration range in chlorophylls a and b, β-carotene, anthocyanin, lutein, and vitamin C. Whereas in white-light treatment for lettuce, there is 251.418 ± 19.239 mg/g, 105.358 ± 22.468 mg/g, 277.807 ± 26.796 mg/g, 0.0067 ± 0.0009 mg/g, 252.276 ± 29.009 mg/g, and 558.950 ± 55.256 mg/g concentration range in chlorophylls *a* and *b*, beta-carotene, anthocyanin, lutein, and vitamin C.

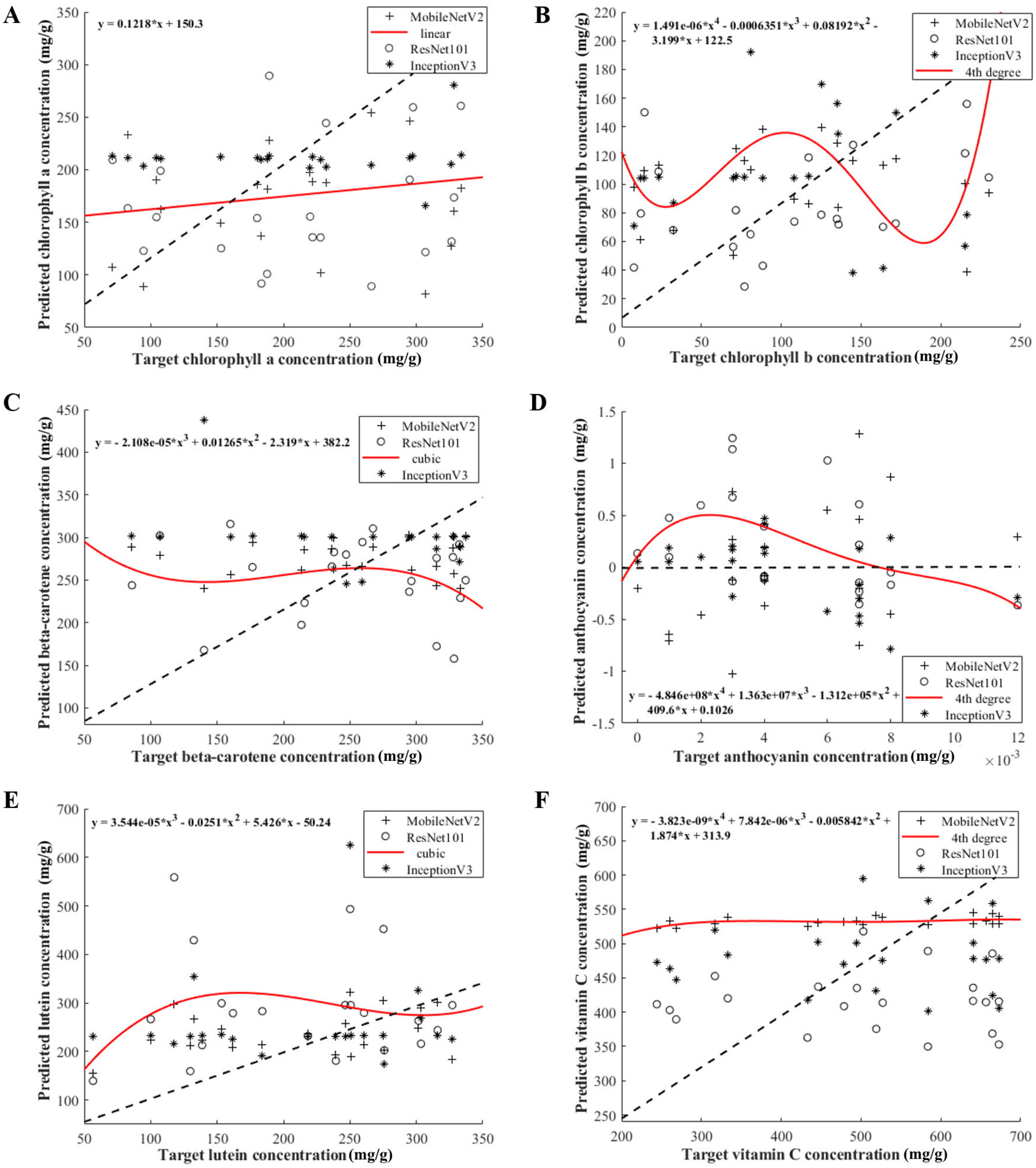

Fig. 8. Prediction curve of the developed image-based deep transfer networks in predicting (A-B) chlorophylls a and b, (C) β-carotene, (D) anthocyanin, (E) lutein, and (F) vitamin C concentrations in lettuce leaf canopy with raw leaf images as input

Table 5. Accuracy of developed predicted models as compared with other techniques in estimating chlorophylls a and b, β-carotene, anthocyanin, lutein, and vitamin C concentrations of leaf canopy

Author and date	Technique/model	Accuracy (%)					
		Chl. a	Chl. b	β-carotene	Anthocyanin	Lutein	Vit. C
Steidle Neto et al., 2016	Partial least square	80.0	80.0	76.0	89.0	*np*	*np*
Yang et al., 2016	Quadratic regression using RGB and HSI color spaces	*np*	*np*	*np*	80.4	*np*	*np*
This study	Hybrid DT-MSRGP based on lettuce leaf signatures	80.9	89.9	83.5	85.5	81.3	83.4

Remarks: *np* means not provided

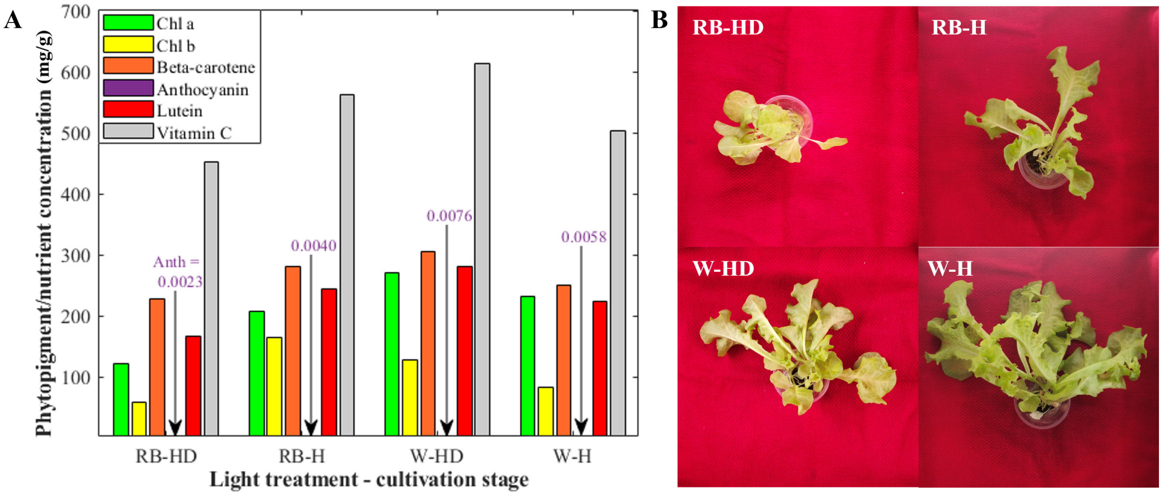

Fig. 9. (a) Impact of light treatment to phytonutrients and pigments in the head development and harvest life stages of green loose-leaf lettuce, and (b) samples of lettuce in each light treatment and cultivation stage. RB is red-blue light; W is white light; HD is head development; H is harvest

Leaf canopy area has an erratic correlation with leaf pigment and nutrient, hence, it is not considered a highly significant leaf feature for prediction. However, the observed canopy area has a weak positive correlation with all pigments and nutrients for white light-treated head development lettuces only. The Chl a/Chl b ratio is found to be 2.106, 1.261, 2.117, and 2.801 for red-blue light-treated head development and harvest lettuces, and white light-treated head development and harvest lettuces, respectively. This chlorophyll ratio signifies that red-blue light-treated harvest lettuces experience drought stress as induced by the light spectrum that is visible to the growth of leaves and root systems. As the Chl a/Chl b ratio is increased by a factor of 1.323 from head development to harvest stage lettuce in the white light treatment, it corresponds to the reduced emphasis on the absorbance of photon energy concerning the photochemistry of chloroplast photosystem II where the main function is the adenosine triphosphate (ATP) synthesis. This claim is supported by Gholami, Rahemi, Kholdebarin, & Rastegar (2012). Both beta-carotene and anthocyanin concentrations of red-blue light treated lettuces increased from head development to harvest confirming that these pigments protect the stressed lettuce as sunscreens. Beta-carotene and anthocyanin also act as active oxygen species (AOS) in stressed lettuce due to exposure to the red-blue light spectrum, and it was previously documented by Gholami, Rahemi, Kholdebarin, & Rastegar (2012). The amount extracted pigment and nutrient concentration in this study is higher than that in the study of Alsiņa, Dūma, Dubova, Šenberga, & Daģis (2016), Kapur et al. (2012), Olotu et al. (2020), and Steidle Neto et al. (2016). This positive deviation is mainly due to the difference in the closed environment agricultural setting. Hence, the cultivated lettuces in both light treatments yielded considerable concentrations of essential phytopigments and nutrients.

CONCLUSION

This study scientifically confirmed that white light promotes a higher concentration of chlorophyll a, β-carotene, anthocyanin, lutein, and vitamin C on head development green loose-leaf lettuce. The red-blue light treatment induces the concentration of chlorophyll b on harvest stage lettuce. The increase in chlorophylls a and b ratio in white light treatment signified the reduced emphasis on photon energy absorbance rate in the chloroplast photosystem II. The increase in beta-carotene and anthocyanin concentration in red-blue light treatment signified that these pigments act as protective molecules for light-stressed lettuce. Hybrid decision tree and multigene symbolic regression genetic programming model exhibited the highest prediction accuracy for all selected phytopigment and nutrient concentrations using the multidimensionally reduced spectro-textural-morphological lettuce leaf canopy signatures. The efficiency and sensitivity of DT-MSRGP enhance the non-destructive onsite evaluation of lettuce leaves. For future studies, it is recommended to employ the same approach for determining macronutrient (nitrate, phosphate, and potassium) concentrations in leaves.

ACKNOWLEDGEMENT

The authors would like to acknowledge the support of the Engineering Research and Development for Technology (ERDT) of the Department of Science and Technology (DOST) of the Philippines, and the Intelligent Systems Laboratory and the Molecular Biology Laboratory of the De La Salle University, Philippines. Also, special gratitude to Sandy Lauguico and Jonnel Alejandrino who helped the main author during the actual laboratory experiments.

REFERENCES

Alsiņa, I., Dūma, M., Dubova, L., Šenberga, A., & Daģis, S. (2016). Comparison of different chlorophylls determination methods for leafy vegetables. *Agronomy Research*, *14*(2), 309–316. Retrieved from https://www.cabdirect.org/cabdirect/abstract/20163195243

Ben Ghnaya, A., Charles, G., Hourmant, A., Ben Hamida, J., & Branchard, M. (2007). Morphological and physiological characteristics of rapeseed plants regenerated in vitro from thin cell layers in the presence of zinc. *Comptes Rendus - Biologies*, *330*(10), 728–734. https://doi.org/10.1016/j.crvi.2007.07.004

Concepcion II, R. S., & Dadios, E. P. (2021). Bioinspired optimization of germination nutrients based on *Lactuca sativa* seedling root traits as influenced by seed stratification, fortification and light spectrums. *AGRIVITA Journal of Agricultural Science*, *43*(1), 174-189. https://doi.org/10.17503/agrivita.v43i1.2843

Concepcion, R. S., Lauguico, S. C., Tobias, R. R., Dadios, E. P., Bandala, A. A., & Sybingco, E. (2020). 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)*. Taipei, Taiwan: IEEE. https://doi.org/10.1109/ARIS50834.2020.9205787

Gholami, M., Rahemi, M., Kholdebarin, B., & Rastegar, S. (2012). Biochemical responses in leaves of four fig cultivars subjected to water stress and recovery. *Scientia Horticulturae*, *148*, 109–117. https://doi.org/10.1016/j.scienta.2012.09.005

Hägele, F., Baur, S., Menegat, A., Gerhards, R., Carle, R., & Schweiggert, R. M. (2016). Chlorophyll fluorescence imaging for monitoring the effects of minimal processing and warm water treatments on physiological properties and quality attributes of fresh-cut salads. *Food and Bioprocess Technology*, *9*(4), 650–663. https://doi.org/10.1007/s11947-015-1661-2

Henderson, I. M. (2015). Ratios, proportions, and mixtures of chlorophylls: Corrections to spectrophotometric methods and an approach to diagnosis. *Limnology and Oceanography: Methods*, *13*(11), 617–629. https://doi.org/10.1002/lom3.10052

Kapur, A., Haskovic, A., Copra-Janicijevic, A., Klepo, L., Topcagic, A., Tahirovic, A., & Sofic, E. (2012). Spectrophotometric analysis of total ascorbic acid contetnt in various fruits and vegetables. *Bulletin* of *the Chemists and Technologists of Bosnia and Herzegovina*, *38*, 39–42. Retrieved from http://hemija.pmf.unsa.ba/glasnik/files/Issue%2038/38%20-%208-Kapur.pdf

Lauguico, S. C., Concepcion II, R. S., Alejandrino, J. D., Tobias, R. R., & Dadios, E. P. (2020). Lettuce life stage classification from texture attributes using machine learning estimators and feature selection processes. *International Journal of Advances in Intelligent Informatics, 6*(2), 173-184. https://doi.org/10.26555/ijain.v6i2.466

Lee, J., Durst, R. W., & Wrolstad, R. E. (2005). Determination of total monomeric anthocyanin pigment content of fruit juices, beverages, natural colorants, and wines by the pH differential method: Collaborative study. *Journal of AOAC International*, *88*(5), 1269–1278. https://doi.org/10.1093/jaoac/88.5.1269

Li, Y., Scales, N., Blankenship, R. E., Willows, R. D., & Chen, M. (2012). Extinction coefficient for red-shifted chlorophylls: Chlorophyll d and chlorophyll f. *Biochimica et Biophysica Acta*, *1817*(8), 1292–1298. https://doi.org/10.1016/j.bbabio.2012.02.026

Liu, H., Fu, Y., Wang, M., & Liu, H. (2016). Green light enhances growth, photosynthetic pigments and CO 2 assimilation efficiency of lettuce as revealed by ‘knock out’ of the 480–560 nm spectral waveband. *Photosynthetica*, *54*, 144–152. https://doi.org/10.1007/s11099-016-0233-7

Loresco, P. J. M., & Dadios, E. P. (2018). A scale-invariant lettuce leaf area calculation using machine vision and knowledge-based methods. *International Journal of Engineering & Technology*, *7*(4), 4880-4885. https://doi.org/ 10.14419/ijet.v7i4.26553

Odabas, M. S., Simsek, H., Lee, C. W., & İseri, İ. (2017). Multilayer perceptron neural network approach to estimate chlorophyll concentration index of lettuce (*Lactuca sativa* L.). *Communications in Soil Science and Plant Analysis*, *48*(2), 162–169. https://doi.org/10.1080/00103624.2016.1253726

Olotu, P. N., Zechariah, J., Ibochi, J. I., Shaibu, K. B., Onche, E. U., Olotu, I. A., & Ajima, U. (2020). Vitamin C quantification and elemental analysis of five local vegetables in jos-south local government area of Plateau State, Nigeria. *Journal of Natural Product and Plant Resources, 10*(3), 1–5. Retrieved from https://www.scholarsresearchlibrary.com/articles/vitaminc-quantification-and-elemental-analysis-of-five-local-vegetables-in-jossouth-local-government-area-of-plateau-sta.pdf

Pacumbaba, R. O., & Beyl, C. A. (2011). Changes in hyperspectral reflectance signatures of lettuce leaves in response to macronutrient deficiencies. *Advances in Space Research*, *48*(1), 32–42. https://doi.org/10.1016/j.asr.2011.02.020

Ritchie, R. J. (2006). Consistent sets of spectrophotometric chlorophyll equations for acetone, methanol and ethanol solvents. *Photosynthesis Research*, *89*(1), 27–41. https://doi.org/10.1007/s11120-006-9065-9

Ritchie, R. J. (2008). Universal chlorophyll equations for estimating chlorophylls a, b, c, and d and total chlorophylls in natural assemblages of photosynthetic organisms using acetone, methanol, or ethanol solvents. *Photosynthetica*, *46*(1), 115–126. https://doi.org/10.1007/s11099-008-0019-7

Santos, D. A., Lima, K. P., Março, P. H., & Valderrama, P. (2016). Vitamin C determination by ultraviolet spectroscopy and multiproduct calibration. *Journal of the Brazilian Chemical Society*, *27*(10), 1912–1917. https://doi.org/10.5935/0103-5053.20160071

Shah, S., Buraidah, M. H., Teo, L. P., Careem, M. A., & Arof, A. K. (2016). Dye-sensitized solar cells with sequentially deposited anthocyanin and chlorophyll dye as sensitizers. *Optical and Quantum Electronics*, *48*(3), 1–8. https://doi.org/10.1007/s11082-016-0492-3

Song, J., Huang, H., Hao, Y., Song, S., Zhang, Y., Su, W., & Liu, H. (2020). Nutritional quality, mineral and antioxidant content in lettuce affected by interaction of light intensity and nutrient solution concentration. *Scientific Reports*, *10*(1), 1–9. https://doi.org/10.1038/s41598-020-59574-3

Steidle Neto, A. J., de Oliveira Moura, L., de Carvalho Lopes, D., de Almeida Carlos, L., Martins, L. M., & de Castro Louback Ferraz, L. (2016). Non-destructive prediction of pigment contents in lettuce based on Vis-NIR spectroscopy. *Journal of the Science of Food and Agriculture*, *97*(7), 2015-2022. https://doi.org/10.1002/jsfa.8002

Sublett, W. L., Casey Barickman, T., & Sams, C. E. (2018). Effects of elevated temperature and potassium on biomass and quality of dark red ‘lollo rosso’ lettuce. *Horticulturae*, *4*(2), 11. https://doi.org/10.3390/horticulturae4020011

Tarrago-Trani, M. T., Phillips, K. M., & Cotty, M. (2012). Matrix-specific method validation for quantitative analysis of vitamin C in diverse foods. *Journal of Food Composition and Analysis*, *26*(1–2), 12–25. https://doi.org/10.1016/j.jfca.2012.03.004

Thrane, J. E., Kyle, M., Striebel, M., Haande, S., Grung, M., Rohrlack, T., & Andersen, T. (2015). Spectrophotometric analysis of pigments: A critical assessment of a high-throughput method for analysis of algal pigment mixtures by spectral deconvolution. *PLoS ONE*, *10*(9), 1–24. https://doi.org/10.1371/journal.pone.0137645

Yang, X., Zhang, J., Guo, D., Xiong, X., Chang, L., Niu, Q., & Huang, D. (2016). Measuring and evaluating anthocyanin in lettuce leaf based on color information. *IFAC-PapersOnLine*, *49*(16), 96–99. https://doi.org/10.1016/j.ifacol.2016.10.018

Zeng, C. (2013). Effects of different cooking methods on the vitamin C content of selected vegetables. *Nutrition and Food Science*, *43*(5), 438–443. https://doi.org/10.1108/NFS-11-2012-0123