

Article

Not peer-reviewed version

Assessing Contents of Sugars, Vitamins, and Nutrients in Baby Leaf Lettuce from Hyperspectral Data with Machine Learning Models

[Sulaymon Eshkabilov](#)^{*} and [Ivan Simko](#)^{*}

Posted Date: 30 April 2024

doi: 10.20944/preprints202404.1969.v1

Keywords: baby leaf lettuce; nutrients; composition; fertilizers; hyperspectral data; machine learning



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Assessing Contents of Sugars, Vitamins, and Nutrients in Baby Leaf Lettuce from Hyperspectral Data with Machine Learning Models

Sulaymon Eshkabilov ^{1,*} and Ivan Simko ^{2,*}

¹ Agricultural and Biosystems Engineering Department, North Dakota State University, Fargo, ND 58102, USA

² U.S. Department of Agriculture, Agricultural Research Service, Crop Improvement and Protection Research Unit, Salinas, CA 93905, USA

* Correspondence: sulaymon.eshkabilov@ndsu.edu (S.E.); ivan.simko@usda.gov (I.S.)

Abstract: Lettuce (*Lactuca sativa*) is a leafy vegetable that provides a valuable source of phytonutrients for a healthy human diet. Assessment of plant growth and composition is vital for determining crop yield and overall quality, however, classical laboratory analyses are slow and costly. Therefore, new, less expensive, more rapid, and non-destructive approaches are being developed, including those based on (hyper)spectral reflectance. Additionally, it is also important to determine how plant phenotypes respond to fertilizer treatments and whether these differences in response can be detected from analyses of hyperspectral image data. In the current study, we demonstrate the suitability of hyperspectral imaging in combination with machine learning models to estimate the content of chlorophyll (SPAD), anthocyanins (ACI), glucose, fructose, sucrose, vitamin C, β -carotene, N, P, K, dry matter content, and plant fresh weight. The implemented five classification and regression machine learning models showed high accuracy in classifying the lettuces by the applied fertilizers treatments and estimating nutrient concentrations. To reduce the input (predictor data, i.e., hyperspectral data) dimension, 13 principal components were found and applied in models. The implemented artificial neural network models of the machine learning algorithm demonstrated high accuracy ($r = 0.85 \dots 0.99$) in estimating fresh leaf weight, and contents of chlorophyll, anthocyanins, N, P, K, and β -carotene. The four applied classification models of machine learning demonstrated 100% accuracy in classifying the studied baby leaf lettuces by phenotype when certain fertilizer treatments were applied.

Keywords: baby leaf lettuce; nutrients; composition; fertilizers; hyperspectral data; machine learning

1. Introduction

Lettuce (*Lactuca sativa* L.) is the most commercially produced leafy vegetable grown in moderate climates worldwide. Fresh lettuce is a desirable, but highly perishable product that should be consumed almost immediately after harvest or maintained in conditions that prevent its wilting, weight loss, enzymatic discoloration, senescence, and tissue deterioration (Peng and Simko, 2023). Lettuce produced in the United States can be grouped into three broad categories: whole heads and bulk harvest (for salad processing) that are commonly harvested at market maturity, and baby leaf grown at very high density and harvested at early stages of plant development (Simko et al., 2014). All lettuce products, including baby leaf lettuce, are a good source of bioactive compounds (Kim M.J. et al., 2016) beneficial to human health (Dillard and German, 2000).

Evaluation of lettuce composition is important for determining its overall quality; however, laboratory analyses are slow and costly. Therefore, more rapid, less expensive, and non-destructive approaches based on (hyper)spectral reflectance are being developed to assess plant phenotypes (Abebe et al., 2023), response to stress (Kumar et al., 2021), and postharvest quality (Simko et al., 2015). Recently, a substantial progress has been made in using reflectance data to estimate lettuce quality and composition, including pH value and the content of chlorophyll, anthocyanins, soluble

solids, water, nitrogen (N), phosphorus (P), potassium (K), calcium (Ca), magnesium (Mg), and sulfur (S) (Ahsan et al., 2021; Eshkabilov et al., 2021; Eshkabilov et al., 2022; Kim C. and van Iersel, 2023; Martins et al., 2021; Pandey et al., 2023; Simko et al., 2016; Yu et al., 2022; Yu et al., 2023; Zhao et al., 2022). However, there is a lack of studies modeling the content of sugars, vitamins, and other compounds in baby leaf lettuce. Moreover, previous studies did not perform classification of lettuce composition based on the applied fertilizers treatments using hyperspectral data and machine learning algorithms. Therefore, the current study focuses on assessing the content of certain sugars, vitamins, and nutrients in baby leaf lettuce from hyperspectral data using partial least squares regression and machine learning models

2. Materials and Methods

2.1. Plant Material

Analyses were performed on a set of 300 greenhouse-grown plants from 50 accessions, which included 36 cultivars, seven plant introductions, six breeding lines, and an accession of *Lactuca serriola* L., the wild progenitor of cultivated lettuce (Appendix A - Table A1). These accessions were selected for the study based on our previous preliminary analyses and published information (Mou, 2005; Mou, 2009) to represent a broad range of phenotypic and composition diversity in lettuce.

Lettuce seeds were sown in potting soil (Premium Growers Mix, Sun Land Garden Products, Watsonville, CA, USA), then covered with ~5 mm of sand, watered, and kept in dark at 10 °C to enhance uniformity of germination. After two days the temperature was increased to 20 °C with 16h/8h light/dark photoperiod. Two weeks later, well-established, and uniformly looking plants were transplanted into 7.6 cm (~514 mL volume) pots. At this point, the plants were split into two groups that were planted into different substrates. One group of plants, later used to analyze the content of sugars and vitamins, was transplanted into a 1:1 mix of potting soil and sand and fertilized with 1.5 g of Osmocote Smart-Release Plant Food Flower & Vegetable (Scotts, Marysville, OH, USA). The second group of plants, later used for nutrient analyses, was transplanted into a 1:2 mix of Espoma VM8 8-Quart Organic Vermiculite (Espoma, Millville, New Jersey, USA) and sand. This group of plants was fertilized with five different combinations of nitrogen (N), phosphorus (P), and potassium (K) fertilizers to ensure a substantial difference in their nutrient composition. The 'NPK' (control) treatment provided a full recommended dose of N, P, and K; the 'nPK' treatment provided 1/3 of N and a full dose of P and K; the 'PK' treatment provided a full dose of P and K, but no N; the 'NP' treatment provided a full dose of N and P, but no K; and the 'NK' treatment provided a full dose of N and K, but no P. N, P, and K macronutrients were provided with Non-Coated Ammonium Nitrate 34-0-0 Prill Form Fertilizer (Intermountain Farmers Association, Salt Lake City, Utah, USA), Triple Super Phosphate 0-46-0 Easy Peasy Plants 99% pure (Easy Peasy Plants, Alvin, Illinois, USA), and All-Natural Muriate of Potash- Easy Peasy 0-0-60 Potassium (Easy Peasy Plants, Alvin, Illinois, USA). All plants were sprayed with the identical amount of micronutrient mix solution that was prepared by dissolving 1 g of Axilo Mix 5 (0-0-0) Micronutrient Mix (Valagro USA, Houston, Texas, USA) in 1 L of distilled water.

Plants from both groups were grown together in a greenhouse in the randomized complete block design with three replications. During the growing period, the average daily temperature in the greenhouse ranged from 20 to 24 °C with a natural day length between 13 to 14h. Outdoor average daily light integrals ranged from 40 to 55 mol m⁻² d⁻¹. Plants were watered daily as needed. When most plants from Osmocote and NPK treatments developed four true leaves about 10 cm long, all plants were evaluated for the content of pigments, scanned with a hyperspectral camera, and subsequently harvested for composition analyses. Plants used in composition analyses were evaluated for biomass production (fresh weight in g).

2.2. Content of Pigments

The content of pigments was determined two days before harvest with hand-held meters that use light transmittance to provide good in situ estimates of chlorophyll (SPAD-502 from Konica

Minolta Sensing, Tokyo, Japan) and anthocyanins (ACM-200 plus from Opti-Sciences, Hudson, New Hampshire, USA). Measurements were taken on three leaves of similar age (avoiding youngest and oldest leaves) about 1 cm from the leaf edge and averaged for each plant. The content of chlorophylls is expressed in SPAD units; the content of anthocyanins is expressed in ACI (anthocyanins content index) units.

2.3. Hyperspectral Imaging

Lettuce plants' reflectance was measured with ASD FieldSpec 3 (Analytical Spectral Devices Inc., Boulder, Colorado, USA) which collects reflectance from 350 nm to 2,500 nm, with a 1.4 nm sampling interval between 350 and 1050, and a 2 nm sampling interval between 1000 and 2500 nm. Hyperspectral data were subdivided into 1 nm bandwidth using an ASD self-driven interpolation method. Spectral jump correction for 725 – 1000 nm and 1800 - 1950 nm bands was performed using the parabolic correction equations (Hueni and Bialek, 2017). The hyperspectral image (HSI) data acquisition set (Figure 1) comprised ASD-Pro halogen lamps, ASD FieldSpec, computer, Spectralon reflectance standard, and connection wires.

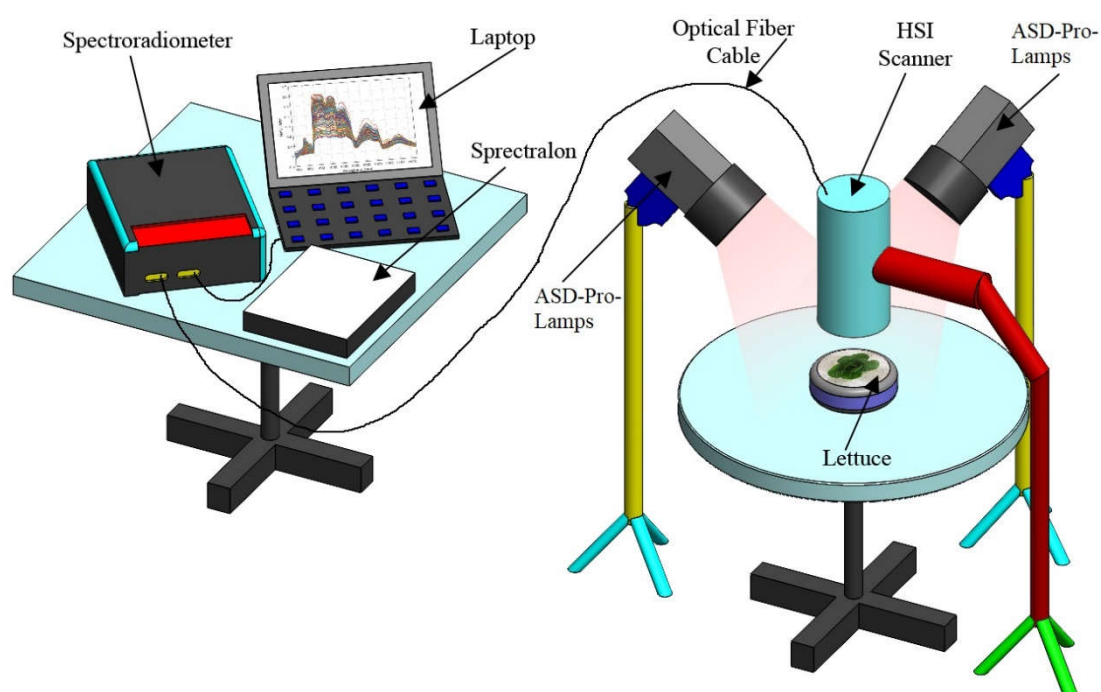


Figure 1. Hyperspectral image data acquisition set-up.

The reflectance measurements were taken for individual plants in a dark room illuminated with four ASD-Pro-Lamps (Analytical Spectral Devices Inc., Boulder, USA). Black, non-reflective cloth was used to cover pots and soil. Hyperspectral images were taken at a height of about 10 cm above the plant. Each plant was scanned four times, with a 90° turn between each measurement. Reflectance calibration was performed using the Spectralon SRT- MS-100 reflectance standard (Labsphere, North Sutton, New Hampshire, USA).

2.4. Laboratory Quantification of Compounds

Immediately after taking reflectance measurements, all leaves from plants were harvested and submitted to the University of California Davis Analytical Laboratory to determine the content of sugars, vitamin C, β -carotene, N, P, and K using common quantification procedures developed for plant tissue samples (<https://anlab.ucdavis.edu>). A detailed description of all analytical methods was previously provided (Simko, 2019; Simko, 2020), therefore they will be mentioned only briefly.

Samples for analyses of sugars were dried at 55°C, ground, and extracted with hot deionized water (Johansen et al., 1996). The amounts of glucose, fructose, and sucrose in extracts were determined using PerkinElmer Series 200 Quaternary HPLC (PerkinElmer, Waltham, Massachusetts, USA) with Sciex API 200 mass spectrometer (Sciex, Redwood City, California, USA). Vitamin C (ascorbic acid) was quantified in leaf tissue according to the previously developed protocol (Bouzari et al., 2015) with minor modifications. β -carotene was quantified from a homogenized sample prepared from fresh tissue and deionized water using HPLC analysis with an isocratic mobile phase of methanol: acetonitrile (90:10). N, P, and K were quantified using the methods based on extraction of soluble nitrate ($\text{NO}_3\text{-N}$) (Miller, 1998), phosphorus ($\text{PO}_4\text{-P}$) (Prokopy, 1995), and potassium (Jones, 2001) from plant material with a solution of 2% acetic acid. The content of glucose, fructose, sucrose, N, P, and K is reported in g per kg of fresh weight (g kg^{-1} FW), and vitamin C and β -carotene are reported in mg per kg of fresh weight (mg kg^{-1} FW). Another part of each sample was oven-dried at 105°C to determine dry matter (DM) content. DM content at both 55°C and 105°C were expressed in percentages of fresh weight.

2.5. Hyperspectral Image Indexing and Extraction Models

The first order derivative of reflectance (FDR) of the filtered HSI data was computed using the formulas elaborated in our previous studies (Eshkabilov, et. al., 2021, Eshkabilov, et. al., 2022). The computed FDR values were explored to find the 13 most important principal components. The reflectance values at the found bandwidth indices were subsequently used to develop multivariate regression models based on the filtered HSI data and laboratory (ground truth) data by applying the PLS and PCA modeling approaches. The ground truth data were those for fresh leaf weight (FLW), percent dry matter content at 55°C (55C DM) and 105°C (105C DM), SPAD, ACI, content of glucose (Glu), fructose (Fru), sucrose (Suc), vitamin C (Vit-C), β -carotene (β -Carot), N, P, and K. A multivariate linear regression modeling (Rangkuti et al., 2017) was applied to establish predictive models using PLSR and PCR approaches.

2.6. Machine Learning Models for Classification and Prediction

To classify and estimate the responses of the studied lettuces to fertilizer treatments using their hyperspectral reflectance data, five classification and regression models of machine learning algorithms were applied: support vector machine (SVM), ensemble, linear and quadratic discriminant, and artificial neural network (ANN). The classification of the studied lettuces based both on their response to fertilizers treatments (PK, NK, NP, NPK, nPK, Osmocote) and HSI data was performed. The estimation and prediction models were found using ANN algorithms. For that 80% of the input HSI data was used for model training, 10% of data for testing and 10% of data for validation. The predictive models were based on the least gradient of mean squared errors over simulation using 1000 epochs with 10 hidden layers and 1 output layer. The implementation of the ANN algorithm for multivariate linear predictive models (Figure 2) composes of five major steps: data import, data preparation, model selection, model evaluation, and regression model performance.

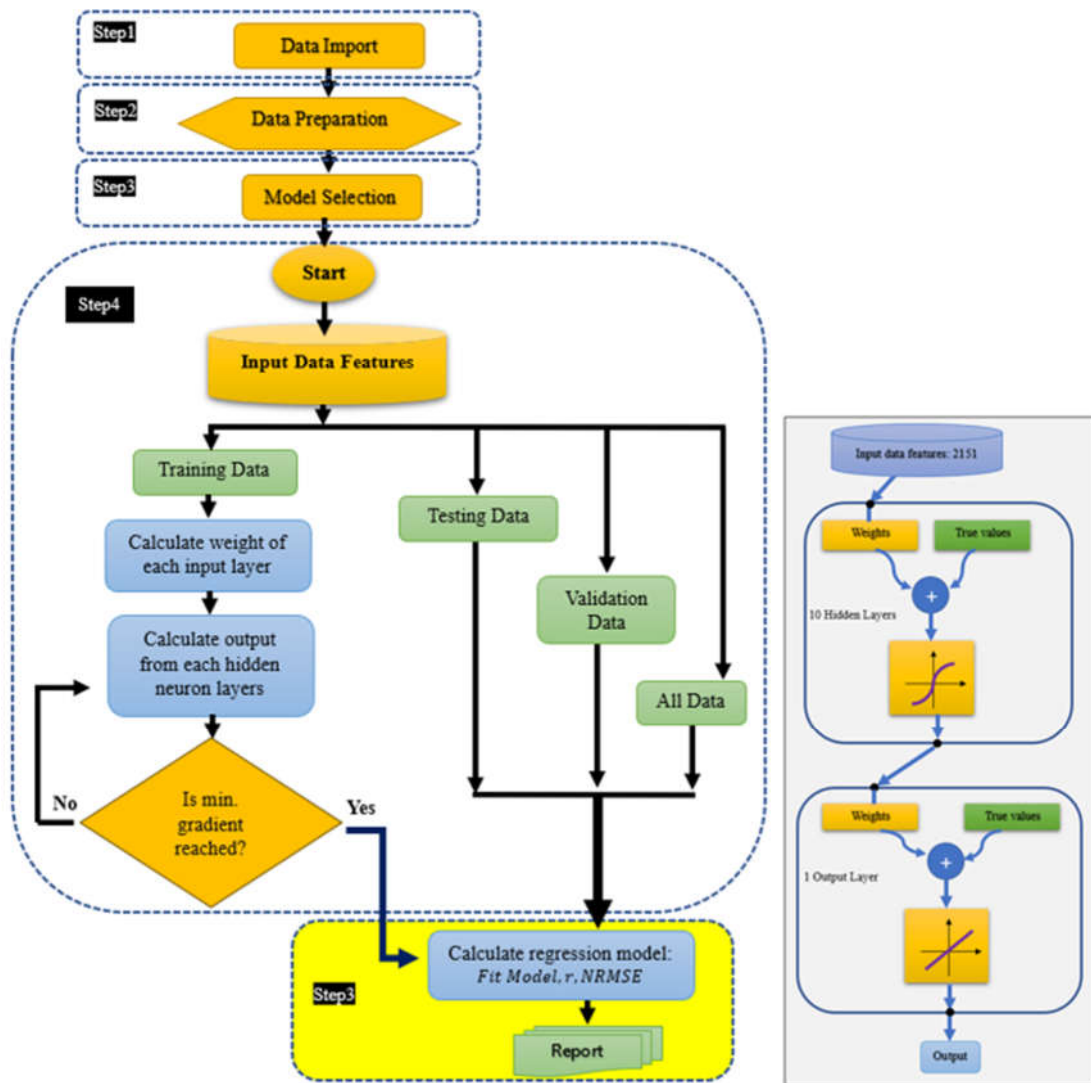


Figure 2. Machine learning algorithm implementation for predictive models.

In the training phase of the ANN model, we applied the Levenberg-Marquardt method (Moré, 1977) which is a built-in function in the MATLAB package (MathWorks, 2023) for linear and non-linear least squares methods.

2.7. Model Accuracy Metrics

Three accuracy metrics of the developed regression and ANN models were used that are correlation coefficient (r), normalized root mean squared error (NRMSE), and accuracy (%). The correlation between the ground truth and predicted values of response variables was computed from Eq (1).

$$r = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^N (y_i - \bar{y})^2}} \quad (1)$$

where x_i and \bar{x} are the laboratory-measured values of a response variable and their mean value, respectively; y_i and \bar{y} are the predicted/estimated values of a response variable and their mean value, respectively.

Another metric used to assess the accuracy of the regression models was NRMSE which was computed from the following formulation – Eq (2).

$$NRMSE = \frac{\sqrt{\frac{\sum_{i=1}^N (x_i - y_i)^2}{N}}}{\bar{y}} \quad (2)$$

where N is the number of data points in the response variable.

Four different classification models (SVM, ensemble, quadratic and linear discriminant) of machine learning algorithms were employed to classify the studied lettuce by the six applied fertilizers treatment (PK, NK, NP, NPK, nPK, Osmocote). The accuracy of the classification models was estimated using true positive (TP), true negative (TN), false positive (FP), and false negative (FN) values and computed using Eq (3) in % (Chollet, 2021).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} * 100\% \quad (3)$$

The accuracy metric given in Eq (3) was applied for all six fertilizer treatments.

3. Results and Discussion

3.1. HSI Data

The captured hyperspectral data of the whole lettuce plant as elaborated in Section 2.3 were filtered using a seven-point moving average digital filter. The filtered hyperspectral reflectance data of all studied lettuce accessions (Figure 3) exhibit six volatile bandwidth regions as found using FDR simulations.

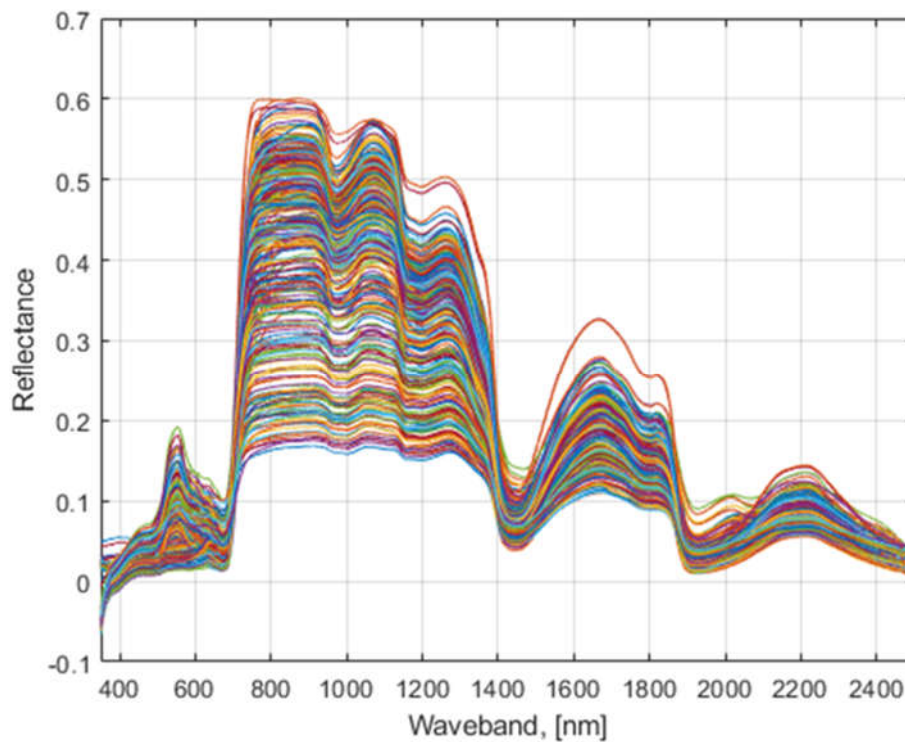


Figure 3. Filtered hyperspectral image data.

3.2. PLSR and PCR Model Performances

In addition to the laboratory-measured values for 13 response variables, two more adjusted values were calculated: a square root of SPAD and logarithm (base 2) of ACI as previously recommended (Simko 2020). The computed FDR of reflectance values show that the 13 most important principal components were found in seven different bandwidth regions, i.e., 500 – 575, 600 – 675, 700 – 775, 950 – 1000, 1100 – 1150, 1300 – 1400, and 1875 – 2000 nm based on the maximum or

minimum reflectance derivative values (Figure 4). These bandwidth regions are in close agreement with those determined previously on lettuce (Falcioni et al., 2022). Correlations between the predicted values and laboratory measurements (ground truth) for multivariate linear PLSR and PCR models ranged from 0.5450 to 0.9476 (Figure 5), while NRMSE values ranged between 0.3397 to 0.9038. The highest correlation (0.9476 and 0.9306) between the predicted values found using PLS and PCR multivariate models and laboratory-measured values was observed for fresh leaf weight (FLW). The lowest correlation (0.6442 and 0.5445) between modeled and measured data was observed for the content of potassium (K). Correlation values of PLSR and PCR models to estimated SPAD, N, P, and K nutrients were lower than those found in our previous studies (Eshkabilov, et. al., 2021, Eshkabilov, et. al., 2022). This difference may be related to the scanning approach used on analyzed plants. In the current study, the whole lettuce plants were scanned using the hyperspectral camera, whereas in our previous study, hyperspectral images were collected from individual lettuce leaves placed in a fully isolated chamber. Correlations for N, P, and K content determined in the current study using PLSR and PCR models are, however, in close agreement with the results from other laboratories (Pandey, et. al., 2023). The accuracy of the PLSR and PCR models were higher for N and P, but lower for K in our study than previously reported accuracies for the PLSR models (Taha, et. al., 2022).

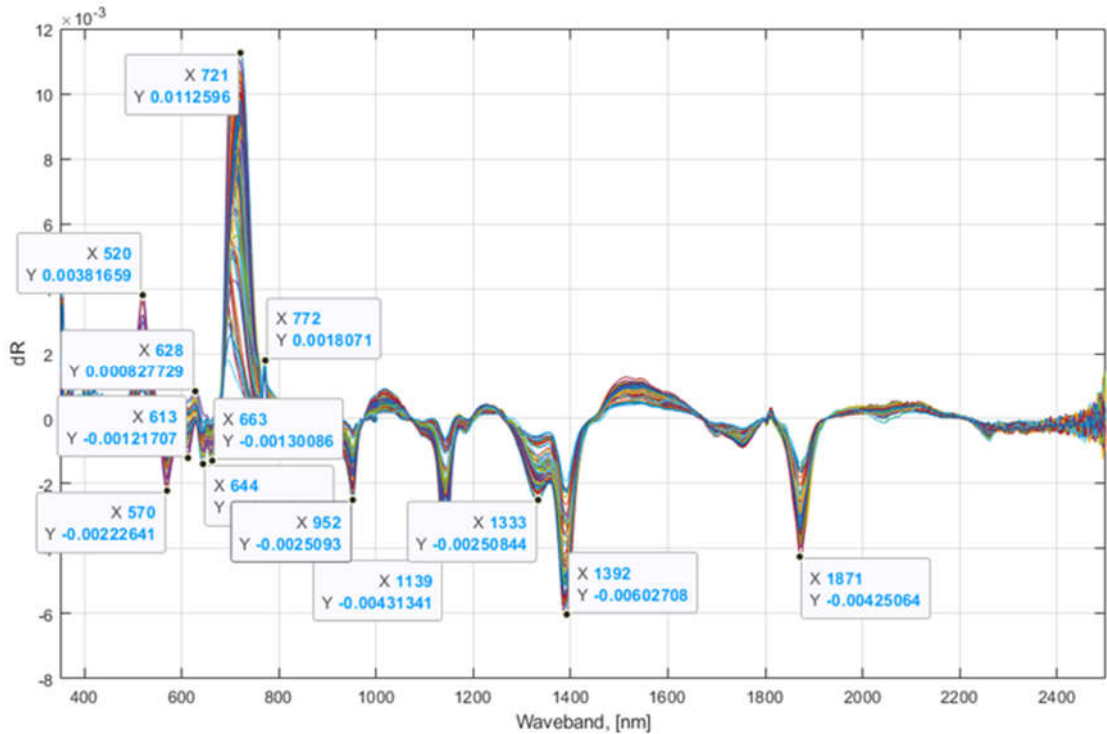


Figure 4. Principal components of multivariate PLS and PCR models.

SPAD	0.9099	0.4147	0.9094	0.4159	0.9 0.8 0.7 0.6 0.5 0.4
ACI	0.6745	0.7382	0.6693	0.743	
sqrt(SPAD)	0.8325	0.7446	0.8258	0.7556	
log ₂ ACI	0.8532	0.5225	0.8456	0.5489	
FLW	0.9476	0.3397	0.9306	0.3586	
Glu	0.8402	0.5422	0.7987	0.6017	
Fru	0.8579	0.5138	0.8155	0.5788	
Suc	0.7092	0.8398	0.5771	0.9038	
55C DM	0.905	0.5585	0.9028	0.5983	
Vit-C	0.8289	0.7556	0.8012	0.7823	
β-Carot	0.655	0.4254	0.6229	0.4301	
105C DM	0.7029	0.7113	0.6819	0.7314	
N	0.8216	0.5701	0.8198	0.5727	
P	0.9094	0.4159	0.8911	0.4539	
K	0.6442	0.7649	0.545	0.8385	
					0.9 0.8 0.7 0.6 0.5 0.4
					0.9 0.8 0.7 0.6 0.5 0.4

Figure 5. Performances of multivariate PLS and PCR models.

3.3. Classification of Lettuces by the Applied Fertilizers Treatments Using Machine Learning Algorithms

To classify the applied six fertilizer treatments (PK, NK, NP, NPK, nPK, and Osmocote) for all studied lettuces and the collected hyperspectral image data, four different machine learning classification algorithms (SVM, ensemble, linear and quadratic discriminant) were implemented and tested. The filtered HSI data was taken as an input variable and the applied six treatments, converted to categorical arrays, were used as an output variable. To reduce the dimension of the input (predictor) variable data, i.e., HSI data, 13 wavebands of HSI, 13 principal component values found using FDR reflectance values (Figure 4) were used. HSI data were classified based on the six applied fertilizer treatments. 90% of 300 data sets were used for model training and 10% for model testing/validation. The employed classification scales were individual classes - TP, FP, TF, FN, and the combined classes - TPR (true positive rates) and FNR (false negative rates). The plotted confusion matrices of the employed algorithms (Figure 6) show that the highest accuracy (77.8% for model training and 73.3% for model testing/validation) in classifying the applied treatments was achieved with the ensemble algorithm (Figure 6, a). The second highest accuracy (75.9% for model training and 70.0% for model testing/validation) of classification was reached using the SVM algorithm (Figure 6, b). The other two algorithms, quadratic discriminant (Figure 6, c) and linear discriminant (Figure 6, d), resulted in 74.7% and 80% for model training, and 70.0% and 66.7% accuracy for model testing/validation, respectively. 100% accuracy was achieved in classifying lettuces using SVM, ensemble, linear, and quadratic discriminant algorithms when NK, Osmocote, and PK treatments were applied (Figure 6, a, b, c, d). This demonstrates that NK, Osmocote, and PK fertilizers have significant effects on the leaf pigment content. Similar conclusions (Figure 6, a, b) can be drawn with adequate confidence about the nPK treatment effect. On the other hand, NP and NPK treatments (Figure 6, a, b, c, d) had a very small or no effect on lettuce leaf pigments and therefore, the hyperspectral imaging technique could not detect the impact of NP and NPK treatments.

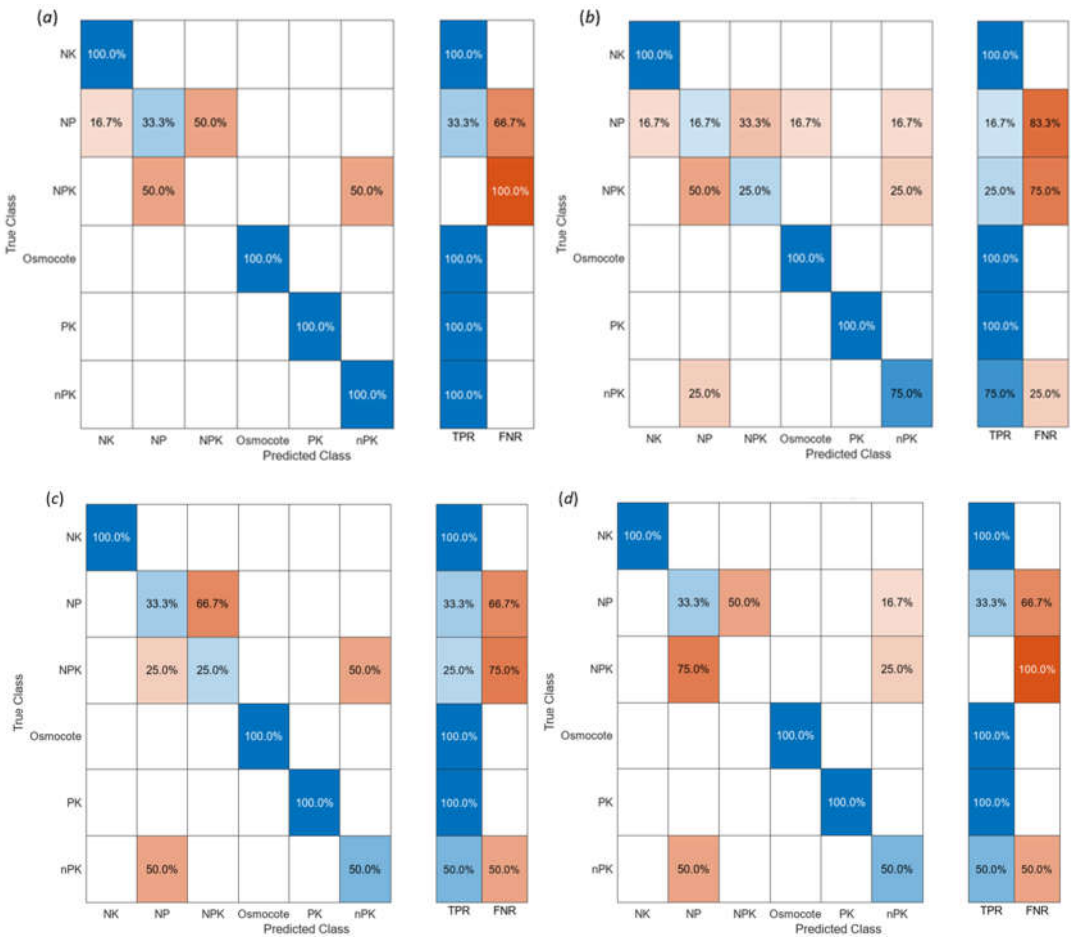


Figure 6. Confusion matrix of the classification accuracy using (a) SVM; (b) Ensemble; (c) Quadratic Discriminant; (d) Linear Discriminant algorithms.

3.4. Performances of Machine Learning Algorithms to Estimate Nutrients of Lettuces

To estimate and predict nutrient concentration levels in lettuces based on hyperspectral image data using artificial neural network algorithms of machine learning, the HSI data were treated as a predictor variable, and laboratory-measured values were treated as a response variable. The acquired HSI data of lettuces were filtered using a Gaussian seven-point moving average filter. Two additional transformed values, the square root of SPAD and the logarithm of ACI, were included in ANN models. Thus, a total of 15 response variables were simulated.

The predictor variable (filtered HSI data) and response variable values were randomly split into 80%, 10%, and 10% for training, testing, and validation sets. 1000 epochs simulations with 10 hidden layers and 1 output layer for ANN models were applied to find the best fitting multivariate predictive linear model based on three performance metrics: mean squared error, error gradient, and correlation coefficient (*r*) between the predicted values and laboratory-measured (ground truth) values. For the ANN model simulations, two different sets of response variable values were tested, original data (laboratory-measured) (Figure 7), and a normalized version of the laboratory-measured data (Figure 8).

When the original data and the ANN model were applied, the highest correlations between the estimated nutrient values and laboratory-measured values were detected for fresh leaf weight (FLW), log₂(ACI), N, P, K, SPAD, and β-carotene in the model training phase (*r* = 0.9808 ... 0.9998), testing (*r* = 0.7500...0.9357), and validation (*r* = 0.8480...0.9400) (Figure 7). The accuracy of the estimation models for N, P, K contents are in good agreement with the accuracy of the models developed using back propagation neural network and random forest algorithms (Taha, et. al., 2022). The lowest correlation between the estimated and ground truth values was found for Vit-C, Glu, Fru, 55C DM, 105C DM, and the square root of SPAD, which also showed much higher NRMSE values (Figure 7).

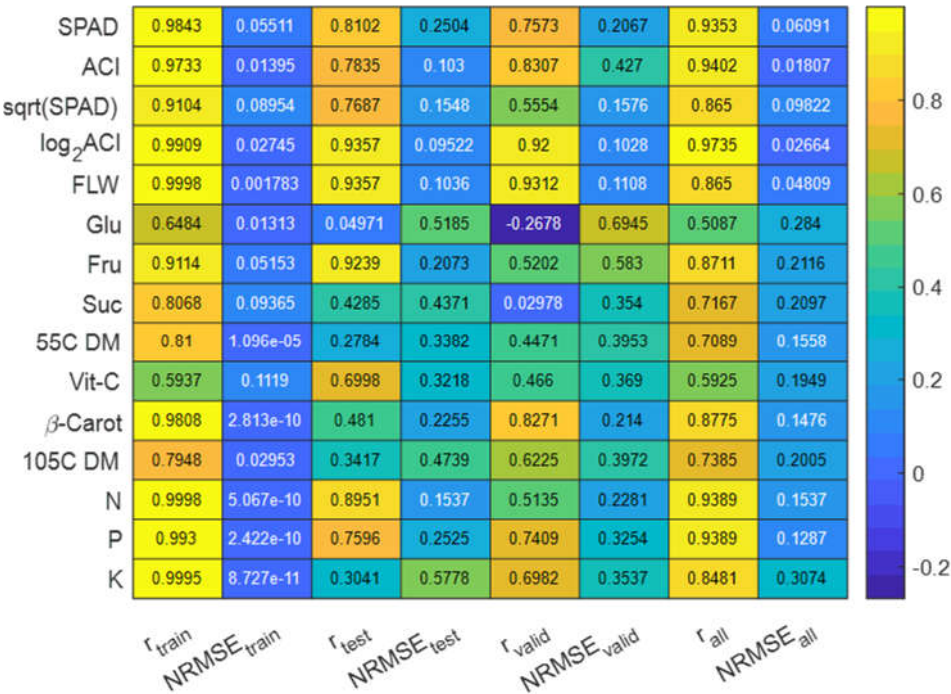


Figure 7. Performances of multivariate ANN models with original data.

The ANN model's performance in estimating nutrient levels and other traits, when using normalized values of response variables, showed a slight decline in overall accuracy in terms of the correlation between predicted and measured values for all nutrients except Glu and 55C DM in the

training set (Figure 8). Conversely when evaluating the validation data sets, the models demonstrated improved accuracy for most traits, and the use of normalized response variables resulted in reduced normalized error margins (NRMSE).

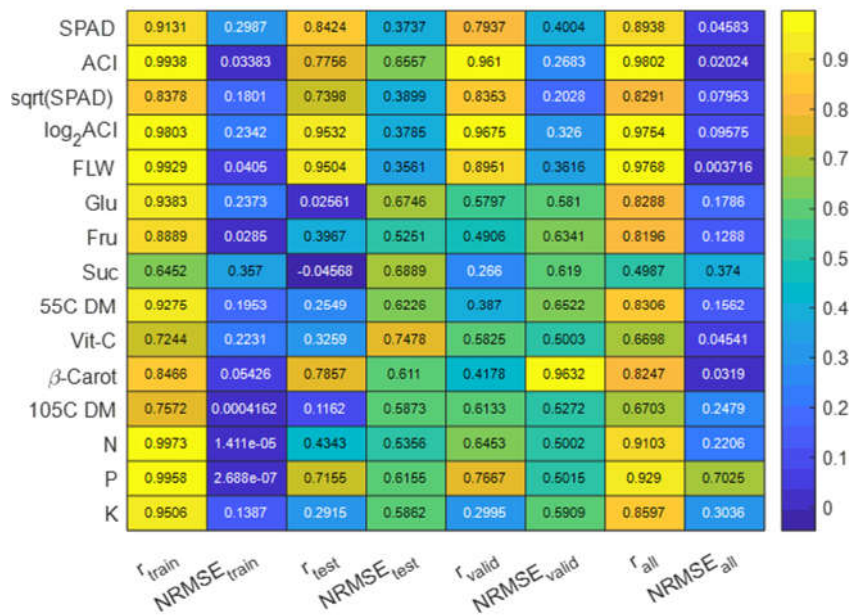


Figure 8. Performances of multivariate ANN models with normalized data.

4. Conclusions

Based on the data processing approach and the estimated performance of models that were developed from hyperspectral image data, the following can be concluded.

1. The performed studies demonstrate that the accuracy of the PLSR and PCR models degrades when the hyperspectral data of the whole lettuce plant is collected. Therefore, machine learning algorithms have become a more reliable solution for model development.
2. Application of machine learning algorithms for classification of the lettuces by the applied treatments based on hyperspectral image data of whole lettuce plant is reliable and provides sufficiently high accuracy, in particular for NK, Osmocote, and PK applied nutrients. This approach can be used to eventually develop a valuable and practical model to identify nutrient deficiency in plants using hyperspectral image data.
3. The accuracy of models to assess nutrient levels of the studied lettuce cultivars using machine learning algorithms demonstrated that a larger number of predictor (input) and response (output) datasets will likely result in a higher accuracy of the estimation models. For example, SPAD and ACI data collected on 300 samples provided better results than traits with smaller datasets. Additionally, data skewness of response variables may hinder the performances of the machine learning models.

Authorship Contribution: Sulaymon Eshkabilov: Methodology, Software, Validation, Formal analysis, Investigation, Data Analysis, Writing – original draft, Writing – review & editing, Visualization. Ivan Samko: Introduction, Methodology, Investigation, Data collection, Writing – original draft, Writing – review & editing.

Acknowledgments: The authors would like to thank Leah Rosental for growing plants, Prof. Susan Ustin and Leah Rosental for organizing the hyperspectral scanning, and Mr. Prashant Lakhemaru, a graduate student at NDSU, for drawing Figure 1. The mentioned trade names or commercial products in this publication are solely to provide specific information and do not imply recommendation or endorsement by the U.S. Department of Agriculture (USDA).

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

Appendix A

Table A1 List of tested accessions, fertilizer treatments, and composition analyses.

Table A1. List of tested accessions, fertilizer treatments, and composition analyses.

Accession	Description	Osmocote	NPK	nPK	PK	NP	NK
Alpi	cultivar	P, C, V, W					
Annapolis	cultivar	P, C, V, W	P, E, W	P, E, W	P, E, W	P, E, W	P, E, W
Balady Barrage	cultivar	P, C, V, W	P, E, W	P, E, W	P, E, W	P, E, W	P, E, W
Balady Cairo	cultivar	P, C, V, W					
Bandit	cultivar	P, C, V, W					
Barcarole	cultivar	P					
Caesar	cultivar	P, C, V, W					
Clemente	cultivar	P					
Conquistador	cultivar	P					
Dark Green Romaine	cultivar	P, C, V, W					
Darkland	cultivar	P, C, V, W	P, E, W	P, E, W	P, E, W	P, E, W	P, E, W
Eruption	cultivar	P, C, V, W	P, E, W	P, E, W	P, E, W	P, E, W	P, E, W
FLA24069	breeding line	P, C, V, W					
Floricos	cultivar	P, C, V, W					
Gardenia	cultivar	P, C, V, W					
Green Forest	cultivar	P, C, V, W					
Green Towers	cultivar	P, C, V, W	P	P	P	P	P
Heavy Heart	cultivar	P, C, V, W					
Inverno De Mall	cultivar	P, C, V, W					
La Brillante	cultivar	P, C, V, W	P, E, W	P, E, W	P, E, W	P, E, W	P, E, W
Lee Tal	cultivar	P, C, V, W					
Little Gem	cultivar	P, C, V, W					
Little Lepricon	cultivar	P, C, V, W					
Merlot	cultivar	P, C, V, W	P, E, W	P, E, W	P, E, W	P, E, W	P, E, W
Parris Island Cos	cultivar	P, C, V, W					
PI 257288	plant introduction	P, C, V, W					
PI 278074	plant introduction	P, C, V, W					
PI 278100-COS	plant introduction	P, C, V, W					
PI 358033-COS	plant introduction	P, C, V, W					
PI 491038	plant introduction	P					
PI 491086	plant introduction	P, C, V, W					
PI 665200	plant introduction	P, C, V, W					
Queen of Hearts	cultivar	P, C, V, W					
Romanor	cultivar	P					
Romaserra	cultivar	P, C, V, W					
Rubicon	cultivar	P, C, V, W					
Salinas	cultivar	P, C, V, W	P, E, W	P, E, W	P, E, W	P, E, W	P, E, W
Salinas 88	cultivar	P					
SalVal-321	breeding line	P, C, V, W					
Siskiyou	cultivar	P, C, V, W					
SM09A	breeding line	P					
SM09B	breeding line	P, C, V, W					
SM13-R2	breeding line	P, C, V, W					
SM13-R3	breeding line	P, C, V, W					
Solar	cultivar	P, C, V, W					
Taiwan	cultivar	P, C, V, W	P	P	P	P	P
Triple Threat	cultivar	P					
UC96US23	wild species	P, C, V, W	P, E, W	P, E, W	P, E, W	P, E, W	P, E, W
Ultegra	cultivar	P, C, V, W					
Valmaine	cultivar	P, C, V, W					

Abbreviations for analyzed composition: P (pigments): anthocyanins, chlorophyll; C (carbohydrates): glucose, fructose, sucrose; V (vitamins and precursors): vitamin C, β-carotene; E (elements): nitrogen, phosphorus, potassium; W (weight): fresh weight.

References

1. Abebe AM, Kim Y, Kim J, Kim SL, Baek J. 2023. Image-based high-throughput phenotyping in horticultural crops. *Plants*. 12:2061.
2. Ahsan M, Eshkabilov S, Cemek B, Küçüktopcu E, Lee CW, Simsek H. 2021. Deep learning models to determine nutrient concentration in hydroponically grown lettuce cultivars (*Lactuca sativa* L.). *Sustainability*. 14:416.
3. Bouzari A, Holstege D, Barrett DM. 2015. Vitamin retention in eight fruits and vegetables: a comparison of refrigerated and frozen storage. *Journal of Agricultural and Food Chemistry*. 63:957-962.
4. Chollet, F. (2021). *Deep learning with Python* (2nd ed.). Shelter Island, NY: Simon & Schuster.
5. Dillard CJ, German JB. 2000. Phytochemicals: nutraceuticals and human health. *Journal of the Science of Food and Agriculture*. 80:1744-1756.
6. Eshkabilov S, Lee A, Sun X, Lee CW, Simsek H. 2021. Hyperspectral imaging techniques for rapid detection of nutrient content of hydroponically grown lettuce cultivars. *Computers and Electronics in Agriculture*. 181:105968.
7. Eshkabilov S, Stenger J, Knutson EN, Küçüktopcu E, Simsek H, Lee CW. 2022. Hyperspectral image data and waveband indexing methods to estimate nutrient concentration on lettuce (*Lactuca sativa* L.) cultivars. *Sensors*. 22:8158.
8. Falcioni R, Gonçalves JVF, Oliveira KMd, Antunes WC, Nanni MR. VIS-NIR-SWIR Hyperspectroscopy Combined with Data Mining and Machine Learning for Classification of Predicted Chemometrics of Green Lettuce. *Remote Sensing*. 2022; 14(24):6330. <https://doi.org/10.3390/rs14246330>
9. Hueni A, Bialek A. 2017. Cause, effect, and correction of field spectroradiometer interchannel radiometric steps. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 10:1542-1551.
10. Johansen HN, Glitsø V, Bach Knudsen KE. 1996. Influence of extraction solvent and temperature on the quantitative determination of oligosaccharides from plant materials by high-performance liquid chromatography. *Journal of Agricultural Food Chemistry*. 44:1470-1474.
11. Jones BJJ. 2001. Extraction of Chloride (Cl), Nitrate (NO₃), Orthophosphate (PO₄), Potassium (K), and Sulfate (SO₄) from Plant Tissue Using 2% Acetic Acid, p. 228-229 *Laboratory Guide for Conducting Soil tests and Plant Analysis*. CRC Press, Boca Raton, FL.
12. Kim C, van Iersel MW. 2023. Image-based phenotyping to estimate anthocyanin concentrations in lettuce. *Frontiers in Plant Science*. 14:1155722.
13. Kim MJ, Moon Y, Tou JC, Mou B, Waterland NL, Analysis. 2016. Nutritional value, bioactive compounds and health benefits of lettuce (*Lactuca sativa* L.). *Journal of Food Composition*. 49:19-34.
14. Kumar P, Eriksen RL, Simko I, Mou B. 2021. Molecular mapping of water-stress responsive genomic loci in lettuce (*Lactuca* spp.) using kinetics chlorophyll fluorescence, hyperspectral imaging and machine learning. *Frontiers in Genetics*. 12:634554.
15. Martins GD, Silva OFd, Carmo GJdS, Castoldi R, Santos L, Charlo HCdO. 2021. Estimation of biometric, physiological, and nutritional variables in lettuce seedlings using multispectral images. *Revista Brasileira de Engenharia Agrícola e Ambiental*. 25:689-695.
16. MathWorks 2023. MATLAB 9.14 and Statistics and Machine Learning Toolbox - Ver 12.5, Release 2023a, The MathWorks, Inc., Natick, MA, USA.
17. Miller RO. 1998. Extractable Chloride, Nitrate, Orthophosphate, Potassium, and Sulfate-Sulfur in Plant Tissue: 2% Acetic Acid Extraction, p. 115-118. In: Kalra YP (ed.). *Handbook of Reference Methods for Plant Analysis*. CRC Press, Boca Raton, FL.
18. Moré, J.J., "The Levenberg-Marquardt Algorithm: Implementation and Theory," *Numerical Analysis*, ed. G. A. Watson, *Lecture Notes in Mathematics* 630, Springer Verlag, pp 105–116, 1977.
19. Mou B. 2005. Genetic variation of beta-carotene and lutein contents in lettuce. *Journal of the American Society for Horticultural Science*. 130:870-876.
20. Mou B. 2009. Nutrient content of lettuce and its improvement. *Current Nutrition & Food Science*. 5:242-248.
21. Pandey P, Veazie P, Whipker B, Young S. 2023. Predicting foliar nutrient concentrations and nutrient deficiencies of hydroponic lettuce using hyperspectral imaging. *Biosystems Engineering*. 230:458-469.
22. Peng H, Simko I. 2023. Extending lettuce shelf life through integrated technologies. *Current Opinion in Biotechnology*. 81:102951.
23. Prokopy WR. 1995. Phosphorus in acetic acid extracts. Lachat Instruments, Milwaukee, WI.
24. Rangkuti M, Saputro, A, Imawan C. 2017. Prediction of soluble solid contents mapping on *Averrhoa carambola* using hyperspectral imaging. *IEEE* (2017), pp. 414-419
25. Simko I. 2020. Predictive modeling of a leaf conceptual midpoint quasi-color (CMQ) using an artificial neural network. 20:3938.

26. Simko I, Hayes RJ, Mou B, McCreight JD. 2014. Lettuce and Spinach, p. 53-86. In: Smith S, Diers B, Specht J, Carver B (eds.). *Yield Gains in Major U.S. Field Crops*. American Society of Agronomy, Inc., Crop Science Society of America, Inc., and Soil Science Society of America, Inc. <https://doi.org/10.2135/cssaspecpub33.c4>.
27. Simko I, Jimenez-Berni JA, Furbank RT. 2015. Detection of decay in fresh-cut lettuce using hyperspectral imaging and chlorophyll fluorescence imaging. *Postharvest Biology and Technology*. 106:44-52.
28. Simko I, Hayes RJ, Furbank RT. 2016. Non-destructive phenotyping of lettuce plants in early stages of development with optical sensors. *Frontiers in Plant Science*. 7:1985. 10.3389/fpls.2016.01985.
29. Simko I. 2019. Genetic variation and relationship among content of vitamins, pigments, and sugars in baby leaf lettuce. *Food Science & Nutrition*. 7:3317-3326.
30. Simko I. 2020. Genetic variation in response to N, P, or K deprivation in baby leaf lettuce. *Horticulturae*. 6:15.
31. Taha MF, ElManawy AI, Alshallash KS, ElMasry G, Alharbi K, Zhou L, Liang N, Qiu Z. Using Machine Learning for Nutrient Content Detection of Aquaponics-Grown Plants Based on Spectral Data. *Sustainability*. 2022; 14(19):12318. <https://doi.org/10.3390/su141912318>
32. Yu S, Fan J, Lu X, Wen W, Shao S, Guo X, Zhao C. 2022. Hyperspectral technique combined with deep learning algorithm for prediction of phenotyping traits in lettuce. *Frontiers in Plant Science*. 13:927832.
33. Yu S, Fan J, Lu X, Wen W, Shao S, Liang D, Yang X, Guo X, Zhao C. 2023. Deep learning models based on hyperspectral data and time-series phenotypes for predicting quality attributes in lettuces under water stress. *Computers and Electronics in Agriculture*. 211:108034.
34. Zhao J, Li H, Chen C, Pang Y, Zhu X. 2022. Detection of water content in lettuce canopies based on hyperspectral imaging technology under outdoor conditions. *Agriculture*. 12:1796.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.