Scientia Agropecuaria

RESEARCH ARTICLE

Rice leaf disease detection using the Stretched Neighborhood Effect Color to Grayscale method and Machine Learning

Elen Yanina Aguirre-Rodríguez[1]* [ID]; Alexander Alberto Rodriguez Gamboa[2] [ID];
Elias Carlos Aguirre Rodríguez[3] [ID]; Juan Pedro Santos-Fernández[1] [ID]; Luiz Fernando Costa Nascimento[3] [ID];
Aneirson Francisco da Silva[3] [ID]; Fernando Augusto Silva Marins[3] [ID]

[1] Facultad de Ingeniería, Universidad Nacional de Trujillo, Av. Juan Pablo II s/n – Ciudad Universitaria, Trujillo, Peru.
[2] Programa de Investigación Formativa e Integridad Científica, Universidad César Vallejo, Trujillo 13001, Peru.
[3] Department of Production, São Paulo State University (UNESP), Guaratinguetá, 12516-410, São Paulo, Brazil.

* Corresponding author: eaguirrer@unitru.edu.pe (E. Y. Aguirre-Rodríguez).

Abstract

The emergence of Machine Learning (ML) technologies and their integration into agriculture has demonstrated a significant impact on disease detection in crops, enabling continuous monitoring and enhancing risk planning and management. This study applied image processing techniques such as thresholding, gamma correction, and the Stretched Neighborhood Effect Color to Grayscale (SNECG) method, alongside ML, to develop a predictive model for identifying five types of rice diseases. The ML techniques used included Logistic Regression, Multilayer Perceptron, Support Vector Machines, Decision Trees, and Random Forests (RF). Hyperparameters were optimized and evaluated through 5-fold cross-validation. In the results, the SNECG method successfully converted images to grayscale, capturing essential features of lesions on rice leaves. The ML models developed with these techniques showed evaluation metrics exceeding 80%, with the RF model (precision = 88.31%) demonstrating superior performance. Additionally, the RF model was integrated into an interface designed for agricultural decision-making. The practical application of the developed model could significantly improve the ability to detect and manage diseases in rice crops.

Keywords: leaf disease; disease classification; disease detection; image processing; machine learning; random forest.

Cite this article:
Aguirre-Rodríguez, E. Y., Rodriguez Gamboa, A. A., Aguirre Rodríguez, E. C., Santos-Fernández, J. P., Nascimento, L. F. C., da Silva, A. F., & Marins, F. A. S. (2025). Rice leaf disease detection using the Stretched Neighborhood Effect Color to Grayscale method and Machine Learning. *Scientia Agropecuaria*, *16*(1), 123-136.

1. Introduction

Globally, rice constitutes a significant portion of the diet for more than half of the population. From 1994 to 2019, Asia was the largest producer of rice, accounting for approximately 90.6% of the total production, followed by the Americas (5.2%), Africa (3.5%), Europe (0.6%), and Oceania (0.1%) (FAO, 2019; Carcea, 2021).

According to the Foreign Agricultural Service, global rice production in 2022/2023 reached 516.73 million tons, with an annual growth of 1% projected for 2023/2024, reaching 522.65 million tons (USDA, 2025). Additionally, among the top ten rice-producing countries in 2024 were China (28%), India (26%), Bangladesh (7%), Indonesia (6%), Vietnam (5%), Thailand (4%), the Philippines (2%),

Myanmar (2%), Pakistan (2%), and Cambodia (1%). Notably, approximately 86% of global rice production came from predominantly Asian countries, while in the Americas, the leading producers were Brazil (1%) and the United States (1%) (USDA, 2025). On the other hand, global projections emphasize the necessity of increasing staple food production by 70% between 2005 and 2050 to ensure nutritional security, given the world's population expansion (FAO, 2009). Projections specific to rice indicate a 26% production increase by 2035, particularly in Africa and Latin America, to meet the growing demand (Seck et al., 2012).

In this way, several factors, including pests and diseases, pose significant threats to rice crops, resulting in substantial yield losses (Nakandakari,

2017; Savary et al., 2019). Diseases and pests account for approximately 30% of these losses (Savary et al., 2019). Moreover, susceptibility to infections throughout the growth stages leads to decreased productivity, increased production costs, and unmet demand, jeopardizing food security (Kawtrakul et al., 2015).

Thus, in the event of an infection, it is important to promptly diagnose the type of rice disease so it can be controlled and treated on time. This ensures efficient and high-quality rice production while minimizing losses and negative impacts on yield.

A correct diagnosis for appropriate treatment necessitates specialists with extensive experience to accurately identify the type of disease (Lu et al., 2017). Consequently, less experienced young farmers may misdiagnose the problem, potentially leading to the application of incorrect pesticides (Sethy et al., 2020).

For this reason, the advancement of technology and its application in agriculture has led to the use of digital technologies to monitor agricultural production. This integration has evolved, and agriculture is currently immersed in the era of Agriculture 4.0, also known as Digital Agriculture. This era is characterized by the incorporation of computer science and robotics, as well as the use of current technologies such as the Internet of Things, cloud computing, big data, and artificial intelligence to significantly enhance agricultural activities (Zhai et al., 2020).

A literature review has revealed a growing interest in using artificial intelligence technologies, like Machine Learning (ML), across different areas of the production chain (Rodríguez et al., 2024a; Rodríguez et al., 2024b), and the agricultural sector has not been an exception (Rodríguez et al., 2022). This interest is due to their efficiency and effectiveness in decision-making, as well as their applicability for disease detection and classification across various crop types (Kartikeyan & Shrivastava, 2021).

Therefore, ML is a powerful technique that, when used correctly, can be highly efficient for developing models that produce reliable results (Rodríguez et al., 2024b). This makes the decision-making process simpler and allows for conclusions to be reached in less time.

Regarding the use of ML for disease identification in rice production, the literature review revealed that, in addition to these techniques, image processing methods can be utilized to enhance certain features. Furthermore, the main ML techniques used include Logistic Regression (LR) (Feng et al., 2020), Support Vector Machines (SVM) (Lu et al.,

2017; Feng et al., 2020; Tian et al., 2021; Sharma et al., 2022), Decision Trees (DT) (Sharma et al., 2022), ensemble methods such as extreme gradient boosting (Azim et al., 2021), Random Forest (RF) (Reddy et al., 2022; Sharma et al., 2022), AdaBoost (Kumar & Kannan, 2022), as well as Neural Networks and their variations (Sethy et al., 2020; Jiang et al., 2021; Elmitwally et al., 2022).

On the other hand, regarding the application of these technologies for disease identification, it was observed that ML techniques were applied to identify around 20 diverse types of diseases (Rodríguez et al., 2022). However, the diseases that appeared most frequently were brown spot, blast, bacterial blight, and leaf smut (Lu et al., 2017; Feng et al., 2020; Sethy et al., 2020; Azim et al., 2021; Jiang et al., 2021; Elmitwally et al., 2022), among others.

In this context, this study aimed to apply image processing techniques, such as segmentation, gamma correction, and the Stretched Neighborhood Effect Color to Grayscale (SNECG) method, as well as ML techniques to develop a predictive model for the early detection of diseases in rice fields. The images used in this study were sourced from secondary data, and various supervised ML techniques were applied and compared.

The article is structured as follows: Section 2 presents the materials and methods used in this study. Section 3 describes the results and discussion, detailing the image processing methods and ML techniques applied. Finally, Section 4 provides the conclusions, followed by the references.

2. Methodology

The flowchart depicted in **Figure 1** illustrates the methodological approach employed in the development of this study, delineated into three principal steps, as explained below.

2.1. Data collection

The dataset for this study comprised images collected from five types of rice diseases obtained from secondary sources (Mendeley, 2022; IPM, 2023; IRRI, 2023). The collected image dataset consisted of 1538 images in JPEG (joint photographic experts' group) format, with an original resolution of 4160 pixels in length and 1952 pixels in width.

The dataset included images of five major rice diseases: sheath blight (28.48%), rice blast (25.75%), leaf scald (18.60%), rice tungro (15.47%), and brown spot (11.70%). These images were systematically organized into separate folders based on the disease type. The labels for each class were generated using the folder structure, ensuring that all images within the same folder were assigned the same class label.

Additionally, to streamline computational efficiency during model development, the images were resized to a resolution of 128 × 128 pixels.

2.2. Image preprocessing

To capture rice disease-specific characteristics, various image processing methods such as thresholding, gamma correction, and conversion to grayscale were applied, as illustrated in **Figure 1**.

Thresholding

A highly popular image segmentation technique is thresholding, which involves separating the foreground from the background of the image by creating binary images (**Nixon & Aguado, 2019**; **Rajinikanth et al., 2020**).

Image thresholding involves selecting a threshold T such that any pixel (x, y) satisfying $f(x, y) > T$ can be termed an object pixel (**Gonzales & Wintz, 2017**). Among advanced thresholding techniques is optimal thresholding with Otsu's method, which aims to find an optimal value, T, for separating the object from the background by employing a generalized grayscale histogram, where the number of pixels at each gray level is divided by the total number of pixels in the image (**Nixon & Aguado, 2019**;

Gonzales & Wintz, 2017). Gaussian smoothing was employed to remove noise from the images (**Gonzales & Wintz, 2017**).

In the thresholding process, the probability distribution for the L gray levels of an image with dimensions $M \times N$ is provided by the mathematical formula in **Equation 1**:

$$p_\ell = \frac{n_\ell}{MN} \tag{1}$$

where the gray level ℓ ranges from 0 to 255, n_ℓ represents the number of pixels in the image with gray level ℓ, and $MN = n_0 + n_1 + ... + n_{L-1}$ denotes the total number of pixels. The threshold $T = k$ (gray level) lies within the intensity range $0 < k < L - 1$, dividing the image into classes $c_1 = [0, k]$ and $c_2 = [k + 1, L - 1]$. Thus, given the threshold k, the probability of a gray level being classified into classes c_1 or c_2 is provided in **Equations 2 and 3**.

$$P_1(k) = \sum_{\ell=0}^{k} p_\ell \tag{2}$$

$$P_2(k) = \sum_{\ell=k+1}^{L-1} p_\ell = 1 - P_1(k) \tag{3}$$

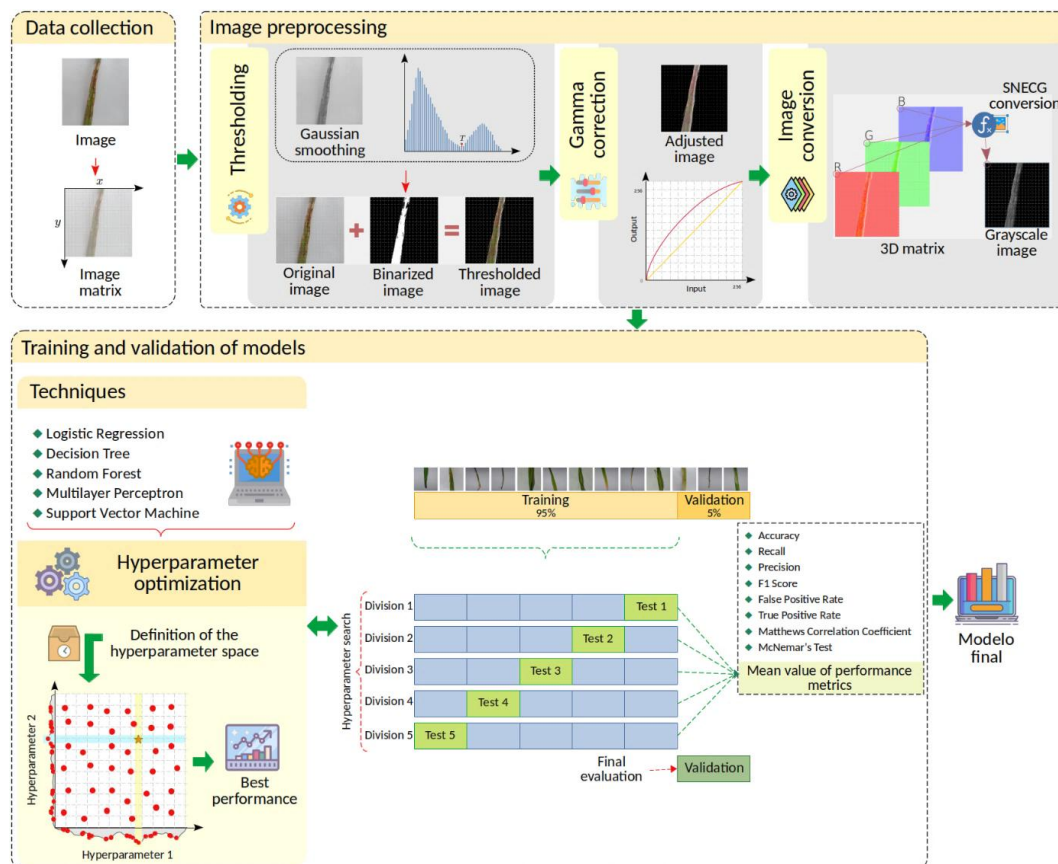where $P_1(k)$ is the probability of class c_1 and $P_2(k)$ is the probability of class c_2.

Figure 1. Methodological approach.

The average of gray levels for classes c_1 and c_2 is calculated as shown in **Equations 4 and 5**. Furthermore, the global mean of gray levels is calculated by **Equation 6**:

$$m_1(k) = \frac{1}{P_1(k)} \sum_{\ell=0}^{k} \ell p_\ell \qquad (4)$$

$$m_2(k) = \frac{1}{P_2(k)} \sum_{\ell=k+1}^{L-1} \ell p_\ell \qquad (5)$$

$$m_G = \sum_{\ell=0}^{L-1} \ell p_\ell \qquad 6)$$

where it must hold that $P_1(k)\, m_1(k) + P_2(k)\, m_2(k) = m_G$ and $P_1(k) + P_2(k) = 1$.

Thus, the global variance of the gray level across all image pixels is computed by $\sigma_G^2(k)$ (**Equation 7**), while the between-class variance is calculated by $\sigma_B^2(k)$ (**Equation 8**). These variance measures are then used to calculate $\eta(k)$ (**Equation 9**), which facilitates the evaluation of the effectiveness of the chosen threshold in separating the image components.

$$\sigma_G^2 = \sum_{\ell=0}^{L-1} (\ell - m_G)^2 \, p_\ell \qquad (7)$$

$$\sigma_B^2(k) = P_1(k)\left[m_1(k) - m_G\right]^2 + P_2(k)\left[m_2(k) - m_G\right]^2 \qquad (8)$$

$$\eta(k) = \frac{\sigma_B^2(k)}{\sigma_G^2} \qquad (9)$$

where the values of $\eta(k)$ are within the range $0 \le \eta(k) \le 1$. In this way, the optimal threshold $T = k^*$ must satisfy **Equation 10**:

$$\sigma_B^2(k^*) = \max_{0 \le k \le L-1} \sigma_B^2(k) \qquad (10)$$

where $\sigma_B^2(k^*)$ maximizes the between-class variance (**Nixon & Aguado, 2019; Gonzales & Wintz, 2017**).

Gamma correction

Gamma correction (γ) adjusts the variations in luminance levels between individual pixels in an image, thereby enhancing visual appearance and highlighting specific features (**Gonzales & Wintz, 2017**). Thus, gamma correction in an image is calculated using the **Equation 11**:

$$s = c r^{\gamma} \qquad (11)$$

where s allows for a non-linear transformation of the color levels in the RGB (red, green, blue) channels, and the variables c and γ are positive constants (**Nixon & Aguado, 2019**).

Conversion to grayscale

Considering that an image is commonly represented in the RGB bands (**Rafael et al., 2020; Gonzalez & Woods, 2018**), the adjusted color images were simplified for processing through conversion to grayscale.

A simple and popular method is linear projection $f(x)_{GS} = \alpha_R \times R + \alpha_G \times G + \alpha_B \times B$, where the values of α_R, α_G, α_B are non-negative and satisfy the constraint $\alpha_R + \alpha_G + \alpha_B = 1$ (**Kanan & Cottrell, 2012; Gonzalez & Woods, 2018**). The simplest method for converting images to grayscale is the average method, where the coefficient values correspond to one-third of the original values of the RGB channels (**Ma et al., 2015**).

In this study, the SNECG method was utilized, which is based on the pixel neighborhood approach and is distinguished by its adaptive characteristics that enhance the brightness, contrast, and details of the images (**Lim % Isa, 2011**). The SNECG method is obtained through the formula presented in **Equation 12**:

$$f(X)_{GS} = W_R \times R^* + W_G \times G^* + W_B \times B^* \qquad (12)$$

where the coefficients W_R, W_G, and W_B are estimated by averaging the coefficients T_R, T_G, and T_B, which are obtained by summing the intensity levels of each RGB channel, as presented in **Equations 13 and 14**:

$$W_R = \frac{T_R}{T_R + T_G + T_B} , \; W_G = \frac{T_G}{T_R + T_G + T_B} \text{ and}$$

$$W_B = \frac{T_B}{T_R + T_G + T_B} \qquad (13)$$

$$T_R = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} R , \; T_G = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} G \text{ and}$$

$$T_B = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} B \qquad (14)$$

The modified values of the coefficients R^*, G^*, and B^* determine the extended values (**Lim and Isa, 2011**), which are calculated using the minimum and maximum values of the RGB channels, as presented in **Equations 15 - 17**.

$$R_{max} = \max(R) , \; G_{max} = \max(G) \text{ and}$$

$$B_{max} = \max(B) \qquad (15)$$

$$R_{min} = \min(R) , \; G_{min} = \min(G) \text{ and}$$

$$B_{min} = \min(B) \qquad (16)$$

$$R = 255 \times \frac{R - R_{min}}{R_{max} - R_{min}} , \; G = 255 \times \frac{G - G_{min}}{G_{max} - G_{min}} \text{ and}$$

$$B = 255 \times \frac{B - B_{min}}{B_{max} - B_{min}} \qquad (17)$$

2.3. Training and validation of models

In addition to image processing, this study was based on applying ML techniques due to their ability to identify complex patterns in high-dimensional datasets (**Marsland, 2015**; **Alpaydin, 2021**). ML is a subfield of artificial intelligence that uses computational and statistical techniques to create mathematical models that can recognize specific features in images (**Jordan & Mitchell, 2015**). Supervised ML classification techniques such as LR, DT, RF, Multilayer Perceptron (MLP), and SVM were used in this work. On the other hand, the optimal combination of hyperparameters for each model was obtained using the Random Search method.

The image dataset was divided into 95% for model training using the 5-fold cross-validation method, with its optimal hyperparameter configuration obtained via random search optimization. On the other hand, due to the class imbalance in the collected dataset, the oversampling technique was employed in each subdivision. This was done to mitigate the inherent bias in the unbalanced data (**Rahman et al., 2015**). The remaining 5% of images were utilized for model validation, as illustrated in **Figure 1**.

Logistic Regression

A classic statistical technique that fits an "S"-shaped curve, akin to linear regression, is utilized. This fitted curve is employed to compute the probability that the output *y* predicted by the model belongs to one of the *k* classes (**Alpaydin, 2021**). The logistic function for models aimed at calculating the probability of two classes is presented in **Equation 18**:

$$p(x_i) = \frac{e^{\beta_0 + \sum \beta_i x_i}}{1 + e^{\beta_0 + \sum \beta_i x_i}} = \frac{1}{1 + e^{-\beta_0 - \sum \beta_i x_i}} \tag{18}$$

Where β_0 and β_1 are the parameters or coefficients of the model, which are calculated using the maximum likelihood method, and x_i are the input variables (**Alpaydin, 2021**).

Decision Trees

The DT technique enables the construction of predictive models through a sequence of binary splits of the training dataset, structured with nodes and leaves. Specifically, in the context of classification trees, their construction involves an iterative process that begins with selecting the most significant feature from the dataset to form the root node (**Alpaydin, 2021**). Based on the root node, internal binary subdivisions occur, leading to the creation of internal nodes (*t*) until reaching the

model's leaves, where predictions are made (**Marsland, 2015**). These data subdivisions must satisfy $\left(X_t^a \cup X_t^b = X_t \wedge X_t^a \cup X_t^b = \varnothing \right)$, continuing until a reduction in error is achieved at each leaf node (**Alpaydin, 2021**).

The leaves are also known as predictors *d(x)*, and their prediction performance can be evaluated using the Gini index, a measure of homogeneity aimed at reducing data impurities from the root node to the leaf nodes (**Alpaydin, 2021**; **Marsland, 2015**). The Gini index represents a measure of total variance among *K (k = 1, 2, ..., K)* classes (**Gareth et al., 2013**; **Alpaydin, 2021**; **Marsland, 2015**) and is calculated as shown in **Equation 19**:

$$G = \sum_{k=1}^{K} p(X_{tk})\left[1 - p(X_{tk})\right] \tag{19}$$

where $p(X_{tk})$ denotes the probability that a proportion of training data in node *t* belongs to class *k* (**Gareth et al., 2013**). A small value of *G* indicates that predictor *d(x)_t* consists of data from one class (**Gareth et al., 2013**; **Alpaydin, 2021**). Another method to evaluate the constructed tree's performance is by assessing the quality of data split using entropy calculation (**Alpaydin, 2021**), shown in **Equation 20**:

$$D = \sum_{k=1}^{K} p(X_{tk}) \log p(X_{tk}) \tag{20}$$

Like the Gini index, the entropy value *D* will be small if $p(X_{tk})$ is close to zero or one, indicating a high purity split (**Gareth et al., 2013**). Furthermore, the classification error rate, depicted in **Equation 21**, can be utilized to evaluate the accuracy of the final model, post-pruning using either entropy or the Gini index.

$$E = 1 + \max_k \left[p(X_{tk}) \right] \tag{21}$$

Multilayer Perceptron

The MLP models the relationship between input signals (input variables) and the output signal (target variable). Its basic structure consists of three main layers: the input layer, one or more hidden layers, and the output layer (**Alpaydin, 2021**).

The neurons in the input layer of the model receive the values to be processed, which flow through several hidden layers, considered the main computational drive. The output layer then performs the prediction or classification based on the information from the input layer (**Khan et al., 2022**). The interconnected layers of the neural network use the backpropagation technique to enhance prediction accuracy. The gradient is

calculated using the error function with the neuron weights in this process. Additionally, each neuron in the model employs a nonlinear activation function (Alpaydin, 2021).

The output of each neuron in the MLP specifically depends on the preceding neurons and the network weights, which can be represented by Equation 22:

$$f(x) = \varnothing \left(w_{ij} x_i + \theta_i \right) \qquad (22)$$

where $\varnothing$ is the activation function, making its outputs known as activations, x_i are the input variables, w_{ij} are the synaptic weights, and $w_{i0} = \theta_i$ is the bias (Alpaydin, 2021; Khan et al., 2022).

Random Forest

The RF technique is an extension of DT and is also known as one of the ensemble methods (Sheykhmousa et al., 2020; Gareth et al., 2013). This technique is characterized by its internal structure, which consists of a collection of classification models using DT, where each model is trained on a random subset of data (Sheykhmousa et al., 2020; Marsland, 2015). Each internal tree generated results in a class, this outcome is the class with the highest frequency (Belgiu & Drăguţ, 2016; Gareth et al., 2013).

In the RF, the L classification trees ($f^1(x)$, $f^2(x)$, ..., $f^L(x)$) trained are used to predict the final class $f(x)$ (Belgiu & Drăguţ, 2016; Sheykhmousa et al., 2020). Specifically, $f(x)$ is obtained through majority voting among the L classifiers (Belgiu & Drăguţ, 2016; Alpaydin, 2021).

Support Vector Machines

The SVM is noted for its efficiency with high-dimensional data, even when the number of dimensions exceeds the number of instances in the dataset (Marsland, 2015). This technique maps the input data into a higher-dimensional nonlinear feature space to find a hyperplane that maximizes the margin between classes by minimizing the distance between them (Islam et al., 2017; Sheykhmousa et al., 2020).

For constructing binary models, the classifier is obtained by solving a regularization problem that maximizes the margin between classes through the minimization of the associated objective function, as depicted in Equation 23 and 24:

$$\min_{f \in H} C \sum_{i=1}^{l} \xi_i + \frac{1}{2} |f|_k^2 \qquad (23)$$

subject to:

$$y_i f(x_i) \geq 1 - \xi_i \quad , \quad \xi_i \geq 0 \quad \wedge \quad i = 1, 2, \ldots, l \quad (24)$$

where x_i and y_i denote the training data and labels, $f(x)$ represents the classifier, C is a constant representing the regularization factor, and ξ_i are slack variables (Islam et al., 2017).

Evaluation and validation metrics

Evaluating and validating predictive models is crucial, as it helps measure the accuracy and overall performance of the models by assessing their error rates and efficiency. The cross-validation technique enables the evaluation of a model's robustness by preventing overfitting and helping to estimate its performance in practical applications (Marsland, 2015; Alpaydin, 2021). In its application, the training dataset is randomly divided into k equally sized subsets (5-fold).

In this way, although there are various evaluation metrics available, no single standard metric has been established (Rodriguez et al., 2022). The formulas for the evaluation metrics adopted, based on the confusion matrix, are presented below in Equations 25 to 30.

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \qquad (25)$$

$$recall = \frac{TP}{TP + FN} \qquad (26)$$

$$precision = \frac{TP}{TP + FP} \qquad (27)$$

$$F1 = 2 \times \frac{precision \times recall}{precision + recall} \qquad (28)$$

$$FPR = \frac{FP}{FP + TN} \qquad (29)$$

$$TPR = \frac{TP}{TP + FN} \qquad (30)$$

The Matthews correlation coefficient (MCC) was used to measure the quality of the classifications generated by the models (Zhu, 2020). The MCC is calculated as depicted in Equation 31.

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \qquad (31)$$

The MCC value varies between {+1, -1}, with values close to +1 indicating perfect classification and those near -1 indicating perfect misclassification. Conversely, a value of 0 indicates random prediction, signifying the model's inability to predict (Chicco, 2020).

Additionally, the McNemar test was utilized to validate the selection of the best final model by comparing and analyzing the frequency of errors or successes of each one. Hence, the comparison of

models is grounded on the null hypothesis that both models possess the same error rate (H_0: $error_{M1} = error_{M2}$). The McNemar statistic (X^2) is calculated according to **Equation 32**.

$$X^2 = \frac{\left(\left|error_{M1} - error_{M2}\right| - 1\right)^2}{error_{M1} - error_{M2}} \tag{32}$$

The null hypothesis is rejected at a significance level of α if the value of X^2 is greater than $X_{\alpha,1}$ (**Alpaydin, 2021**).

3. Results and discussion

The optimal threshold values of the analyzed images per class showed a close variation within their value ranges. **Figure 2 (A-B)** illustrates the distribution of the optimal threshold values obtained for each image per class.

For the brown spot class, the threshold value k^* varied between 103 and 149. For the leaf scald class, the range was between 105 and 147, while for rice blast, it ranged between 104 and 152. In the case of rice tungro, the threshold spanned from 105 to 149, and for the sheath blight class, it varied between 102 and 149 (**Figure 2**). Additionally, a gamma value of $\gamma = 0.65$ was utilized in the gamma correction method. **Figure 2 (C)** displays the curve of color levels of individual pixels corrected with the gamma value considered, resulting in brightness adjustment of the dataset images. Improving the visual appearance of the images through thresholding

and gamma correction methods emphasized specific characteristics of rice leaves, such as the lesions caused by the evaluated diseases.

The application of the SNECG method was chosen because, unlike the average method, SNECG demonstrated superior performance in the conversion of grayscale images, effectively capturing the inherent characteristics related to the color of each lesion associated with several types of diseases (**Figure 3A**).

The SNECG method enabled a more detailed capture of the brightness and characteristics of each color image during the conversion to grayscale, as shown in **Figure 3 (B)**.

Next, the **Table 1** presents the labels and the number of images per class for the training and validation processes.

Table 1
Dataset division for training and validation

Class	Label	Training	Validation
Leaf scald	0	271	15
Rice blast	1	375	21
Brown spot	2	173	7
Sheath blight	3	413	25
Rice tungro	4	229	9

Among the training outcomes, each model exhibited varying performance based on the results of 5-fold cross-validation. **Figure 4** presents the performance metrics, including accuracy, recall, precision, and F1-score.
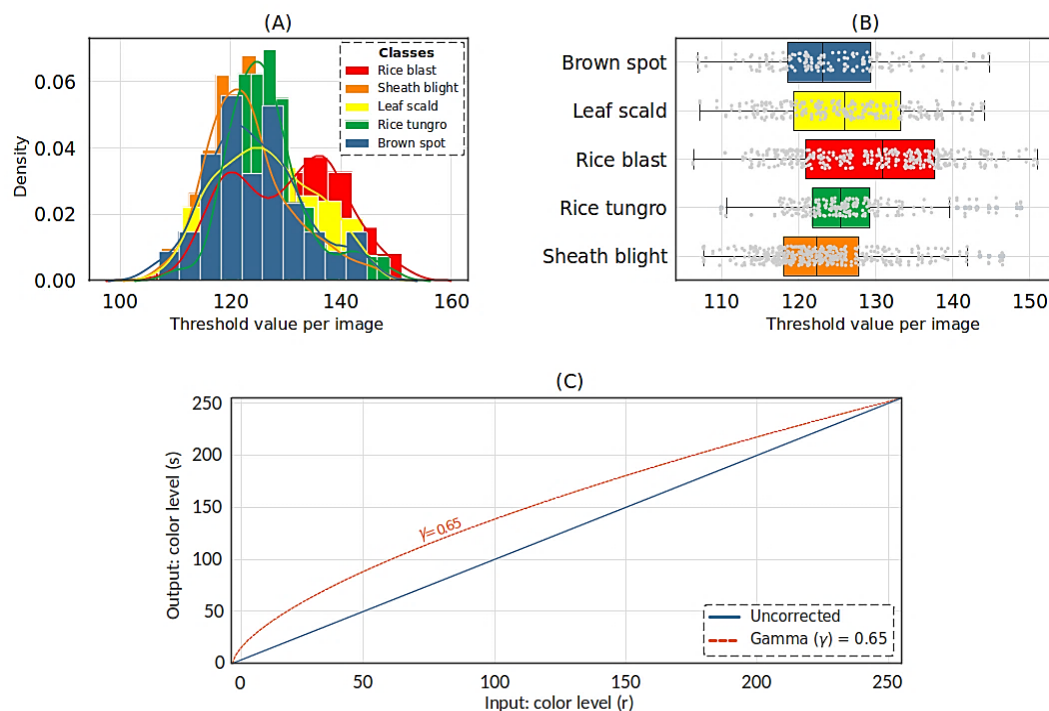

Figure 2. Distribution of optimal threshold values and color curve generated with gamma correction.
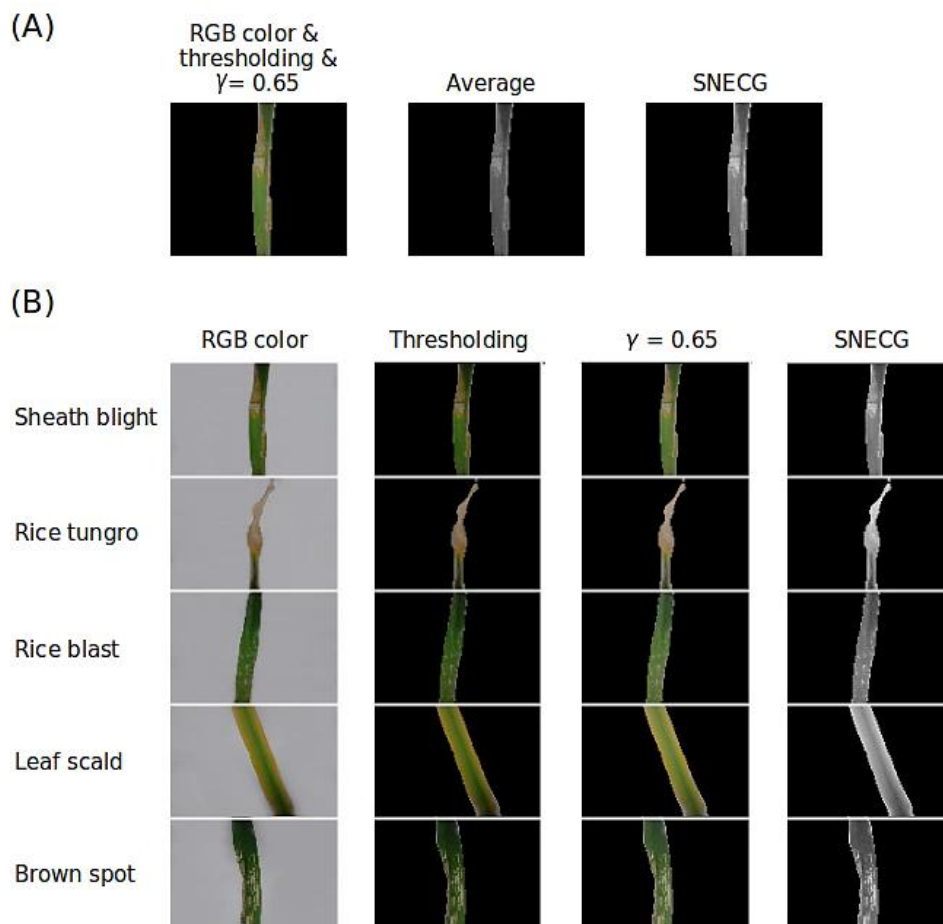
Figure 3. (A) Color-processed images and grayscale images produced using the average method and the SNECG method, along with (B) original and processed images using thresholding, gamma correction, and the SNECG method for each type of disease.

The performance analysis of the models trained using 5-fold cross-validation showed that all tested techniques achieved metrics above 78%, which can be considered acceptable. **Figure 5** illustrates the mean values for the evaluation metrics FPR, TPR, and MCC, along with their corresponding confidence intervals (95% CI) at a 5% significance level.

Upon inspection of the model training results (**Figures 4 and 5**), it was observed that models generated using SVM and RF techniques achieved the highest number of correctly classified images. To facilitate the comparison of error metric results for each evaluated technique, **Table 2** presents the mean values of the evaluation metrics obtained during model training.

As observed in **Table 2**, for the metrics of accuracy, precision, recall, F1-score, and TPR, all tested techniques exceeded 82.29%, while the FPR metric presented values below 4.30%. The results indicated that the models that used MLP, DT, and LR exhibited the lowest values in the evaluated performance metrics. Concerning the model trained with the LR technique, it was observed that

it presented values similar to those obtained by **Feng et al. (2020)**.

Additionally, the models trained with RF and SVM techniques achieved the best error metric values, with very close values to each other compared to models trained with other techniques (**Table 2**).

The MCC was used to measure the quality of the predictions of each trained model. It was found that all tested models obtained a mean MCC value above 78.05%. The MCC results demonstrated that models using RF (83.84%, with CI95%: 88.41 – 79.28) and SVM (83.58%, with CI95%: 88.14 – 79.02) showed the best prediction performance compared to the models using MLP (80.74%, with CI95%: 85.67 – 75.82), DT (78.76%, with CI95%: 82.95 – 74.57), and LR (78.05%, with CI95%: 84.36 – 71.74). Based on the results obtained during the cross-validation training stage, the models using RF and SVM were retrained with the remaining 95% of the dataset to validate and select the final model. **Table 3** presents the optimal hyperparameters for each model trained with each final ML technique.
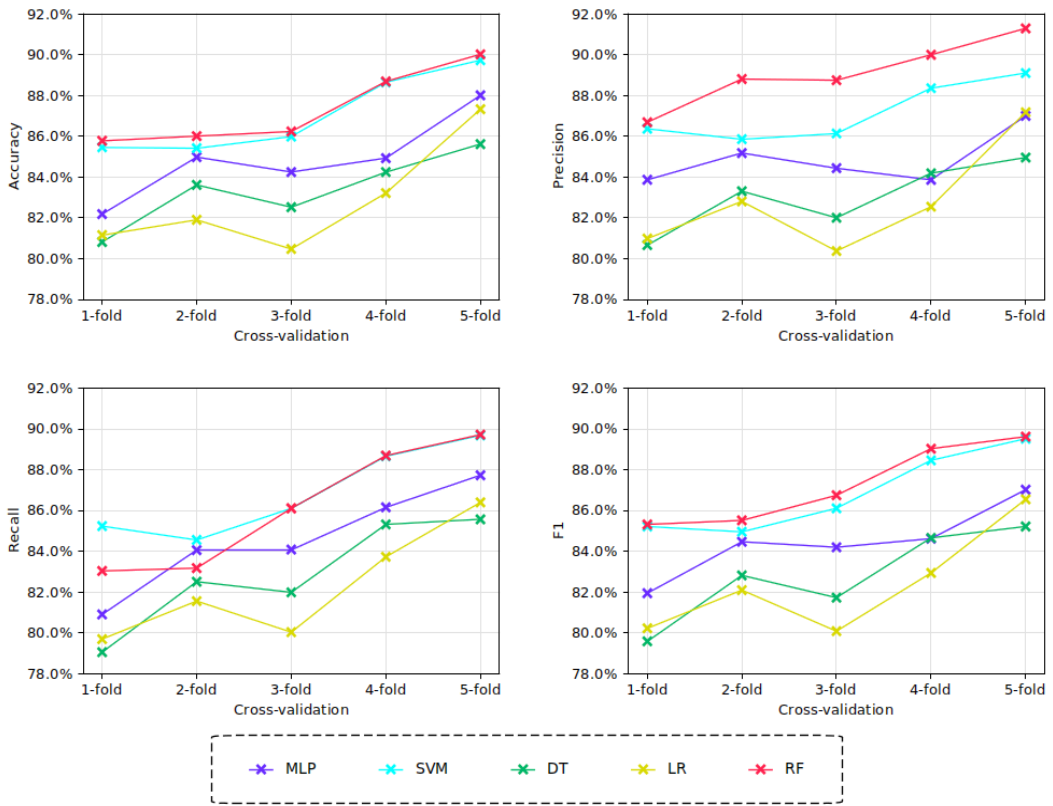
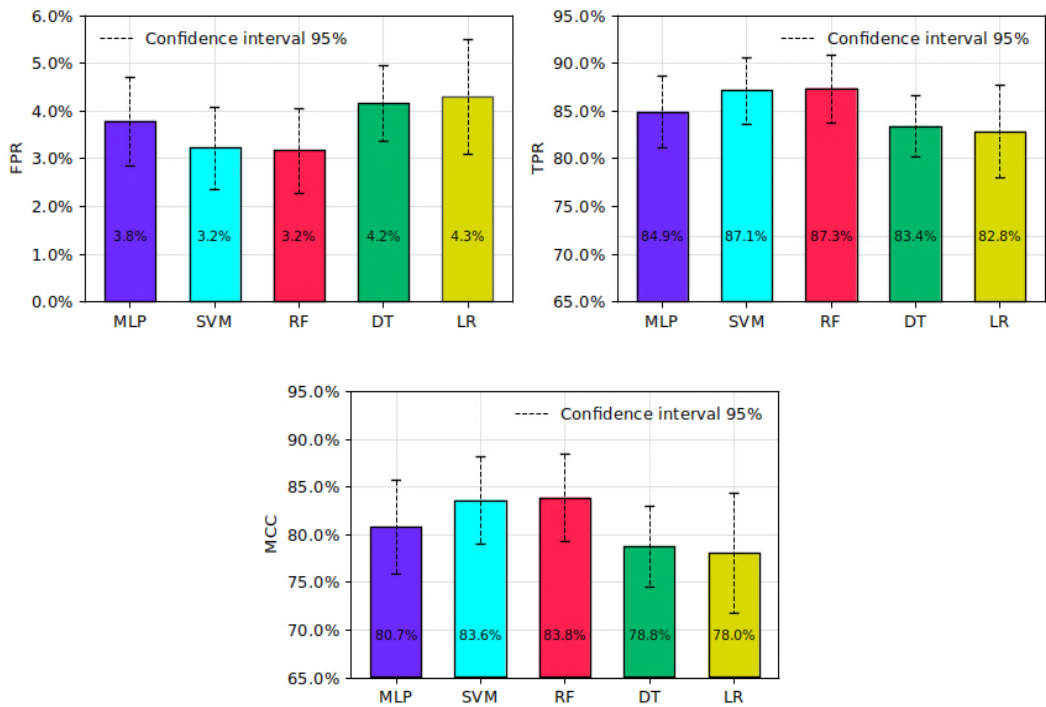**Figure 4.** Performance metrics for each fold of the cross-validation.

Figure 5. Mean values of the evaluation metrics FPR, TPR, and MCC.

Table 2
Mean values of evaluation metrics during models training

| Metrics	Models				
	MLP (CI$_{95\%}$)	SVM (CI$_{95\%}$)	RF (CI$_{95\%}$)	DT (CI$_{95\%}$)	LR (CI$_{95\%}$)
Accuracy	84.87% (88.61 - 81.14)	87.05% (90.63 - 83.47)	87.35% (90.74 - 83.95)	83.37% (86.61 - 80.13)	82.82% (87.68 - 77.96)
Precision	84.87% (87.21 - 82.54)	87.17% (89.78 - 84.55)	89.11% (92.16 - 86.06)	83.03% (86.10 - 79.96)	82.78% (87.55 - 78.02)
Recall	84.59% (89.19 - 79.99)	86.85% (90.82 - 82.88)	86.15% (91.64 - 80.66)	82.89% (87.68 - 78.10)	82.29% (87.29 - 77.28)
F1	84.45% (87.67 - 81.23)	86.85% (90.49 - 83.22)	87.25% (90.80 - 83.70)	82.80% (86.88 - 78.73)	82.38% (87.08 - 77.68)
FPR	3.78% (4.71 - 2.85)	3.22% (4.09 - 2.35)	3.17% (4.06 - 2.28)	4.16% (4.96 - 3.36)	4.30% (5.51 - 3.08)
TPR	84.87% (88.61 - 81.14)	87.13% (90.63 - 83.64)	87.34% (90.90 - 83.78)	83.37% (86.61 - 80.13)	82.82% (87.68 - 77.96)
MCC	80.74% (85.67 - 75.82)	83.58% (88.14 - 79.02)	83.84% (88.41 - 79.28)	78.76% (82.95 - 74.57)	78.05% (84.36 - 71.74)
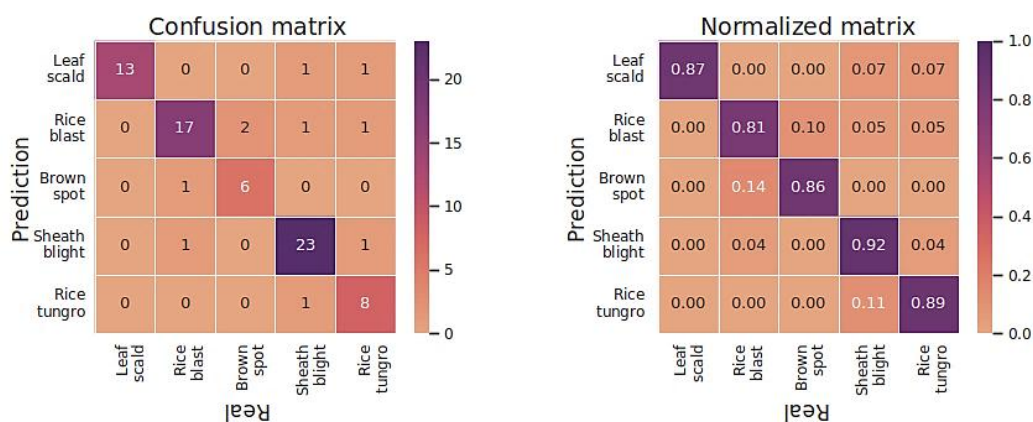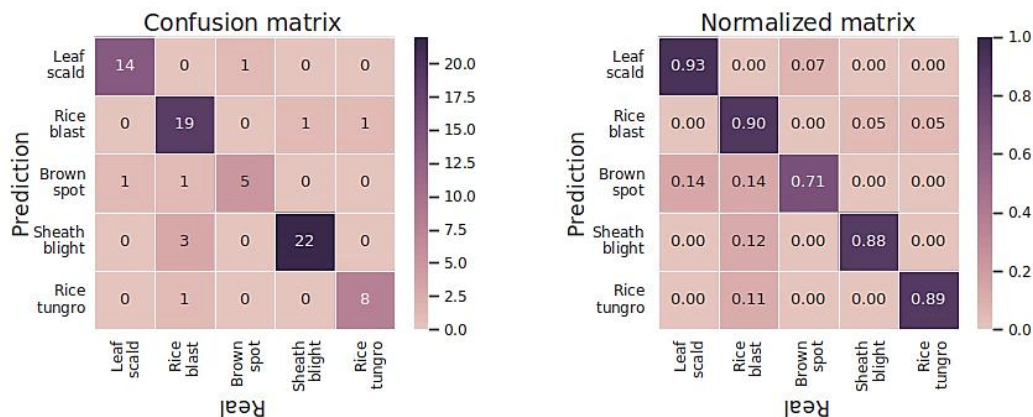
Figure 6. Confusion matrix for the validation of the trained model with SVM and RF.

Considering that model validation was conducted using 5% (77) of the images, Figure 6 (A) presents the confusion matrix of classifications made with the SVM model, demonstrating that the model correctly classified 87.01% of the images. Figure 6 (B) presents the confusion matrix for the RF model validation, demonstrating that the model developed using this technique correctly classified 88.31% of the images. This technique exhibited behavior closely resembling the results obtained with SVM.

Table 3
Hyperparameters of the best ML models

Model	Hyperparameter	Description	Parameter
RF	n_estimators	Number of classification trees.	128
	max_features	Maximum number of features used to split a node.	sqrt
	criterion	Function to measure the quality of the split.	gini
	min_samples_split	Minimum number of features in a node before splitting.	2
	min_samples_leaf	Minimum number of features in a leaf node.	1
	bootstrap	Sampling method for constructing the trees.	False
SVM	C	Regularization parameter.	125
	decision_function	Decision function.	ovr
	max_iter	Maximum iteration limit.	-1
	kernel	Kernel function.	rbf
	gamma	Kernel coefficient.	scale

The McNemar test results (Table 4) indicated no significant differences between the errors of the models. These findings suggested the need for a more detailed evaluation to select the final model.

Table 4
Results of the McNemar test in the validation process

Models	X^2	$X_{0.05,1}$	Description
RF			
SVM	0.001	3.84	Accept H_0

The metrics results of both models were remarkably similar; however, the RF model exhibited higher precision and accuracy exceeding 88%, which outperformed the SVM technique (Table 5). Additionally, the Matthews correlation coefficient indicated that the RF model achieved better performance, slightly surpasssing the SVM model with a coefficient close to 85%.

Table 5
Evaluation metrics results with validation images

Models	SVM	RF
Accuracy	87.01%	88.31%
Precision	86.84%	86.43%
Recall	85.13%	88.07%
F1	85.61%	87.05%
TPR	87.01%	88.31%
FPR	3.25%	2.92%
MCC	83.13%	84.74%

Contrasting these results with some observed in the literature, it is noteworthy that the studies by Tian et al. (2021), Azim et al. (2021), and Feng et al. (2020) developed models using ML techniques that achieved lower accuracy than the SVM and RF techniques used in this study.

Moreover, it is important to note that one of the main differences between those studies and this research is that their models were developed to classify only one to three types of rice diseases, while this work evaluated five types of diseases. Table 6 presents the types of diseases considered by various authors for the development of classification models. ML techniques have been widely applied to the identification of various rice diseases. Among these, the most frequently studied include brown spot, blast, bacterial blight, and leaf smut (Lu et al., 2017; Feng et al., 2020; Sethy et al., 2020; Azim et al., 2021; Jiang et al., 2021; Elmitwally et al., 2022). However, when comparing the findings of these studies with the results of the present investigation, no prior research was identified that comprehensively evaluated the diseases sheath blight, tungro, blast, leaf scald, and brown spot in an integrated manner.

Table 6
Types of diseases analyzed in various published research

Research	Rice diseases														
	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)	(k)	(l)	(m)	(n)	(o)
Lu et al. (2017)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓					
Sethy et al. (2020)	✓		✓				✓				✓				
Jiang et al. (2021)			✓				✓					✓			
Tian et al. (2021)	✓														
Azim et al. (2021)			✓				✓					✓			
Upadhyay y Kumar (2021)			✓									✓			
Feng et al. (2020)	✓			✓			✓								
Sharma et al. (2022)	✓		✓				✓				✓				
Latif et al. (2022)	✓		✓				✓						✓	✓	
Kumar y Kannan (2022)			✓				✓					✓			
Rallapalli y Saleem (2021)	✓		✓				✓					✓			✓
Elmitwally et al. (2022)			✓				✓					✓			
Akyol (2023)			✓				✓					✓			
Reddy et al. (2022)			✓												✓

NOTE. (a): Blast, (b): False smut, (c): Brown spot, (d): Bakanae, (e): Sheath blight, (f): Sheath rot, (g): Bacterial blight, (h): Bacterial sheath brown rot, (i): Seeding blight, (j): Bacterial wilt, (k): Tungro, (l): Leaf smut, (m): Leaf scald, (n): Narrow brown spot, (o): Hispa

Finally, **Figure 7** illustrates an interface developed by integrating the best model with RF to diagnose the type of disease present in the rice leaf. The developed interface allows for observing the classification result, in addition to the probability that the leaf presents the identified disease. In **Figure 7 (A)**, the identification of brown spot disease with a 97.66% probability of accuracy, and in **Figure 7 (B)**, the identification of sheath blight disease with a 75.78% probability of accuracy.

4. Conclusions

Rice is one of the most widely consumed foods globally, and its production faces significant challenges due to the risk of infections from pests and diseases that can negatively impact efficiency and quality. This study applied various image processing methods, such as thresholding, gamma correction, and grayscale conversion using the SNECG method, to characterize disease lesions on rice leaves. Additionally, ML techniques were explored to develop a predictive model for the early detection of crop issues.

Five ML techniques were applied, including RF, LR, MLP, DT, and SVM, with cross-validation and random search for hyperparameter optimization. The results showed that all models correctly classified more than 80% of the images; however, the models with the best performance during training were the SVM and RF models. Validation of the top models indicated that the classifications made with the RF model were more accurate, making it the final selected model, with an accuracy exceeding 88%.

These results highlight that the integration of advanced image processing methods with ML techniques enables the development of highly efficient classification models and opens new possibilities for the precise detection of specific issues in rice crops, leading to a significant improvement in their production.
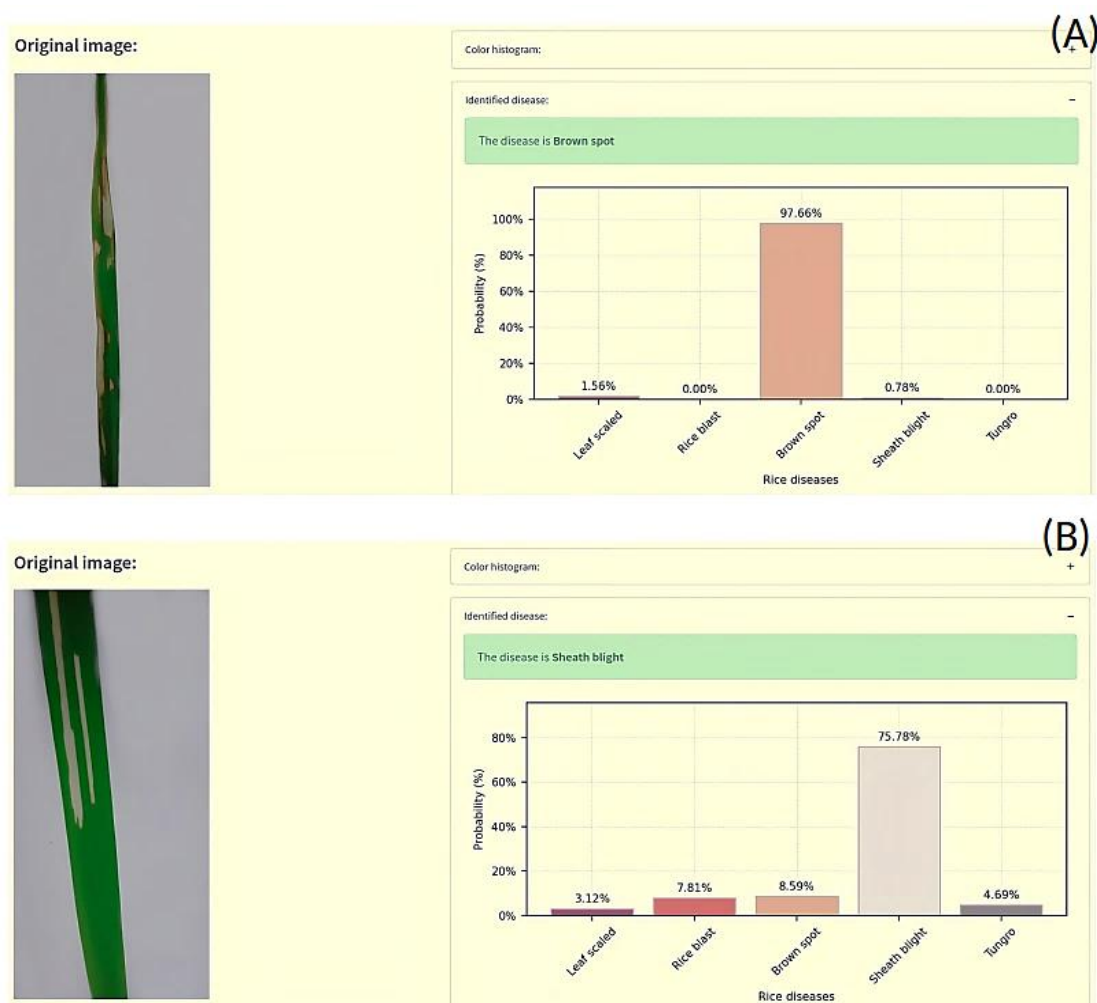

Figure 7. Interface for the identification of diseases in rice leaves.

Moreover, by facilitating more effective management of agricultural diseases, these emerging tools play a crucial role in strengthening food security through more informed and proactive agricultural practices. This approach with ML models can not only optimize the management of agricultural resources but also promote more responsible and effective practices in food production, thereby fostering a more robust and secure agricultural system for future generations.

Finally, for future research, the application and comparison of other ML techniques are recommended, as well as the integration of other types of diseases affecting rice crops, to develop a more robust tool.

Author contributions

Aguirre-Rodríguez E. Y.: Conceptualization; Data curation; Formal Analysis; Investigation; Methodology; Software; Validation; Visualization; Writing – original draft; Writing - review & editing. **Gamboa A. A. R.:** Formal Analysis; Methodology; Validation; Visualization; Writing - review & editing. **Rodríguez E. C. A.:** Formal Analysis; Methodology; Validation; Visualization; Writing - review & editing. **Santos-Fernández J. P.:** Conceptualization; Data curation; Formal Analysis; Investigation; Methodology; Software; Supervision; Validation; Visualization; Writing – original draft; Writing - review & editing. **Nascimento L. F. C.:** Formal Analysis; Methodology; Visualization; Writing - review & editing. **Silva A. F.:** Formal Analysis; Methodology; Visualization; Writing - review & editing. **Marins F. A. S.:** Conceptualization; Data curation; Formal Analysis; Investigation; Methodology; Software; Supervision; Validation; Visualization; Writing – original draft; Writing - review & editing.

Conflict of interest statement

The authors declare that they have no conflict of interest.

ORCID

E. Y. Aguirre-Rodríguez [iD] https://orcid.org/0000-0002-3829-4118
A. A. Rodriguez Gamboa [iD] https://orcid.org/0000-0002-0102-4253
E. C. Aguirre Rodríguez [iD] https://orcid.org/0000-0003-1120-1708
J. P. Santos-Fernández [iD] https://orcid.org/0000-0002-8882-9256
L. F. Costa Nascimento [iD] https://orcid.org/0000-0001-9793-750X
A. F. da Silva [iD] https://orcid.org/0000-0002-2215-0734
F. A. S. Marins [iD] https://orcid.org/0000-0001-6510-9187

References

Alpaydin, E. (2021). *Introduction to machine learning*. MIT press. https://doi.org/10.7551/mitpress/13811.001.0001

Akyol, K. (2023). Handling hypercolumn deep features in machine learning for rice leaf disease classification. *Multimedia Tools and Applications*, *82*(13), 19503-19520. https://doi.org/10.1007/s11042-022-14318-5

Azim, M. A., Islam, M. K., Rahman, M. M., & Jahan, F. (2021). An effective feature extraction method for rice leaf disease classification. *Telkomnika (Telecommunication Computing Electronics and Control)*, *19*(2), 463-470. http://doi.org/10.12928/telkomnika.v19i2.16488

Belgiu, M., & Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS journal of photogrammetry and remote sensing*, *114*, 24–31. https://doi.org/10.1016/j.isprsjprs.2016.01.011

Carcea, M. (2021). Value of wholegrain rice in a healthy human nutrition. *Agriculture*, *11*(8), 720. https://doi.org/10.3390/agriculture11080720

Chicco, D. (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics*, *21*, 6. https://doi.org/10.1186/s12864-019-6413-7

Elmitwally, N. S., Tariq, M., Khan, M. A., Ahmad, M., Abbas, S., & Alotaibi, F. M. (2022). Rice Leaves Disease Diagnose Empowered with Transfer Learning. *Computer Systems Science & Engineering*, *42*(3). https://doi.org/10.32604/csse.2022.022017

Food and Agriculture Organization of the United Nations, FAO (2019). *Crops and livestock products*. FAOSTAT. https://www.fao.org/faostat/en/#data/QCL

Food and Agriculture Organization of the United Nations, FAO (2009). *La agricultura mundial en la perspectiva del año 2050*. https://www.fao.org/fileadmin/templates/wsfs/docs/Issues_papers/Issues_papers_SP/La_agricultura_mundial.pdf

Feng, L., Wu, B., Zhu, S., Wang, J., Su, Z., Liu, F., He, Y., & Zhang, C. (2020). Investigation on data fusion of multisource spectral data for rice leaf diseases identification using machine learning methods. *Frontiers in plant science*, *11*, 577063. https://doi.org/10.3389/fpls.2020.577063

Foreign Agricultural Service, USDA (2025). *Production - Rice*. https://www.fas.usda.gov/data/production/commodity/0422110

Gareth, J., Daniela, W., Trevor, H., & Robert, T. (2013). *An introduction to statistical learning: with applications in R*. Springer. https://doi.org/10.1080/24754269.2021.1980261

Gonzales, R. C., & Wintz, P. (2017). *Digital image processing*. Addison Wesley Longman Publishing Co., Inc.

Gonzalez, R. C. & Woods, R. E. (2018). *Digital image processing*. Pearson.

IPM Images, IPM (2023). *Agricultural Systems : Rice*. IPM Images. https://www.ipmimages.org/browse/Areasubs.cfm?area=125

International Rice Research Institute, IRRI (2023). *Pests and diseases: Diseases*. IRRI Knowledge Bank. http://www.knowledgebank.irri.org/step-by-step-production/growth/pests-and-diseases/diseases

Islam, M., Dinh, A., Wahid, K., & Bhowmik, P. (2017). Detection of potato diseases using image segmentation and multiclass support vector machine. In *2017 IEEE 30th canadian conference on electrical and computer engineering (CCECE)* (pp. 1-4). IEEE. https://doi.org/10.1109/CCECE.2017.7946594

Jiang, Z., Dong, Z., Jiang, W., & Yang, Y. (2021). Recognition of rice leaf diseases and wheat leaf diseases based on multi-task deep transfer learning. *Computers and Electronics in Agriculture*, *186*, 106184. https://doi.org/10.1016/j.compag.2021.106184

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, *349*(6245), 255-260. https://doi.org/10.1126/science.aaa8415

Kanan, C., & Cottrell, G. W. (2012). Color-to-grayscale: does the method matter in image recognition? *PloS one*, *7*(1), e29740. https://doi.org/10.1371/journal.pone.0029740

Kartikeyan, P., & Shrivastava, G. (2021). Review on emerging trends in detection of plant diseases using image processing with machine learning. *International Journal of Computer Applications*, *975*(8887), 39-48. https://doi.org/10.5120/ijca2021920990

Kawtrakul, A., Tippayarak, P., Andres, F., & Ujjin, S. (2015). Personal warning service for pest management using a crop calendar and bus model. In *Proceedings of the 7th International Conference on Management of Computational and Collective*

Intelligence in Digital Ecosystems (pp. 242–249). https://doi.org/10.1145/2857218.285727

Khan, M. A., Khan, R., & Ansari, M. A. (2022). *Application of Machine Learning in Agriculture*. Academic Press.

Kumar, K. K. & Kannan, E. (2022). Detection of rice plant disease using AdaBoostSVM classifier. *Agronomy journal*, 114(4), 2213-2229. https://doi.org/10.1002/agj2.21070

Latif, G., Abdelhamid, S. E., Mallouhy, R. E., Alghazo, J., & Kazimi, Z. A. (2022). Deep learning utilization in agriculture: Detection of rice plant diseases using an improved CNN model. *Plants*, 11(17), 2230. https://doi.org/10.3390/plants11172230

Lim, W. H. & Isa, N. A. M. (2011). Color to grayscale conversion based on neighborhood pixels effect approach for digital image. In *Proc. int. conf. on electrical and electronics engineering* (pp. 157-161).

Lu, Y., Yi, S., Zeng, N., Liu, Y., & Zhang, Y. (2017). Identification of rice diseases using deep convolutional neural networks. *Neurocomputing*, 267, 378–384. https://doi.org/10.1016/j.neucom.2017.06.023

Ma, K., Zhao, T., Zeng, K., & Wang, Z. (2015). Objective quality assessment for color-to-gray image conversion. *IEEE Transactions on Image Processing*, 24(12), 4673-4685. https://doi.org/10.1109/TIP.2015.246001

Marsland, S. (2015). *Machine learning: an algorithmic perspective*. Chapman and Hall/CRC.

Mendeley (2022). *Mendeley data*. https://data.mendeley.com

Nakandakari, D. L. H. (2017). *Pest management issues in rice cultivation (Oryza sativa L.)* [Master's thesis, National Agrarian University La Molina]. UNALM-Institutional Repository. https://hdl.handle.net/20.500.12996/2988

Nixon, M. & Aguado, A. (2019). *Feature extraction and image processing for computer vision*. Academic press. https://doi.org/10.1016/C2011-0-06935-1

Rafael, C., Richard, E., Woods, L. S., & Steven, L. (2020). *Digital image using matlab processing*. Person Prentice Hall, Lexington.

Rahman, H. A. A., Wah, Y. B., He, H., & Bulgiba, A. (2015). Comparisons of ADABOOST, KNN, SVM and logistic regression in classification of imbalanced dataset. In *Soft Computing in Data Science: First International Conference, SCDS 2015, Putrajaya, Malaysia, September 2-3, 2015, Proceedings 1* (pp. 54-64). Springer. https://doi.org/10.1007/978-981-287-936-3_6

Rajinikanth, V., Raja, N. S. M., & Dey, N. (2020). *A Beginner's Guide to Multilevel Image Thresholding*. CRC Press. https://doi.org/10.1201/9781003049449

Rallapalli, S., & Saleem Durai, M. A. (2021). A contemporary approach for disease identification in rice leaf. *International Journal of System Assurance Engineering and Management*, 1-11. https://doi.org/10.1007/s13198-021-01159-y

Reddy, S. R., Varma, G. S., & Davuluri, R. L. (2022). Deep neural network (dnn) mechanism for identification of diseased and healthy plant leaf images using computer vision. *Annals of Data Science*, 11(1), 243-272. https://doi.org/10.1007/s40745-022-00412-w

Rodriguez, E. Y. A., Gamboa, A. A. R., Rodriguez, E. C. A., da Silva, A., Rizol, P. M. S. R., & Marins, F. A. S. (2022). Comparison of adaptative neuro-fuzzy inference system (ANFIS) and machine learning algorithms for electricity production forecasting. *IEEE Latin America Transactions*, 20(10):2288–2294. https://doi.org/10.1109/TLA.2022.9885166

Rodríguez, E. Y. A., Rodríguez, E. C. A., Silva, A. F. d., Rizol, P. M. S. R., Miranda, R. d. C., & Marins, F. A. S. (2024a). A decision-making framework with machine learning for transport outsourcing based on cost prediction: an application in a multinational automotive company. *International Journal of Information Technology*, 16(3), 1495-15. https://doi.org/10.1007/s41870-023-01707-8

Rodríguez, E. Y. A., Rodríguez, E. C. A., Silva, A. F. d., Rizol, P. M. S. R., Miranda, R. d. C., & Marins, F. A. S. (2024b). Analysis of machine learning integration into supply chain management. *International Journal of Logistics Systems and Management*, 47(3), 327–355. https://doi.org/10.1504/IJLSM.2021.10042452

Rodríguez, E. Y. A., Rodríguez, E. C. A., Fernández, J. P. S., Nascimento, L. F. C., Silva, A. F. d., & Marins, F. A. S. (2022). Inteligencia artificial para detección y diagnóstico de enfermedades en cultivos de arroz. In *5to Congreso estudiantil de Inteligencia Artificial aplicada a la ingeniería y tecnología*. https://virtual.cuautitlan.unam.mx/intar/ceiaait/wp-content/uploads/sites/14/2023/02/Int-Art-10-18.pdf

Savary, S., Willocquet, L., Pethybridge, S. J., Esker, P., McRoberts, N., & Nelson, A. (2019). The global burden of pathogens and pests on major food crops. *Nature ecology & evolution*, 3, 430–493. https://doi.org/10.1038/s41559-018-0793-y

Seck, P. A., Diagne, A., Mohanty, S., & Wopereis, M. C. S. (2012). Crops that feed the world 7: Rice. *Food Security*, 4(1), 7–24. https://doi.org/10.1007/s12571-012-0168-1

Sethy, P. K., Barpanda, N. K., Rath, A. K., & Behera, S. K. (2020). Deep feature-based rice leaf disease identification using support vector machine. *Computers and Electronics in Agriculture*, 175, 105527. https://doi.org/10.1016/j.compag.2020.105527

Sharma, R., Singh, A., Kavita, Jhanji, N. Z., Masud, M., Jaha, E. S., & Verma, S. (2022). Plant disease diagnosis and image classification using deep learning. *Computers, Materials & Continua*, 71(2), 2125–2140. https://doi.org/10.32604/cmc.2022.020017

Sheykhmousa, M., Mahdianpari, M., Ghanbari, H., Mohammadimanesh, F., Ghamisi, P., & Homayouni, S. (2020). Support vector machine versus random forest for remote sensing image classification: A meta-analysis and systematic review. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13, 6308-6325. https://doi.org/10.1109/JSTARS.2020.3026724

Tian, L., Xue, B., Wang, Z., Li, D., Yao, X., Cao, Q., Zhu, Y., Cao, W., & Cheng, T. (2021). Spectroscopic detection of rice leaf blast infection from asymptomatic to mild stages with integrated machine learning and feature selection. *Remote Sensing of Environment*, 257, 112350. https://doi.org/10.1016/j.rse.2021.112350

Upadhyay, S. K., & Kumar, A. (2022). A novel approach for rice plant diseases classification with deep convolutional neural network. *International Journal of Information Technology*, 14(1), 185-199. https://doi.org/10.1007/s41870-021-00817-5

Zhai, Z., Martínez, J. F., Beltran, V., & Martínez, N. L. (2020). Decision support systems for agriculture 4.0: Survey and challenges. *Computers and Electronics in Agriculture*, 170, 105256. https://doi.org/10.1016/j.compag.2020.105256

Zhu, Q. (2020). On the performance of Matthews correlation coefficient (MCC) for imbalanced dataset. *Pattern Recognition Letters*, 136, 71-80. https://doi.org/10.1016/j.patrec.2020.03.030