



Article

Classification Models for Nitrogen Concentration in Hop Leaves Using Digital Image Processing

Lucas Gomes de Brito ¹, Rodrigo Chaves Jorge ¹, Victor Crespo de Oliveira ^{2,*}, Patrícia Ferreira Cassemiro ¹, Alexandre Dal Pai ², Valéria Cristina Rodrigues Sarnighausen ² and Sergio Augusto Rodrigues ²

¹ Department of Bioprocess and Biotechnology, School of Agricultural Sciences, São Paulo State University, Botucatu 18610-034, SP, Brazil; lucas.g.brito@unesp.br (L.G.d.B.); rodrigo.chaves@unesp.br (R.C.J.); patricia.cassemiro@unesp.br (P.F.C.)

² Agricultural Engineering Program, School of Agricultural Sciences, São Paulo State University, Botucatu 18610-034, SP, Brazil; dal.pai@unesp.br (A.D.P.); valeria.sarnighausen@unesp.br (V.C.R.S.); sergio.rodrigues@unesp.br (S.A.R.)

* Correspondence: victor.crespo@unesp.br

Abstract: Hop (*Humulus lupulus* L.) is a climbing plant that contains essential components for beer production. Although Brazil is the third-largest beer producer in the world, it still relies on imports to meet demand. Some hop varieties have already adapted to the tropical climate, but nitrogen fertilization is essential for the proper development of plants. Digital image processing (DIP) and modeling technologies are emerging as fast and economical alternatives for monitoring the nutritional status of plants. This study evaluated the impact of image quality and the performance of models in classifying hop plants in terms of nitrogen concentration, using predictors extracted from leaf images. A total of 24 plants subjected to six levels of fertilization, ranging from 0 to 200% of the optimal level, were analyzed. The leaves were classified into two nitrogen concentration groups and the data organized into two sets: one containing only significant variables and another including all the variables in the model. Linear discriminant analysis (LDA) and quadratic discriminant analysis (QDA) models were estimated. The QDA models demonstrated great efficacy in classifying plants with a high nitrogen concentration, achieving over 80% accuracy, although performance was lower for plants with a lower nitrogen concentration.

Keywords: discriminant analysis; modeling; *Humulus lupulus*; data analysis; fertilization



Academic Editor: Borja Velazquez-Marti

Received: 18 February 2025

Revised: 11 March 2025

Accepted: 13 March 2025

Published: 25 April 2025

Citation: Brito, L.G.d.; Jorge, R.C.; Oliveira, V.C.d.; Cassemiro, P.F.; Dal Pai, A.; Sarnighausen, V.C.R.; Rodrigues, S.A. Classification Models for Nitrogen Concentration in Hop Leaves Using Digital Image Processing. *Appl. Sci.* **2025**, *15*, 4799. <https://doi.org/10.3390/app15094799>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Hop (*Humulus lupulus* L.) is a plant native to the Northern Hemisphere, whose inflorescences, called cones, have palatable and antibacterial properties. Due to its presence of alpha and beta acids and essential oils, hop is mainly used for brewing beer, bringing flavor, bitterness, and aroma to the drink [1–3]. Although Brazil is ranked as the third-largest beer producer globally, with an annual output of approximately 13.3 billion liters, most of the hops used in the process are imported from major global producers such as Germany and the United States [4]. Despite efforts to cultivate hop varieties adapted to warm climates, the cultivated area remains small [5], and significant gaps persist in scientific research regarding its management, production, and quality in tropical climates.

Nevertheless, the cultivation of hop has been expanding in Brazil, with indications that the country experiences a growing harvest each year [6]. Conditions such as Brazil's predominant soil type, latosol [7], extended periods of sunlight, and the climatic requirements of hop cultivation [1] create a favorable environment for crop development in the country.

However, another essential factor for the healthy growth and commercial production of hop is fertilization, particularly nitrogen supplementation. This nutrient is crucial for primary growth, playing a structural function in amino acid synthesis. These characteristics accelerate plant growth, promoting cone flowering [8].

Studies have shown that nitrogen application positively affects early hop development [9], which can also influence cone quality by altering acid content, total oils, dry matter, overall yield, and cone color [10,11]. However, excessive nitrogen application can lead to excessive plant growth and predispose the plant to a greater risk of leaf diseases and insect infestations [1].

At the same time, nitrogen availability is reduced due to urea volatilization and leaching, which may compromise plant development. Therefore, monitoring the available nitrogen levels is essential, and can be carried out through leaf analysis. This method assesses the leaves and verifies the response of the plant's metabolism to a shortage, sufficiency, or excess of this nutrient [12]. Furthermore, biochemical analyses in the laboratory are not always carried out by producers due to high costs or limited access to testing facilities. This highlights the need for alternative nutritional monitoring methods.

The use of digital image processing (DIP) computational tools, such as RGB (Red, Blue, and Green) analysis, offers a lower-cost and time-efficient alternative for assessing plant nutritional status. In addition, compared to visual assessment, it offers greater precision, improving operational and financial capacity [13,14]. Studies have shown that classification models generated via machine learning techniques outperform human evaluation in determining the geometric parameters in leaf images [15].

In the context of hop cultivation, image processing was used by Castro et al. [16] to classify hop species, presenting the first study of image processing and hop variety classification in Brazil. In this work, they developed a classification model based on 12 varieties and 1592 leaf images, achieving an accuracy of 95.69%. However, the application of these technologies remains largely unexplored in hop cultivation, especially in Brazil, where the literature is scarce and still primarily exploratory and observational.

Multivariate statistical and machine learning techniques have been used to classify nutritional deficiencies and phytopathologies using leaf image processing [14,17,18]. Among the statistical models used, discriminant analysis, both linear (LDA) and quadratic (QDA), has been successfully applied. LDA was one of the models used by Alajas et al. [15], achieving the second-highest accuracy in the experiment, as well as the fastest performance in distinguishing the quality of grape leaves damaged by black rot, even though it misclassified four leaves.

Similarly, in Pydipati et al.'s study [19], when used together with the color co-occurrence methods, texture-based hue, saturation, and intensity color features to identify grapefruit leaves (Duncan variety) under laboratory conditions, LDA achieved classification accuracies of over 95% in all classes using hue and saturation texture variables. However, models relying on intensity-based features showed a reduced classification capacity when classifying the front of the leaves, due to their pigmentation.

Regarding post-harvest carrot processing, Jahanbakhshi and Kheiralipour [17] evaluated image processing techniques and discriminant analysis methods. Their study utilized 135 samples with different shapes and regularities (56 regular and 79 irregular) and extracted variables using DIP techniques. LDA and QDA were estimated, achieving classification rates of 92.59% and 96.30%, respectively. The performance of the quadratic analysis was comparable to that in similar studies on potatoes and cucumbers [18,20].

Therefore, developing models to identify and classify potential nitrogen deficiencies in hop plants through leaf image analysis could contribute to the sector. In the long term, this classification tool could be advantageous for producers by enabling the acquisition of

reliable data, facilitating crop management, and reducing costs. In addition, this technology is already widely used in other crops, such as soybean [21,22] and maize [23].

The objective of this study was to evaluate the impact of image quality and model performance in classifying hop plants with low or high nitrogen concentrations, considering predictors derived from digitally processed leaf images obtained with a mobile phone camera (low resolution) and a digital scanner printer (high resolution). The prospect of accurate nutrient monitoring using widely available mobile phones could bring significant benefits in terms of cost and productivity in hop cultivation management.

2. Materials and Methods

2.1. Location of the Experiment

Hop plants of the Cascade variety were cultivated in a protected environment (greenhouse) belonging to the LUPAM group (Hops, Research, Applications, and Management), located in the experimental area of the School of Agricultural Sciences (FCA) of São Paulo State University (UNESP “Julio de Mesquita Filho”) in Botucatu, São Paulo, Brazil.

The greenhouse is located at 22°51' south latitude, 48°26' west longitude, and 786 m altitude. It measures 24 × 7 × 2.5 m and is constructed with an arched metal frame. The roof is made of transparent low-density polyethylene with a diffuser (100 mm thick) and the sides are closed with an anti-snake mesh. The climate of the municipality of Botucatu according to the Köppen classification is Aw, with hot, humid summers and cold, dry winters [24].

2.2. Experimental Characterization

The experiment began in January 2022. Cascade hop plants were arranged in four rows oriented north–south, with each row containing 25 plants spaced 0.8 m apart, and 1 m between rows, totaling 100 plants grown in Red Latosol soil. The chemical and physical characteristics of the experimental soil were analyzed and are presented in Tables 1 and 2, respectively.

Table 1. Chemical properties of the experimental soil.

Depth (m)	pH	OM	V	Al ³⁺	H + Al	K	Ca	Mg	CEC
	CaCl ₂	g/dm ³	%			mmol _c /dm ³			
0–0.20	5.6	17	74	0	23	1.1	51	12	87
0.20–0.40	4.9	12	52	0	38	1.4	32	8	79
Depth (m)	P _{resin}	S	B	Cu	Fe	Mn	Zn		
	mg/dm ³								
0–0.2	309	9	0.22	6.3	72	7.3	24.9		
0.2–0.4	173	12	0.26	4.1	52	4.2	11.5		

OM: organic matter; pH: potential of hydrogen measured in CaCl₂ solution; P_{resin}: phosphorus extracted with ion-exchange resin; V: base saturation; Al³⁺: exchangeable acidity; H + Al: potential acidity; K: potassium; Ca: calcium; Mg: magnesium; CEC: cation exchange capacity; S: sulfur; B: boron; Cu: copper; Fe: iron; Mn: manganese; and Zn: zinc.

Table 2. Physical properties of the experimental soil.

Soil Texture	Sand	Silt	Clay
	g/kg		
Medium	631	98	271

The lateral rows were considered border rows, while the inner rows were longitudinally divided into 6 treatment groups (T0 to T5), with 4 plants per group, totaling

24 experimental plants. Each treatment group received a different dosage of nitrogen in the form of urea.

The optimal nitrogen dose (11 g of nitrogen per plant) was determined via local soil analysis and based on the recommendations of Spósito et al. [5]. To create varying conditions of both nitrogen deficiency and excess, two treatment groups received nitrogen dosages lower than the optimal amount, two groups received higher than optimal dosages, one group received the optimal dosage, and one group served as the control. Thus, the 24 experimental plants were distributed across the six treatments groups as follows: T0 (no nitrogen added), T1 (25% of the optimal nitrogen dose), T2 (50%), T3 (100%), T4 (150%), and T5 (200% of the optimal level). These specific percentages are commonly used in nitrogen nutrition experiments in the literature [25,26]. Other factors such as training, irrigation, soil composition, ventilation, and radiation were controlled and standardized.

Maintenance fertilization was applied equally to all treatments and borders following the recommendations of Spósito et al. [5], and 50 g/plant of NPK 10-10-10 was applied once a month.

2.3. Obtaining Images of Hop Leaves

Hop leaf collection for image processing (DIP) purposes was conducted in two distinct periods. The first collection took place at the beginning of 2022, 15 days after fertilization. In this case, three leaves from the upper third of each plant were selected and photographed using a mobile phone inside a dark chamber with known dimensions, illuminated by a LED light, generating images with a resolution of 960×1280 pixels, i.e., approximately 1.2 megapixels (MP). The hop leaves were photographed alongside a white square (2 cm per side), used as a reference for DIP calculations. In total, 72 images were generated and saved in JPEG format.

The use of a mobile phone camera was intentional, aiming to evaluate the feasibility of using an accessible and cost-effective imaging system for nitrogen classification in hop plants. Mobile devices are widely available and have been increasingly employed in precision agriculture due to their portability, ease of use, and capacity to support real-time monitoring applications.

The second collection took place one year after the first. In this instance, six leaves from the upper third of each plant were collected. If a plant had not reached maturity by the time of collection or had died, leaves were collected from the remaining developed plants within the same treatment, along with three additional leaves from a deficient plant. After collection, the leaves were scanned using an ECOSYS M6295cidn printer (Kyocera Document Solutions Inc., Osaka, Japan), together with a white background and a black reference square (5 cm long), generating images with a resolution of 4960×7014 pixels (~34.8 MP), totaling 150 images, which were saved in JPEG format.

2.4. Laboratory Analysis

For laboratory nutritional analysis, 10 leaves per plant were collected, totaling 40 leaves per treatment, in both the first and second collections. The leaves designated for nutritional analysis were weighed, washed with neutral soap and deionized water, and subsequently dried in an oven at 60 degrees Celsius for approximately 20 h. The samples were then ground, placed in monolucid paper bags, and sent for macro- and micronutrient analysis following the methodology described in [27].

The laboratory data showed median nitrogen concentrations of 24 g.kg^{-1} for the first collection and 20 g.kg^{-1} for the leaves from the second collection. To identify nitrogen-deficient leaves, the lowest concentration was selected as the threshold for classifying the samples in terms of the amount of nitrogen, with the images of leaves with a nitrogen

concentration of up to 20 g.kg^{-1} being allocated to group 1 and the remaining leaves to group 2. The choice of the lowest median for classification is justified by the lack of a specific nitrogen concentration value indicative of a deficiency in hop leaves. Figure 1 shows examples of leaf images from each group taken by mobile phone and printer.

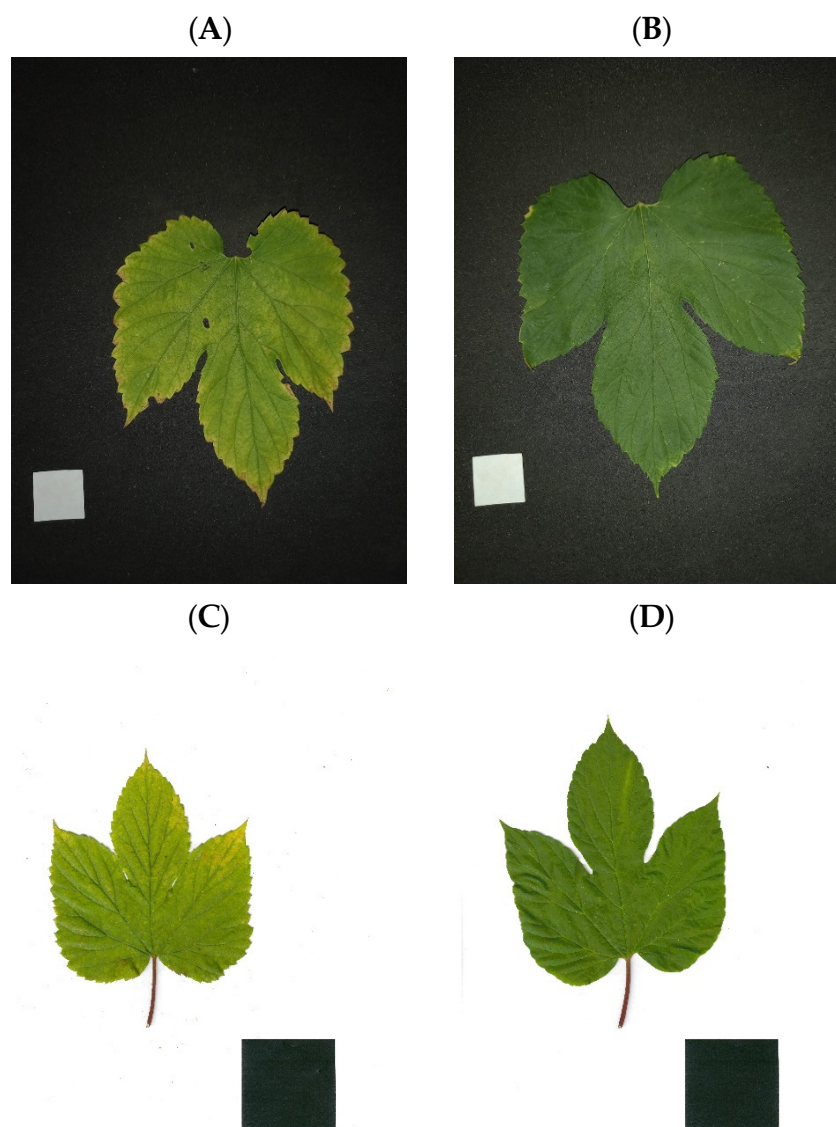


Figure 1. Images of leaves allocated to group 1 (A) and group 2 (B) obtained using a mobile phone, and leaves allocated to group 1 (C) and group 2 (D), obtained using a scanner.

2.5. Image Processing

The images were processed in the R computational environment, version 4.2.2 [28], using the *pliman* package (version 2.2.0) [29] to isolate the leaf area of hop plants and extract information through techniques such as enhancement, outlining, filtering, manipulation, and segmentation. The images were analyzed in terms of RGB color components, including pixel histograms, color matrices, and vegetative indices of interest. Thirteen quantitative variables were obtained: leaf area (A), perimeter (P), maximum and minimum diameters (Dmax and Dmin), mean values of red (R), green (G), and blue (B) color matrices, mean values of normalized color matrices (NR, NG, and NB), as well as the green–red ratio index (GRRI), modified photochemical reflectance index (MPRI), and percentage of yellow–brown area (YBA). The equations for obtaining the NR, NG, NB, GRRI, and MPRI indexes are presented in Equations (1)–(5), respectively.

$$NR = \frac{R}{R + G + B} \quad (1)$$

$$NG = \frac{B}{R + G + B} \quad (2)$$

$$NB = \frac{G}{R + G + B} \quad (3)$$

$$GRRR = \frac{G}{R} \quad (4)$$

$$MPRI = \frac{R - G}{R + G} \quad (5)$$

The MPRI variable normalizes the distance between the R and G colors within a range from -1 to 1 . A value of -1 represents a completely green leaf devoid of red, while a value of 1 represents a leaf that is absent of green and completely red.

To calculate the percentage of yellow–brown leaf area, an adaptation of the `measure_disease` command from the `pliman` package was used. This function was originally developed to identify damaged areas in diseased leaves. A reference color palette was created for the green part of the leaf (Figure 2A), the yellow–brown region (Figure 2B), and the background of the image (Figure 2C).

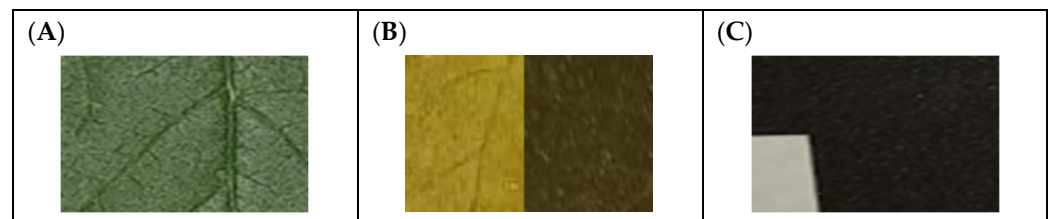


Figure 2. Reference palettes of the green part (A), yellow–brown region (B), and background (C).

In the `pliman` package, the `analyze_objects` command was used to isolate the hop leaf and the reference square as distinct objects, allowing for the extraction of values for the variables `A`, `P`, `Dmax`, `Dmin`, `R`, `G`, `B`, `NR`, `NG`, `NB`, `GRRR`, and `MPRI`. For images obtained using a mobile phone (low resolution), the arguments `index = "G + R - B"`, `tolerance = 900`, and `invert = TRUE`, were empirically defined as the optimal configuration for leaf segmentation and image processing performance, proving to be suitable for most images.

However, in four cases, leaf recognition was inadequate, so it was necessary to define the additional filter and threshold arguments individually. The filter argument applies a median filter of a specific size to remove noise from the image, while threshold sets the cutoff value for converting grayscale images into binary images.

The index argument defines the operation on color matrices that will be the basis for conversion into binary images, and the tolerance argument specifies a minimum object height for recognition. Finally, the invert argument was used to reverse the binary image, which was necessary for mobile phone images due to the black background.

For images obtained using a scanner (high-resolution images), the arguments used were `filter = 10`, `index = "B"`, `tolerance = 900`, and `threshold = 0.6`. The results were stored in two `xlsx` files: one for mobile phone images and another for scanner images. In these files, each row corresponds to an image, with the first column listing the file name, the last column indicating the group classification according to the nutritional analysis of the nitrogen concentration (group 1 or group 2), and the remaining columns containing the values of the evaluated variables.

2.6. Data Analysis

2.6.1. Exploratory Analysis

An initial exploratory analysis was conducted on the data obtained after image processing. The data on the variables were summarized using descriptive statistics (mean, standard deviation, median, and minimum and maximum value), based on the image source (mobile phone or scanner) and nitrogen concentration group (group 1 or group 2), as established from the threshold given by the median nitrogen concentration from the laboratory analysis. For mobile phone images, 12 photos belonged to group 1 and 60 to group 2, totaling 72 images. For scanner images, 72 belonged to group 1 and 78 to group 2, totaling 150 images.

In addition to descriptive statistics, a *t*-test was used to compare the groups within each image source, considering the F-test for homoscedasticity and a significance level of 5% ($p < 0.05$). Correlation matrices were also obtained for both groups using the *Corrplot* package in R (version 0.92) [30], for both the mobile phone and printer data, to visually assess the associations between variables and identify potential variables for exclusion in the development of discriminant models. In addition, correlation coefficients were also used to evaluate the relationship between the nitrogen concentration and the variables obtained from the DIP for each database.

2.6.2. Variable Selection and Discriminant Models

Following the exploratory analysis, the models were estimated to classify images according to the nitrogen concentration in hop leaves.

First, it was necessary to define which explanatory variables would be used in the models. To do this, two sets of preliminary variables were defined. The first set (set 1) included variables that exhibited statistically significant differences between groups 1 and 2, while the second set (set 2) included all variables except the normalized color variables NR, NG, and NB. This was carried out for both mobile phone and scanner image data.

Additional selection was performed based on the multicollinearity of the variables, using the variance inflation factor (VIF). A high VIF value for a variable indicates a strong correlation with the other explanatory variables, suggesting multicollinearity, which can increase uncertainty and complicate model interpretation. A commonly used cut-off point in the literature is 10, above which multicollinearity is considered problematic [20]. A possible solution to this is to exclude the variable with the highest VIF, which should not significantly affect the model, as this variable adds redundant information in the presence of the other variables [31].

Thus, from the sets of preliminary variables, the final variables were selected using a loop as follows: the VIF value for each explanatory variable was calculated, the variable with the highest VIF was excluded, then the VIF values for the remaining variables were calculated again, and the variable with the highest VIF was excluded, repeating this cycle until all the variables had a VIF value < 10 . Therefore, two sets of final variables were obtained for use in building the models (two sets for mobile phone images and two for scanner images). The VIF values were calculated using the R *faraway* package (version 1.0.8) [32].

The linear discriminant analysis (LDA) and quadratic discriminant analysis (QDA) models were estimated using the *MASS* package (version 7.3-60) [33]. LDA is a discriminant model that assumes that the classes have multivariate normal distributions and that the variance and covariance matrices are equal between the classes, finding a linear hyperplane that maximizes the separation between the classes. QDA, on the other hand, assumes that the matrices are different between the groups and uses a quadratic function to maximize the separation between them [34]. Thus, the covariance and sample variance matrices of the

groups were compared using the Box's M test to check for possible significant differences ($p < 0.05$). This was carried out for both the mobile image data and scanner data.

With the data from the mobile phone images (low-resolution images), an LDA model and a QDA model were built for each set of variables, totaling four models: two LDA models and two QDA models. Similarly, with the data from the scanner images (high re-resolution), an LDA model and a QDA model were built for each set of variables, totaling four models: two LDA models and two QDA models. The prior probabilities used were 0.5 for both groups. The general flowchart of the methodology is shown in Figure 3.

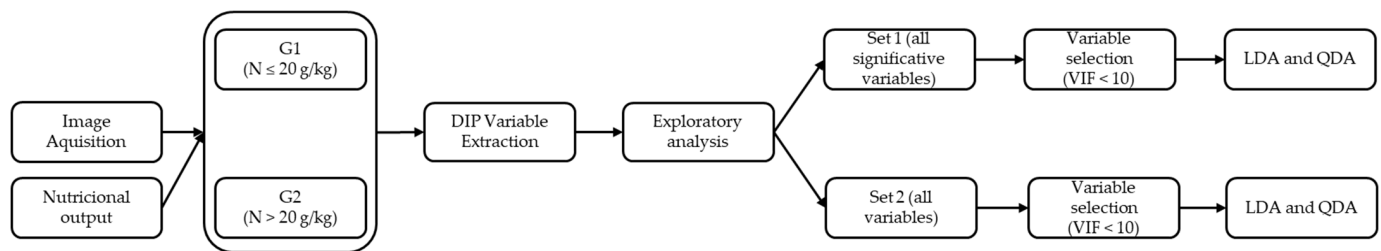


Figure 3. General methodology flowchart for group division and data analysis.

2.7. Performance Evaluation

The performance metrics used to evaluate the image classification models include true positive (TP), true negative (TN), false positive (FP), and false negative (FN). In the context of the models developed, a true positive (TP) is considered when an image of a leaf belonging to group 1 is correctly classified by the model in that group. Similarly, a true negative (TN) is obtained when an image of a leaf belonging to group 2 is correctly classified in that group. These results indicate that the model's diagnosis is aligned with the actual condition of the leaves [35].

On the other hand, a false positive (FP) occurs when the image of a leaf in group 2 is incorrectly classified as belonging to group 1. Conversely, a false negative (FN) occurs when an image of a leaf in group 1 is incorrectly classified as belonging to group 2 [35]. Thus, in terms of TP, TN, FN, and FP, Equations (6), (7), (8), and (9) refer, respectively, to the sensitivity, specificity, accuracy, and balanced accuracy of the models evaluated.

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (6)$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (7)$$

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (8)$$

$$\text{Balanced accuracy} = \frac{\text{Sensitivity} + \text{Specificity}}{2} \quad (9)$$

Thus, sensitivity refers to the chance of correctly classifying leaves belonging to group 1, specificity refers to the chance of correctly classifying leaves belonging to group 2, and accuracy refers to the model's overall success rate. Balanced accuracy is the arithmetic mean of sensitivity and specificity and is a more appropriate indicator than accuracy when the data have unbalanced groups [36]. These indicators were calculated in the R environment using the caret package [37].

Finally, the receiver operating characteristic (ROC) curve was generated for each model, and the area under the ROC curve (AUC) was calculated. In total, eight ROC curves were generated. For LDA and QDA models, the ROC curve is obtained by varying the

posterior probability threshold required to classify an individual into one of the groups and calculating the resulting true positive and false positive rates.

The result is a curve in which the vertical axis expresses sensitivity and the horizontal axis relates to specificity. Therefore, the AUC value is a summary measure of the model's performance across all thresholds, and ranges from 0 to 1. A perfect classifier has AUC = 1, while a random classifier has AUC = 0.5 [31]. The *pROC* package (version 1.18.5) was used to generate the ROC curves and calculate the AUC [38].

2.8. Cross-Validation

The performance of the models was analyzed via cross-validation with the leave-one-out (LOO) method, using the MASS package. The LOO method is a specific case of *k-fold* cross-validation when the number of folds (*k*) is equal to the number of individuals in the data set (*n*), i.e., when *k* = *n*. Thus, the data set is divided into *n* folds, with each fold containing one individual. At each iteration, one individual is separated to serve as validation, the model is trained on the remaining individuals, and it is tested only on the separated individual. This process is repeated until all individuals have been used once to validate the model, totaling *n* iterations. Therefore, 72 iterations were carried out for mobile phone data, while 150 iterations were carried out for scanner data. One of the main advantages of the LOO method is that it generates practically unbiased performance indicators, but it can be computationally intensive when *n* is very large [39].

3. Results

The results of the descriptive summary of the DIP data and *t*-test show that the variables G and NB exhibited a significant difference between the means of the two groups ($p < 0.05$) for both image sources. Additionally, the variables R, NR, GRRI, and MPRI also showed a difference between the groups for the data obtained from the scanner, while the B variable only showed a difference for the data obtained from the mobile phone (Table 3). This suggests that spectral characteristics, particularly those related to colors and vegetation indices, respond differently depending on the image acquisition method, possibly due to variations in light capture and image quality between the mobile phone and the printer.

Table 3. Mean (standard deviation) [first row], median [second row], and minimum–maximum value [third row] of the variables measured in each image type group.

Variables	Mobile Phone Images			Scanner Images		
	G1 (n ₁ = 12)	G2 (n ₂ = 60)	<i>p</i>	G1 (n ₁ = 78)	G2 (n ₂ = 72)	<i>p</i>
A	66.2 (26.1) 70.9 (26.4–99.1)	64.7 (19.0) 60.5 (21.1–107.9)	0.81 *	67.6 (21.9) 63.3 (31.1–123.0)	63.0 (21.1) 61.9 (25.3–104.6)	0.19 *
P	57.1 (14.4) 60.8 (30.1–75.4)	59.6 (10.0) 59.7 (31.3–92.0)	0.47 *	45.7 (9.0) 45.7 (29.1–65.4)	43.7 (9.0) 43.8 (19.9–60.2)	0.19 *
Dmin	2.8 (0.6) 2.7 (1.9–4.3)	2.6 (0.9) 2.4 (1.3–6.1)	0.51 *	3.5 (1.2) 2.9 (1.7–6.4)	3.2 (1.2) 3.2 (1.7–6.4)	0.36 *
Dmax	12.4 (2.7) 13.2 (7.7–15.4)	12.3 (1.6) 12.5 (7.6–16.5)	0.87 †	12.8 (2.1) 13 (8.7–17.5)	12.4 (2.1) 12.2 (7.1–17.1)	0.29 *

Table 3. Cont.

Variables	Mobile Phone Images			Scanner Images		
	G1 (n ₁ = 12)	G2 (n ₂ = 60)	<i>p</i>	G1 (n ₁ = 78)	G2 (n ₂ = 72)	<i>p</i>
R	0.400 (0.073) 0.37 (0.341–0.561)	0.372 (0.049) 0.362 (0.309–0.603)	0.21 [†]	0.438 (0.106) 0.403 (0.309–0.782)	0.373 (0.074) 0.38 (0.119–0.496)	<0.01 [†]
G	0.465 (0.030) 0.461 (0.432–0.506)	0.441 (0.022) 0.438 (0.410–0.514)	<0.01 [*]	0.506 (0.074) 0.499 (0.384–0.715)	0.463 (0.082) 0.477 (0.173–0.620)	<0.01 [*]
B	0.145 (0.026) 0.141 (0.110–0.189)	0.174 (0.033) 0.176 (0.105–0.277)	<0.01 [*]	0.076 (0.026) 0.071 (0.033–0.134)	0.079 (0.026) 0.074 (0.031–0.147)	0.44 [*]
NR	0.394 (0.042) 0.376 (0.360–0.482)	0.376 (0.032) 0.372 (0.326–0.489)	0.08 [*]	0.425 (0.031) 0.42 (0.376–0.520)	0.403 (0.040) 0.405 (0.268–0.497)	<0.01 [†]
NG	0.463 (0.025) 0.466 (0.417–0.502)	0.450 (0.021) 0.452 (0.401–0.501)	0.06 [*]	0.499 (0.019) 0.503 (0.441–0.531)	0.505 (0.033) 0.511 (0.397–0.548)	0.17 [†]
NB	0.142 (0.03) 0.141 (0.100–0.191)	0.174 (0.03) 0.178 (0.102–0.194)	<0.01 [*]	0.076 (0.033) 0.069 (0.028–0.153)	0.092 (0.062) 0.077 (0.024–0.332)	0.05 [†]
GRR1	1.193 (0.154) 1.245 (0.867–1.340)	1.209 (0.121) 1.216 (0.835–1.436)	0.69 [*]	1.182 (0.095) 1.203 (0.850–1.300)	1.265 (0.112) 1.322 (0.902–1.499)	<0.01 [*]
MPRI	−0.082 (0.072) −0.108 (−0.145–0.072)	−0.091 (0.052) −0.097 (−0.177–0.091)	0.62 [*]	−0.080 (0.043) −0.091 (−0.129–0.082)	−0.113 (0.047) −0.125 (−0.197–0.058)	<0.01 [*]
YBA	17.7 (35.8) 2.2 (1.1–97.9)	7.3 (19.3) 1.3 (0.4–100.0)	0.35 [†]	4.5 (12.2) 1.3 (0.6–94.8)	4.2 (13.5) 0.7 (0.2–68.2)	0.87 [*]

p-values superscripted with * are *t*-test results, while those superscripted with † indicate that the nonparametric Mann–Whitney test was used.

It was also found that the nitrogen concentration did not show strong individual correlations with the DIP variables (Table 4). The highest correlation coefficients with nitrogen concentration were found for the following variables: leaf area ($r = -0.37$), perimeter ($r = -0.35$), mean R and NR ($r = -0.35$), as well as MPRI ($r = -0.36$). This pattern is consistent with the correlations observed in the image data matrices.

Table 4. Correlations between nitrogen concentration and DIP variables.

Databases	DIP Variables												
	A	P	Dmin	Dmax	R	G	B	NR	NG	NB	GRR1	MPRI	YBA
Scanner	0.19	0.18	0.23	0.07	0.02	−0.06	−0.18	0.09	0.07	−0.14	−0.03	0.04	−0.03
Mobile Phone	−0.37	−0.35	−0.2	−0.3	−0.35	−0.29	0.17	−0.35	0.04	0.24	0.4	−0.36	−0.03

By analyzing the Pearson correlation matrices of both databases (Figures 4 and 5), it was observed that some pairs of variables exhibited very high correlation coefficients in both groups, indicating that not all variables are necessary for adjusting the discriminant models. Measures A and P showed a high correlation with each other and with Dmax

and Dmin for both mobile phone and scanner images, respectively. A high correlation was also observed between the GRRRI and MPRI leaf indices and the colors R and G, a fact directly associated with the calculation methodology of those indices, which uses the R and G variables for value composition. The strong correlation between the variable R and the variable YBA is due to the red color reflection caused by chlorosis and the damaged area.

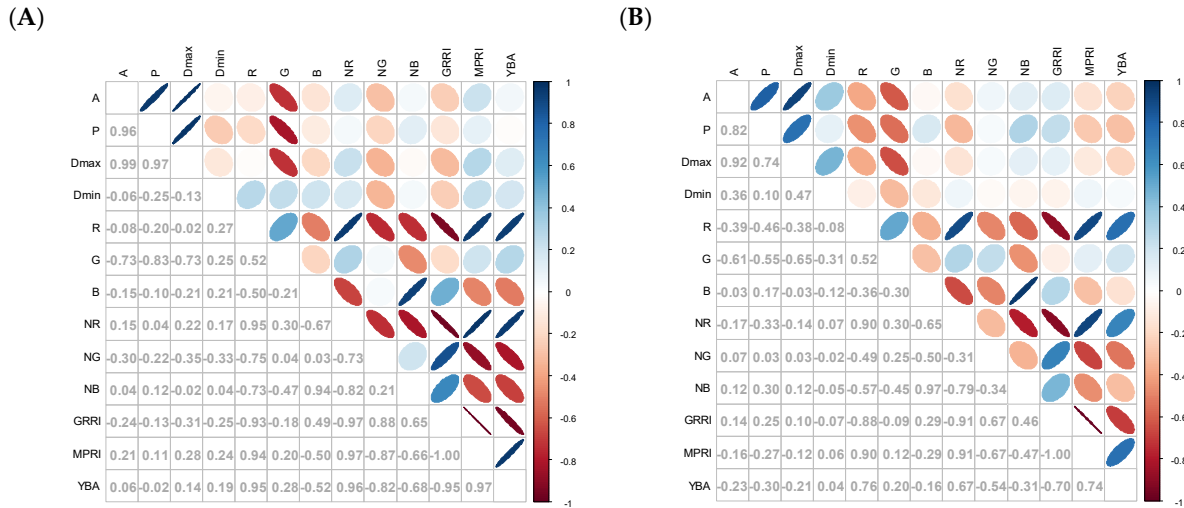


Figure 4. Correlation matrices of mobile phone image data: Group 1 (A) and 2 (B).

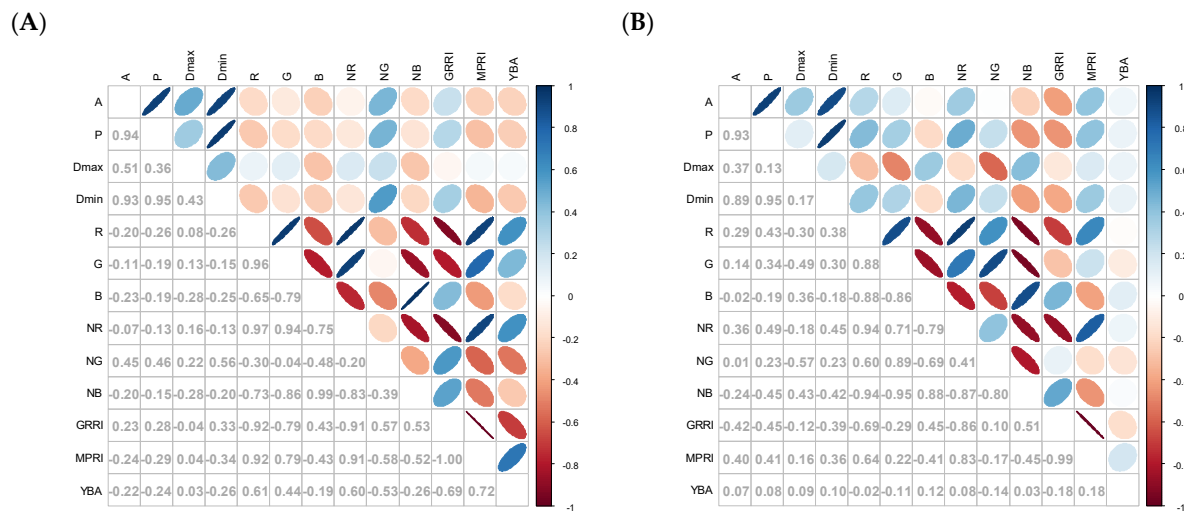


Figure 5. Correlation matrices of scanner image data: Group 1 (A) and 2 (B).

Concomitantly with the correlation matrices, the analysis and selection of variables through VIF were performed. For the set derived from the significant variables of the mobile phone images, the remaining variables were G and B (VIF = 1.147). However, for the data obtained from printer images (R, G, NB, GRRRI, and MPRI), the remaining variables were G, NB, and GRRRI (with VIF values of 5.051; 4.890 and 1.434, respectively), since NR, MPRI, and GRRRI depend on the R variable for their calculation.

In set 2, which included all variables except the highly correlated ones, the remaining variables after VIF analysis were similar for both databases, mobile phone and scanner images. The only exception for the mobile phone data was that perimeter remained (VIF values: P: 4140; Dmin: 1.403; Dmax: 5.007; G: 2.347; B: 1.523; GRRRI: 2.622; and YBA: 2.603), whereas for the scanner data, the area remained (VIF values: A: 7.949; Dmin: 7.185; Dmax: 1.475; G: 3.708; B: 3.255; GRRRI: 1.833; and YBA: 1.236). As observed in the correlation matrices, the two variables (P and A) exhibit a high correlation with each other.

Furthermore, the results of Box’s M test indicated that the variance–covariance matrices of each group, for the set of significant variables obtained from the mobile phone images, can be considered equal ($p = 0.299$). In contrast, for the set of variables selected using VIF, the variance–covariance matrices were different ($p < 0.05$). In the case of scanner images, the matrices of the groups were different in both sets of variables. When the variance–covariance matrices are equal between groups, LDA models are more suitable. Conversely, when the matrices are different, QDA is the most appropriate approach.

Based on this, eight models were estimated, and their performance indicators are presented in Table 5.

Table 5. Model quality indicators for each set of cross-validation predictors.

Database	Set of Variables	Model	Performance Indicators				
			Sen. (%)	Sp. (%)	Ac. (%)	Bal. Ac. (%)	AUC (CI)
Mobile phone (low-resolution images)	G B	LDA	50	78.3	73.6	64.2	0.74 (0.582–0.904)
		QDA	50	73.3	69.4	61.7	0.72 (0.560–0.882)
	P Dmax Dmin	LDA	58.3	81.7	77.8	70	0.73 (0.560–0.893)
	G B GRR1 YBA	QDA	33.3	88.3	79.2	60.8	0.70 (0.544–0.862)
Scanner (high-resolution images)	G NB GRR1	LDA	56.9	84.6	71.3	70.8	0.76 (0.683–0.838)
		QDA	72.2	75.6	74	73.9	0.85 (0.792–0.911)
	A Dmax Dmin	LDA	68.1	69.2	68.7	68.6	0.76 (0.684–0.837)
	G B GRR1 YBA	QDA	79.2	71.8	75.3	75.5	0.84 (0.778–0.902)

Sen: sensitivity; Sp: specificity; Ac: accuracy; Bal. Ac: balanced accuracy; and AUC: area under the ROC curve, with confidence interval (CI) with significance of 95%.

By comparing the models from both image bases (mobile phone and scanner), it is evident that the scanner models (high-resolution images) generally exhibited greater sensitivity and balanced accuracy, while the mobile phone models (low resolution) presented slightly higher specificity. This result highlights that images with better resolutions favor the development of robust models and consequently contribute to better nitrogen content classification.

Regarding the ROC curves (Figures 6 and 7), all curves are above the line of chance (diagonal), indicating that the models have discriminative power. Additionally, the proximity of the curves to the upper left corner indicates a good balance between sensitivity and specificity.

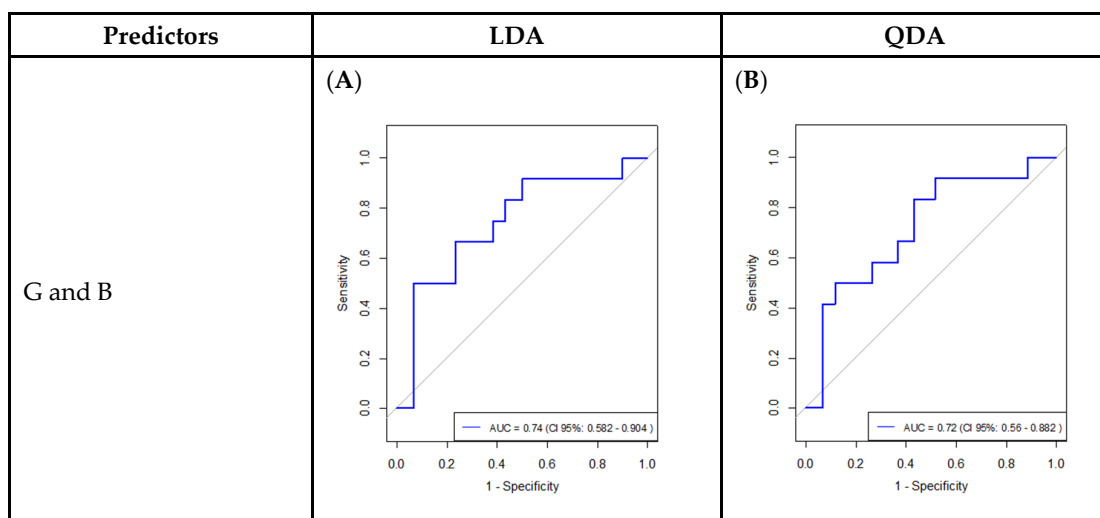


Figure 6. Cont.

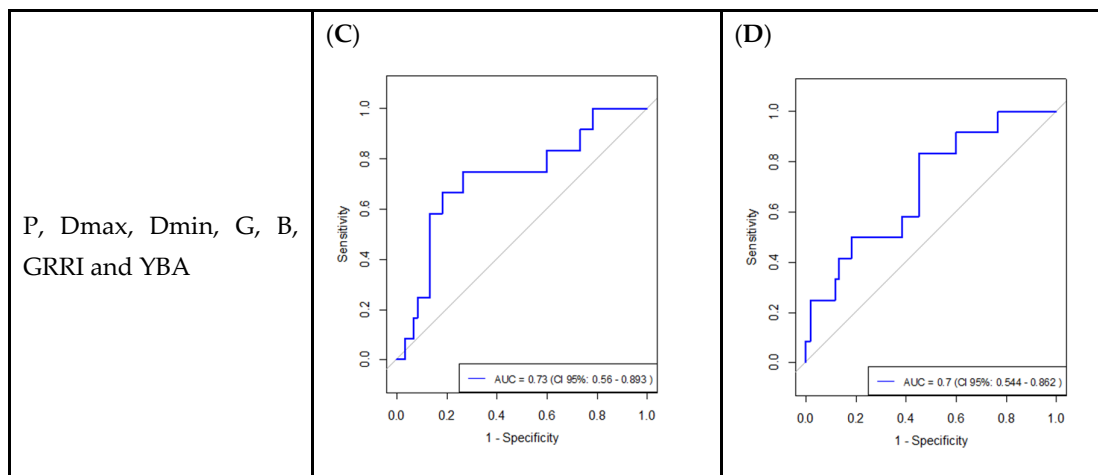


Figure 6. ROC curves of the models with predictor of mobile phone photos (low resolution). (A) and (B) show the ROC curves for the Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA) models, respectively, when using the predictors G and B. (C,D) display the ROC curves for LDA and QDA, respectively, using the predictor set: P, Dmax, Dmin, G, B, GRRI, and YBA.

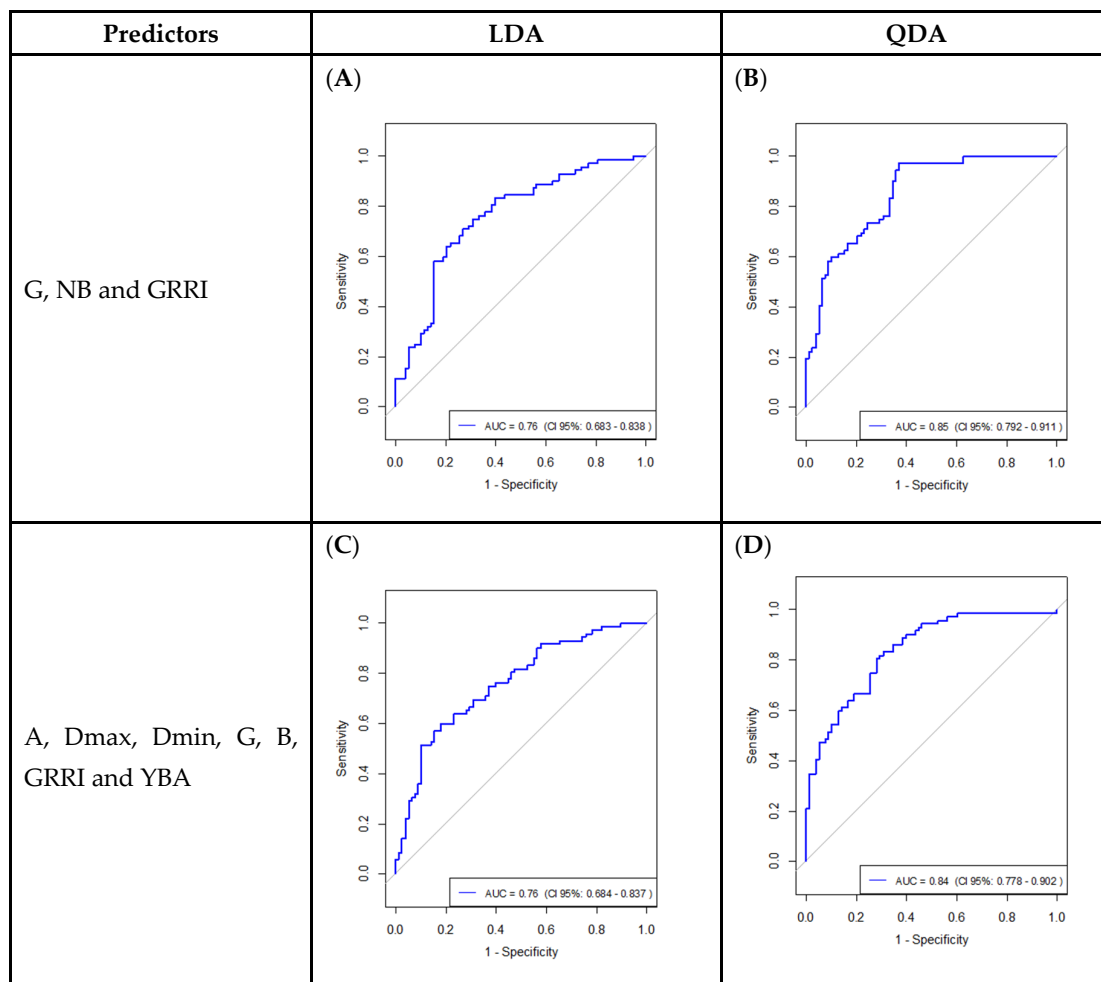


Figure 7. ROC curves of the models with predictor of scanner photos (high resolution). (A,B) show the ROC curves for the Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA) models, respectively, when using the predictors G, NB, and GRRI. (C,D) display the ROC curves for LDA and QDA, respectively, using the predictor set: A, Dmax, Dmin, G, B, GRRI, and YBA.

Figure 6A,B presented slightly higher AUCs (0.74 and 0.72, respectively) compared to the curves in Figure 6B,D (AUC equal to 0.73 and 0.70), suggesting a potentially better classification performance for models generated from set 1, i.e., only predictor variables with significant differences between the groups. However, the confidence intervals (CI) for the AUCs overlap in all situations, indicating that the differences are not statistically significant. The models generated from the data obtained from the higher-resolution images (Figure 7) demonstrated overall superior performance compared to the models estimated using mobile phone data (lower resolution), both in terms of proximity to the upper left corner and relatively higher AUC values.

4. Discussion

One of the visual symptoms of a decreased nitrogen concentration in leaves is chlorosis, the insufficient production of chlorophyll by the plant, which results in a color transition from light green to yellow, depending on the severity [40]. Thus, this factor may be related to the significance of the green color in both databases. Similarly, in some plant species, such as tomato (*Solanum lycopersion*), cauliflower (*B. oleracea* var. botrytis), cabbage (*Brassica oleracea*), and *Arabidopsis thaliana*, nitrogen deficiency induces anthocyanin production, which may be associated with the significance of blue and red colors in the average of the groups [41].

For the low-resolution image database (mobile), LDA models demonstrated higher sensitivity and balanced accuracy compared to QDA models. Additionally, as evidenced by the Box's M test, the LDA models outperformed the QDA models, with accuracies of 73.4% and 69.4%, respectively, for the model using only the significant variables as predictors. In Alajas et al.'s study [15], LDA also achieved the highest accuracy in classifying the health status of grape leaves among the models studied, with an accuracy of 98.99%.

Furthermore, in a study involving classification methods for identifying and quantifying damaged pixels in leaves, the linear discriminant analysis classifier demonstrated the best performance, with an average precision of 95%, in addition to being a computationally efficient and robust model [42]. Similarly, the LDA model for the mobile database starting from set 1 exhibited a considerable classification capacity for leaves with a nitrogen concentration above the threshold, using only two precursor variables (G and B).

In another study on the early detection of excess nitrogen application in tomato leaves using a non-destructive hyperspectral imaging system, when comparing the performance of nine machine classifiers, LDA showed the second-lowest performance, with a correct classification rate (CCR) of 85.5% [43]. However, this rate was still close to that of other classifiers with superior performance, including hybrid artificial neural networks, independent component analysis, harmony search, the bee algorithm, and deep learning-based classifiers using convolutional neural networks, which had CCRs between 88.8% and 91.6%. Additionally, LDA had a CCR 27% higher than the other classic supervised classifier, support vector machines (SVMs).

In the model adjusted using the predictors selected using the VIF from the set containing all the predictors (set 2), QDA showed higher accuracy than LDA, but lower balanced accuracy, due to differences in group sizes between the two groups. For the higher-resolution data (scanner), QDA models performed better in both cases, except for specificity in the models for set 1, with 84.6% for LDA compared to 75.6% for QDA. High-resolution images significantly enhance insect detection models' accuracy, as they provide better detail and clarity, which is crucial for identifying small pests such as whiteflies on tomato leaflets. On the other hand, low-resolution images reduce accuracy, highlighting the importance of training models on images that match the data resolution. Models trained on mixed-resolution datasets performed comparably to those trained exclusively

on high- or low-resolution images, suggesting that incorporating multiple resolutions can be beneficial [44,45].

Both databases yielded models with classification rates for group 2 exceeding 80% specificity, the highest being 88.3% for the lower-resolution (mobile) images and 84.6% for the higher-resolution (scanner) images. Additionally, for higher-resolution images, the QDA model generated from set 2 exhibited a classification rate of 71.8% for leaves with nitrogen levels above the threshold. The study by Jahanbakhshi and Kheiralipour [17] also reported a higher overall accuracy rate for the quadratic discriminant analysis model.

On the other hand, most models had lower sensitivity than specificity, especially those trained on lower-resolution (mobile) images. No such model achieved a sensitivity higher than 60%, and the QDA model generated from set 2 achieved a mere 33.3% sensitivity. Thus, while LDA and QDA models based on mobile images could be useful for detecting higher nitrogen concentrations in the field, their lower sensitivity limits the detection of nitrogen deficiency. However, it is important to note that the image resolution of the mobile phone used in the study is very low, at only 1.2 MP. Most modern smartphones have higher-resolution cameras (e.g., iPhone 15 with 48 MP), with some recent high-end models reaching 200 MP (Samsung Galaxy S25 Ultra). These are much higher than the 34.8 MP scanner used in this study. Therefore, advancements in smartphone camera technology could significantly enhance the sensitivity and overall performance of these models in the future.

In Pathak et al.'s study [46], LDA and QDA models were used to reduce the number of features in plant leaf images, followed by classifiers such as SVM, KNN, and logistic regression for classifying a specific image into a particular plant class. In this case, QDA performed better when applied to classes with different variances.

In oil palm cultivation, by using leaf-scale and machine learning approaches, the performance of discriminant analysis models and SVM was compared in terms of defining nitrogen profiles [47]. Discriminant analysis identified a greater number of optimal spectral bands, and could more accurately discriminate nitrogen levels compared to the other model. However, SVM could achieve reasonable accuracy with fewer spectral bands, and when combined with leaf measurement, it presented the best discriminant function in the study.

In this context, the performance of set 1 models is noteworthy. Despite containing only three variables, they exhibited an AUC equal to or superior to those set 2 models, which contained seven variables. Furthermore, the QDA model in set 1 achieved the highest AUC of all the models evaluated (0.85), indicating superior performance and greater classification power for the images. However, compared to previous studies in the literature focused on disease classification through leaf analysis, both the AUC and accuracy values were lower, with reported accuracy values of 94.35% and AUC values of 94.7% [48].

5. Conclusions

Thus, it can be concluded that image quality had a direct impact on hop leaf classification regarding nitrogen concentration, with high-definition images improving model performance, achieving AUC values ranging from 0.76 to 0.85 and sensitivity between 56% and 79.2%, while images of lower resolution, though slightly inferior (AUC between 0.70 and 0.74), still provided promising classification results. The quadratic discriminant analysis (QDA) model exhibited the best performance, standing out even when using a reduced number of explanatory variables and achieving superior AUC, indicating good discriminative capacity, although still inferior to other studies in the literature focused on disease and nutritional classification in plant leaves.

The key contribution of this study is demonstrating that accurate nitrogen classification can be achieved using a simplified set of image-derived features and a straightforward model, enabling easy application while minimizing the need for complex preprocessing.

For future research, integrating data from different image sources, including additional variables such as spectral and morphological characteristics, and evaluating advanced classifiers, such as neural networks and support vector machines (SVM), are suggested, as they may offer complementary alternatives in leaf nutritional analysis. Additionally, the development of practical tools, such as applications for nutritional classification, could expand the applicability of the results in hop cultivation management in Brazil.

6. Patents

Based on the acquired knowledge, there is a clear intention to develop an intuitive digital application. This application would function as a tool for hop producers, primarily for family farming, integrating image capture via mobile phones with the developed models, enabling the rapid and practical retrieval of nutritional information about the plant. The development of practical tools, such as applications for nutritional classification, could expand the applicability of the results in hop cultivation management in Brazil.

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

Funding: This research was financially supported by the National Council for Scientific and Technological Development (CNPq), grant number 129529/2022-2, the Coordination for the Improvement of Higher Education Personnel (CAPES), funding code 001, and São Paulo State University (UNESP) through resources from the Institutional Development Plan (PDI 2025) and the Agricultural Engineering Program.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available upon request from the corresponding author due to their proprietary nature and the authors' intention to use them in the development of a software application.

Acknowledgments: The authors would like to express their gratitude to the School of Agricultural Sciences of São Paulo State University (FCA/UNESP), Botucatu, for the administrative and technical support, as well as for providing the necessary infrastructure for the development of this study. Additionally, we acknowledge the Coordination for the Improvement of Higher Education Personnel (CAPES) and the National Council for Scientific and Technological Development (CNPq).

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

References

1. Almaguer, C.; Schönberger, C.; Gastl, M.; Arendt, E.K.; Becker, T. *Humulus lupulus*—A Story That Begs to Be Told. A Review. *J. Inst. Brew.* **2014**, *120*, 289–314. [[CrossRef](#)]
2. Chadwick, L.R.; Pauli, G.F.; Farnsworth, N.R. The Pharmacognosy of *Humulus lupulus* L. (Hops) with an Emphasis on Estrogenic Properties. *Phytomedicine* **2006**, *13*, 119–131. [[CrossRef](#)] [[PubMed](#)]
3. Zanolini, P.; Zavatti, M. Pharmacognostic and Pharmacological Profile of *Humulus lupulus* L. *J. Ethnopharmacol.* **2008**, *116*, 383–396. [[PubMed](#)]

4. Arruda, T.R.; Pinheiro, P.F.; Silva, P.I.; Bernardes, P.C. A New Perspective of a Well-Recognized Raw Material: Phenolic Content, Antioxidant and Antimicrobial Activities and α - and β -Acids Profile of Brazilian Hop (*Humulus lupulus* L.) Extracts. *LWT* **2021**, *141*, 110905. [CrossRef]
5. Spósito, M.B.; Ismael, R.V.; Barbosa, C.M.d.A.; Tagliaferro, A.L. *A Cultura Do Lúpulo*; Biblioteca ESALQ: Piracicaba, Brazil, 2019; Volume 68.
6. A prolúpulo—Associação Brasileira de Produtores de Lúpulo. Available online: <https://www.aprolupulo.com.br/noticia/lupulo-e-incluido-no-programa-moderagro-e-ganha-impulso-para-acelerar-crescimento> (accessed on 20 June 2024).
7. Burgess, A.H. *Hops: Botany, Cultivation, and Utilisation*; CABI: Oxfordshire, UK, 1964.
8. Faquin, V. *Nutrição Mineral de Plantas*; Universidade Federal de Lavras: Lavras, Brazil, 2005.
9. Machado, M.P.; Gomes, E.N.; Francisco, F.; Bernert, A.F.; Bessalho Filho, J.C.; Deschamps, C. Micropropagation and Establishment of *Humulus lupulus* L. Plantlets Under Field Conditions at Southern Brazil. *J. Agric. Sci.* **2018**, *10*, 275. [CrossRef]
10. Iskra, A.E.; Lafontaine, S.R.; Trippe, K.M.; Massie, S.T.; Phillips, C.L.; Twomey, M.C.; Shellhammer, T.H.; Gent, D.H. Influence of Nitrogen Fertility Practices on Hop Cone Quality. *J. Am. Soc. Brew. Chem.* **2019**, *77*, 199–209. [CrossRef]
11. Čeh, B. Impact of Slurry on the Hop (*Humulus lupulus* L.) Yield, Its Quality and n-Min Content of the Soil. *Plant Soil Environ.* **2014**, *60*, 267–273. [CrossRef]
12. Cantarella, H. *Fertilidade Do Solo*; Novais, R.F., Victor Hugo Alvarez, V., De Barros, N.F., Lúcio Fontes, R., Cantarutti, R.B., Neves, J.C.L., Eds.; Sociedade Brasileira de Ciência do Solo: Viçosa, Brazil, 2007.
13. Pir, R.M.S. Nutrient Deficiency and Syndrome Recognition in Both Mango Leaf and Cotton Plant Using K-Means Clustering and BPNN. *Int. J. Eng. Dev. Res.* **2016**, *4*, 636–642.
14. de Magalhães, L.P.; Trevisan, L.R.; Gomes, T.M.; Rossi, F. Use of Digital Images to Classify Leaf Phosphorus Content in Grape Tomatoes. *Eng. Agric.* **2021**, *42*, e20210147. [CrossRef]
15. Alajas, O.J.; Sybingco, E.; Mendigoria, C.H.; Dadios, E. Prediction of Grape Leaf Black Rot Damaged Surface Percentage Using Hybrid Linear Discriminant Analysis and Decision Tree. In Proceedings of the 2021 International Conference on Intelligent Technologies (CONIT), Hubli, India, 25–27 June 2021; pp. 1–6.
16. Castro, P.H.N.; Moreira, G.J.P.; Luz, E.J.d.S. An End-to-End Deep Learning System for Hop Classification. *IEEE Lat. Am. Trans.* **2022**, *20*, 430–442. [CrossRef]
17. Jahanbakhshi, A.; Kheiralipour, K. Evaluation of Image Processing Technique and Discriminant Analysis Methods in Postharvest Processing of Carrot Fruit. *Food Sci. Nutr.* **2020**, *8*, 3346–3352. [CrossRef] [PubMed]
18. Kheiralipour, K.; Pormah, A. Introducing New Shape Features for Classification of Cucumber Fruit Based on Image Processing Technique and Artificial Neural Networks. *J. Food Process Eng.* **2017**, *40*, e12558. [CrossRef]
19. Pydipati, R.; Burks, T.F.; Lee, W.S. Identification of Citrus Disease Using Color Texture Features and Discriminant Analysis. *Comput. Electron. Agric.* **2006**, *52*, 49–59. [CrossRef]
20. Elmasry, G.; Cubero, S.; Moltó, E.; Blasco, J. In-Line Sorting of Irregular Potatoes by Using Automated Computer-Based Machine Vision System. *J. Food Eng.* **2012**, *112*, 60–68. [CrossRef]
21. Negrisoni, M.M.; Mereb Negrisoni, R.; Nunes da Silva, F.; da Silva Lopes, L.; de Sales de Souza, F., Jr.; Velini, E.D.; Carbonari, C.A.; Rodrigues, S.A.; Raetano, C.G. Soybean Rust Detection and Disease Severity Classification by Remote Sensing. *Agron. J.* **2022**, *114*, 1–17. [CrossRef]
22. Barbedo, J.G.A. Deep Learning for Soybean Monitoring and Management. *Seeds* **2023**, *2*, 340–356. [CrossRef]
23. Pathak, H.; Igathinathane, C.; Howatt, K.; Zhang, Z. Machine Learning and Handcrafted Image Processing Methods for Classifying Common Weeds in Corn Field. *Smart Agric. Technol.* **2023**, *5*, 100249. [CrossRef]
24. Franco, J.R.; Pai, E.D.; Calça, M.V.C.; Raniero, M.R.; Dal Pai, A.; Sarnighausen, V.C.R.; Sánchez-Román, R.M. Update of Climatological Normal and Köppen Climate Classification for the Municipality of Botucatu-SP. *IRRIGA* **2023**, *28*, 77–92. [CrossRef]
25. Ramos, D.P.; Leonel, S. Atributos de qualidade de frutos de tangerineira ‘Poncã’ adubada com composto orgânico, em dois ciclos agrícolas de avaliação. *Sci. Plena* **2014**, *10*, 1–10.
26. Baesso, M.M.; Leveghin, L.; Sardinha, E.J.d.S.; Oliveira, G.P.d.C.N.; de Sousa, R.V. Deep Learning-Based Model for Classification of Bean Nitrogen Status Using Digital Canopy Imaging. *Eng. Agric.* **2023**, *43*, e20230068. [CrossRef]
27. Malavolta, E.; Vitti, G.C.; De Oliveira, S.A. *Avaliação Do Estado Nutricional Das Plantas: Princípios e Aplicações*, 2nd ed.; POTAFOS: Piracicaba, Brazil, 1997; Volume 1.
28. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2023. Available online: <https://www.R-Project.Org/> (accessed on 22 July 2024).
29. Olivoto, T. Lights, Camera, Pliman! An R Package for Plant Image Analysis. *Methods Ecol. Evol.* **2022**, *13*, 789–798. [CrossRef]
30. Wei, T.; Simko, V.; Levy, M.; Xie, Y.; Jin, Y.; Zemla, J. Package “Corrplot”. Visualization of a Correlation Matrix. *Statistician* **2017**, *56*. Available online: <https://cran.r-project.org/web/packages/corrplot/corrplot.pdf> (accessed on 10 September 2024).
31. James, G.; Witten, D.; Hastie, T.; Tibshirani, R. *An Introduction to Statistical Learning—With Applications in R*; Springer: New York, NY, USA, 2013.

32. Faraway, J. Package ‘Faraway’. R Topics Documented 2015. Available online: <https://cran.r-project.org/package=faraway> (accessed on 10 August 2024).
33. Ripley, B.; Venables, B.; Bates, D.M.; Hornik, K.; Gebhardt, A.; Firth, D. MASS: Support Functions and Datasets for Venables and Ripley’s MASS. R Package Version 7.3-60. R Topics Documented 2019. Available online: <https://cran.r-project.org/web/packages/MASS/index.html> (accessed on 28 July 2024).
34. Johnson, R.A.; Wichern, D.W. *Applied Multivariate Statistical Analysis*; Pearson Prentice Hall: Hoboken, NJ, USA, 2007; ISBN 0-13-187715-1.
35. Zhu, W.; Zeng, N.; Wang, N. Sensitivity, Specificity, Accuracy, Associated Confidence Interval and ROC Analysis with Practical SAS[®] Implementations. In Proceedings of the Northeast SAS Users Group 2010: Health Care and Life Sciences, Baltimore, MD, USA, 14–17 November 2010.
36. Thölke, P.; Mantilla-Ramos, Y.J.; Abdelhedi, H.; Maschke, C.; Dehgan, A.; Harel, Y.; Kemtur, A.; Mekki Berrada, L.; Sahraoui, M.; Young, T.; et al. Class Imbalance Should Not Throw You off Balance: Choosing the Right Classifiers and Performance Metrics for Brain Decoding with Imbalanced Data. *Neuroimage* **2023**, *277*, 120253. [[CrossRef](#)] [[PubMed](#)]
37. Kuhn, M. Building Predictive Models in R Using the Caret Package. *J. Stat. Softw.* **2008**, *28*, 1–26. [[CrossRef](#)]
38. Robin, X.; Turck, N.; Hainard, A.; Tiberti, N.; Lisacek, F.; Sanchez, J.C.; Müller, M. PROC: An Open-Source Package for R and S+ to Analyze and Compare ROC Curves. *BMC Bioinform.* **2011**, *12*, 77. [[CrossRef](#)]
39. Fushiki, T. Estimation of Prediction Error by Using K-Fold Cross-Validation. *Stat. Comput.* **2011**, *21*, 137–146. [[CrossRef](#)]
40. de Bang, T.C.; Husted, S.; Laursen, K.H.; Persson, D.P.; Schjoerring, J.K. The Molecular–Physiological Functions of Mineral Macronutrients and Their Consequences for Deficiency Symptoms in Plants. *New Phytol.* **2021**, *229*, 2446–2469.
41. Jezek, M.; Allan, A.C.; Jones, J.J.; Geilfus, C.M. Why Do Plants Blush When They Are Hungry? *New Phytol.* **2023**, *239*, 494–505.
42. Kruse, O.M.O.; Prats-Montalbán, J.M.; Indahl, U.G.; Kvaal, K.; Ferrer, A.; Futsaether, C.M. Pixel Classification Methods for Identifying and Quantifying Leaf Surface Injury from Digital Images. *Comput. Electron. Agric.* **2014**, *108*, 155–165. [[CrossRef](#)]
43. Benmouna, B.; Pourdarbani, R.; Sabzi, S.; Fernandez-Beltran, R.; García-Mateos, G.; Molina-Martínez, J.M. Comparison of Classic Classifiers, Metaheuristic Algorithms and Convolutional Neural Networks in Hyperspectral Classification of Nitrogen Treatment in Tomato Leaves. *Remote Sens.* **2022**, *14*, 6366. [[CrossRef](#)]
44. Kamei, M. Effect of Image Resolution on Automatic Detection of Whitefly (Hemiptera: Aleyrodidae) Species on Tomato Leaflets Using Deep Learning. *Smart Agric. Technol.* **2023**, *6*, 100372. [[CrossRef](#)]
45. Ruby, E.D.K.; Amirthayogam, G.; Sasi, G.; Chitra, T.; Choubey, A.; Gopalakrishnan, S. Advanced Image Processing Techniques for Automated Detection of Healthy and Infected Leaves in Agricultural Systems. *Mesopotamian J. Comput. Sci.* **2024**, *2024*, 62–70. [[CrossRef](#)] [[PubMed](#)]
46. Pathak, A.; Vohra, B.; Gupta, K. Performance of Feature Extracted on Leaf Images by Discriminant Analysis on Various Classifiers. *Lect. Notes Electr. Eng.* **2021**, *668*, 287–304. [[CrossRef](#)]
47. Amirruddin, A.D.; Muharam, F.M.; Mazlan, N. Assessing Leaf Scale Measurement for Nitrogen Content of Oil Palm: Performance of Discriminant Analysis and Support Vector Machine Classifiers. *Int. J. Remote Sens.* **2017**, *38*, 7260–7280. [[CrossRef](#)]
48. Kurmi, Y.; Gangwar, S.; Agrawal, D.; Kumar, S.; Srivastava, H.S. Leaf Image Analysis-Based Crop Diseases Classification. *Signal Image Video Process.* **2021**, *15*, 589–597. [[CrossRef](#)]

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.