World Journal of Advanced Research and Reviews
eISSN: 2581-9615 CODEN (USA): WJARAI
Cross Ref DOI: 10.30574/wjarr
Journal homepage: https://wjarr.com/

(RESEARCH ARTICLE)

CATCHING RICE DISEASES EARLY: WAVELET TEXTURE, AGGD MODELING, AND AN SVM CLASSIFIER

Alphonse AZOMA [1, *], Boris ZANNOU[2], Ibrahima TOURE [2] and Jean Ouèrè TOUPOUVOGUI [3]

[1] Department of Biomedical Equipment Technology, Mamou Higher Institute of Technology, Guinea.
[2] Department of Computer Science, Mamou Higher Institute of Technology, Guinea.
[3] Department of Instrumentation and Physical Measurements, Mamou Higher Institute of Technology, Guinea.
* Corresponding Author; Email: alphonseazoma@gmail.com

ORCID Details
Alphonse AZOMA: https://orcid.org/0009-0002-4687-9674
Boris ZANNOU: https://orcid.org/0009-0003-1246-7193
Jean Ouèrè TOUPOUVOGUI: https://orcid.org/0009-0008-8142-3430

World Journal of Advanced Research and Reviews, 2026, 32(02), 1478–1486

Publication history: Received on 10 October 2026; revised on 16 November 2026; accepted on 18 November 2026

Article DOI: https://doi.org/10.30574/wjarr.2026.32.2.2962

Copyright © 2026 Author(s) retain the copyright of this article. This article is published under the terms of the Creative Commons Attribution License 4.0

ABSTRACT

Rice diseases caused by fungi, bacteria and viruses threaten food security, and they are hard to catch early. Visual inspection is subjective and often comes too late, while laboratory diagnosis is slow, hard to access and unsuited to real-time use. We propose AGGD-MM-SVM, an early-diagnosis method that analyzes the texture of macroscopic rice-leaf images and extends to rice an approach we first developed for tomato leaves. Working with public images of healthy and infected leaves, we decompose each image with a two-level Haar wavelet transform in MATLAB R2024b and fit the histogram of every detail image with an asymmetric generalized Gaussian distribution (AGGD) whose mean may differ from zero. The parameters of this distribution, estimated by the method of moments (MM), serve as texture markers for a support vector machine (SVM). On a held-out test set, the classifier reaches 98.33% accuracy, compared with 97.6% for Prototypical Networks. A handful of statistical texture markers can therefore compete with deep networks for the early diagnosis of rice diseases, although larger field datasets are needed to confirm this result.

Keywords: Early detection; Rice leaf diseases; Wavelet transform; Asymmetric generalized Gaussian distribution; Texture characterization; Support vector machine

1 INTRODUCTION

Rice is the staple food of more than half of the world's population, so any threat to its production is a threat to food security. Fungal, bacterial and viral diseases rank among the most serious of these threats, especially in countries whose economies depend on agriculture [1]. They are estimated to cut rice yields by 10–15% [2] and also lower the quality of the harvested grain. The sooner a disease is identified, the more of the crop can be saved.

In practice, diagnosis still rests largely on visual inspection, which is slow and calls for expertise that many farmers lack [3]. This has driven a large body of work on recognizing rice diseases automatically from images [4]. Deep learning methods can now classify plant diseases directly from leaf photographs [5] and deliver a diagnosis in the field quickly and cheaply [6]. Their accuracy, however, depends on annotated image collections such as the public rice-leaf disease dataset of Shah et al. [7], and the features they learn come from deep networks that demand substantial data and computing power and are hard to interpret. Far less attention has gone to approaches that describe leaf texture with a few interpretable statistical parameters. In earlier work, we proposed one such approach for tomato leaves: we modeled the histograms of Haar wavelet detail images with an asymmetric generalized Gaussian distribution (AGGD) with a non-zero mean and classified the resulting parameters with k-nearest neighbors [8].

Here, we extend that approach to the early detection of rice diseases from macroscopic leaf images. Our contributions are threefold:

- We transfer the non-zero-mean AGGD texture model of [8] from tomato leaves to rice leaves affected by bacterial leaf blight, brown spot and leaf smut.
- We determine which directional shape parameters β, estimated by the method of moments, best separate healthy from infected rice leaves. This selection differs from the one retained for tomato.
- We replace the k-nearest-neighbor classifier of [8] with an SVM using a radial basis function (RBF) kernel. On a held-out test set, it reaches 98.33% accuracy, ahead of the 97.6% obtained by Prototypical Networks.

Section 1.1 reviews related work, Section 2 describes the data and the method, Section 3 discusses the results, and Section 4 concludes.

1.1 Related work

Most recent studies on rice-disease recognition rely on deep learning. Hasan et al. [4] use convolutional neural networks (CNNs) to extract features from images of damaged leaves and an SVM to classify the disease, with an average recognition rate of 96.8%. Krishnamoorthy et al. [9] combine InceptionResNetV2 with transfer learning and reach 95.67% accuracy. Ritharson et al. [10] assembled their own collection of 5,932 rice-leaf photographs and report 99.94% accuracy with a customized VGG16. Thai-Nghe et al. [11] trained an EfficientNet by transfer learning and deployed it in an Android application, with a validation accuracy of 95%.

Other studies build multi-stage pipelines or focus on explainability. Dubey et al. [12] chain pre-processing, feature extraction, feature selection and classification, ending with an adaptive bidirectional long short-term memory (ABi-LSTM) classifier. Sowmiyan et al. [13] apply explainable AI (XAI) to disease prediction in paddy crops.

Rather than learning features with a deep network, our method describes leaf texture with a compact statistical model of the wavelet coefficients, namely AGGD parameters estimated by the method of moments, and classifies these markers with an SVM. The aim is to let farmers treat only the infected areas, earlier and with fewer inputs.

2 MATERIAL AND METHODS

Figure 1 summarizes the AGGD-MM-SVM pipeline. Up to the classification stage, it follows the pipeline we developed for tomato leaves [8]: the images are collected and pre-processed, their texture is extracted and described by statistical parameters, an SVM classifies them, and four standard measures assess the result.

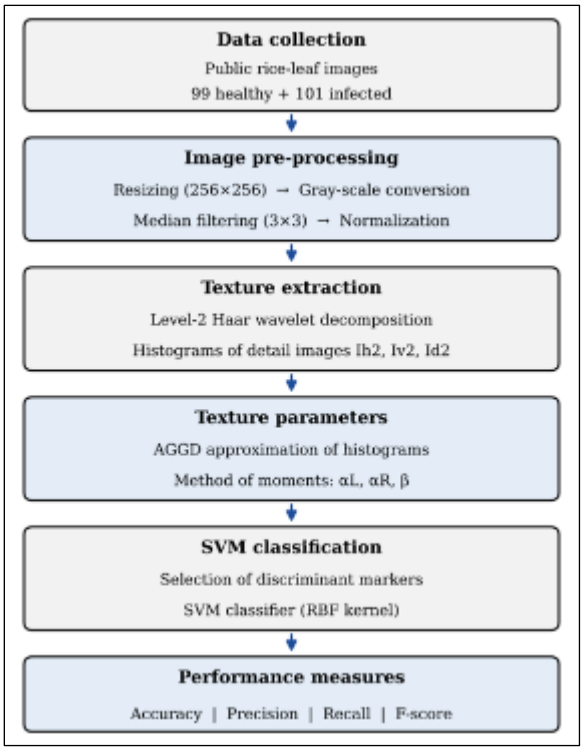


Figure 1: Framework of the proposed AGGD-MM-SVM system.

2.1 Dataset

The dataset contains 200 macroscopic images of rice leaves (Table 1). The 101 infected leaves come from the public Rice Leaf Diseases dataset [7], also available on Kaggle, and show three diseases at an early stage of infection: bacterial leaf blight, brown spot and leaf smut. The 99 healthy leaves are freely licensed images found through a Google Images search. We set aside 60 images for testing (30 per class) and use the remaining 140 for training and validation. All images are RGB JPEG files showing the leaf on a white background; Figure 2 shows examples. All algorithms were implemented and run in MATLAB R2024b.

Table 1: Composition of the image database

Class	Images	Composition	Conditions
Healthy leaves	99	Healthy (99)	White background
Infected leaves	101	Bacterial leaf blight (35), brown spot (35), leaf smut (31)	White background

Sources: infected leaves, Rice Leaf Diseases dataset [7] (UCI/Kaggle); healthy leaves, freely licensed images collected through Google Images.

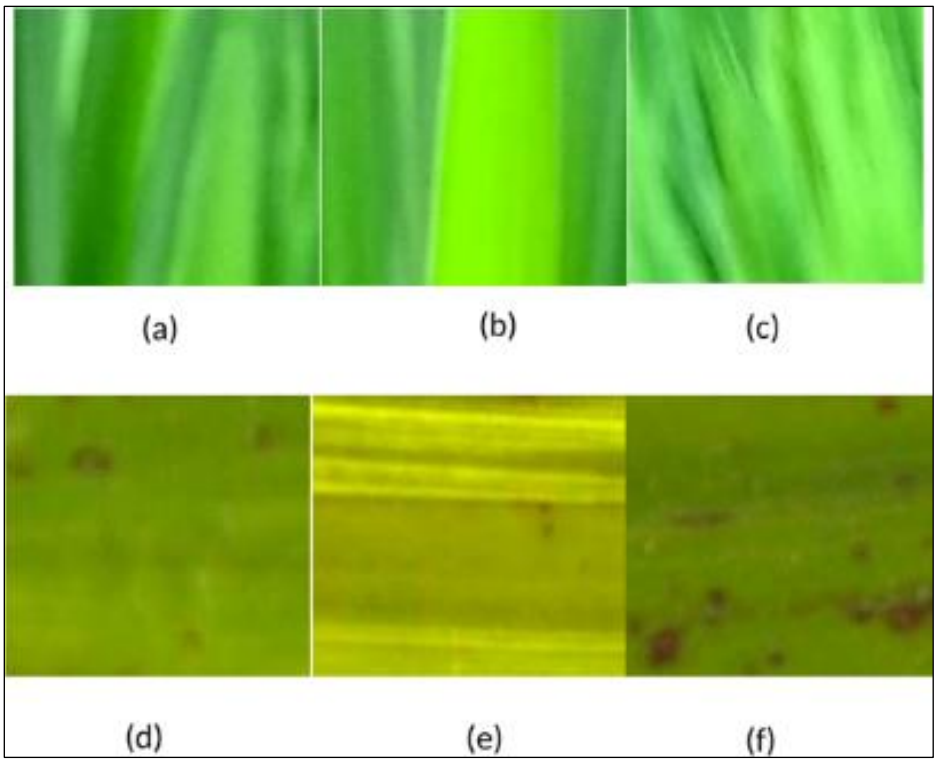


Figure 2: Healthy rice leaves (a)–(c) and infected rice leaves (d)–(f).

2.2 Image pre-processing

As in [8], each image goes through four pre-processing steps before texture extraction.

1) Resizing. Every source image J (height H, width W) is brought to a common size of 256 × 256 pixels; the coordinates of the resized image I map back to J through

$$x_J = x_I \cdot W / 256, \quad y_J = y_I \cdot H / 256 \qquad (1)$$

2) Gray-scale conversion. The three color channels are combined into a single luminance channel:

$$I_{gray}(x,y) = 0.299\, R(x,y) + 0.587\, G(x,y) + 0.114\, B(x,y) \quad (2)$$

with R, G and B the intensities (0–255) of the red, green and blue channels.

3) Filtering. Isolated outlier pixels are suppressed by replacing each pixel with the median of its 3 × 3 neighborhood S(x,y):

$$h(x,y) = \text{median}\{ I(m,n) \mid (m,n) \in S(x,y) \} \qquad (3)$$

4) Normalization. Gray levels are then mapped linearly onto the interval [0, 1]:

$$I_{norm}(x,y) = (I(x,y) - I_{min}) / (I_{max} - I_{min}) \qquad (4)$$

In this expression, I(x,y) denotes the gray level of pixel (x,y), while I_{min} and I_{max} are the minimum and maximum gray levels of the image.

2.3 Texture extraction by Haar wavelet decomposition

A wavelet transform represents a signal x(t) with a family of functions $\psi_{a,b}(t)$, derived from a single mother wavelet $\psi(t)$ by dilation and shift:

$$C_{a,b} = \int x(t)\, \psi_{a,b}(t)\, dt \qquad (5)$$

$$\psi_{a,b}(t) = (1/\sqrt{|a|})\, \psi((t-b)/a), \quad a \neq 0 \qquad (6)$$

where a sets the scale and b the position. The resulting coefficients $C_{a,b}$ give a multi-scale view of x(t) [14]. In two dimensions, the image is filtered row-wise and then column-wise by low-pass (g) and high-pass (h) filters, which yields one approximation and three detail sub-images at each level. Following [8], the pre-processed leaf image is decomposed over two levels with the Haar wavelet (Figure 3) [15]. At level 2, the detail sub-images I_{h2}, I_{v2}, and I_{d2} carry the high-frequency content of the leaf, that is, its texture along the horizontal, vertical and diagonal directions [16].

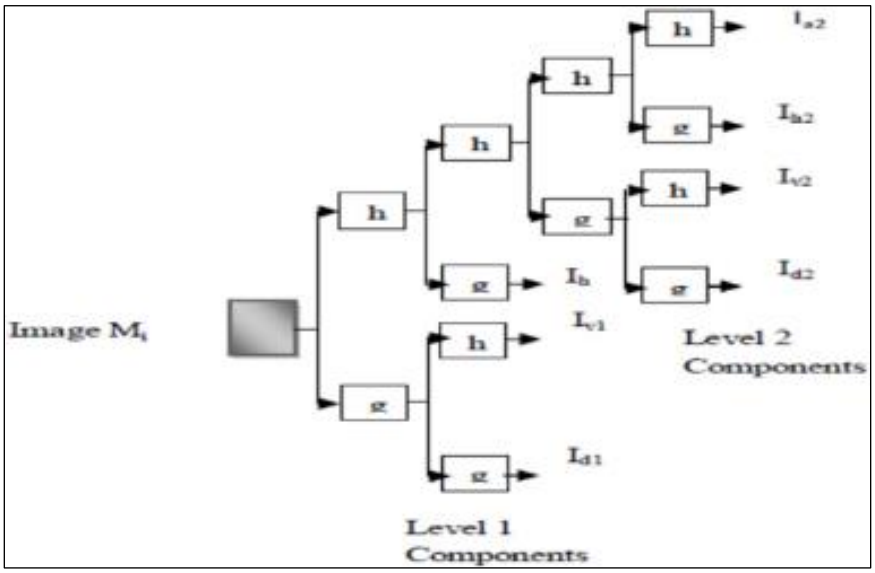


Figure 3: Two-level wavelet decomposition [15]

The histogram h(k) of a detail image of size H × W counts the pixels at each intensity level k:

$$h(k) = \Sigma_{x=0}^{W-1} \Sigma_{y=0}^{H-1} \delta(I(x,y) - k) \quad (7)$$

Here k ranges over the intensity levels (from 0 to 1 after normalization) and δ denotes the Kronecker delta, equal to 1 at 0 and to 0 elsewhere.

2.4 AGGD modeling of the detail histograms

Detail histograms encode the leaf texture [16], yet they are noisy because of outliers and empty bins. Fitting a parametric model to them yields a smoother and more compact description. Detail histograms are commonly modeled by a zero-mean AGGD [17]; as in our tomato study [8], we let the mean μ differ from zero so that the model can follow the shift that disease introduces in the histogram:

$$P(x) = \beta / ((\alpha_L+\alpha_R)\, \Gamma(1/\beta)) \cdot \exp[\, -(|x-\mu| / \alpha_L)^{\beta}], \; x < \mu \qquad (8)$$

$$P(x) = \beta / ((\alpha_L+\alpha_R)\, \Gamma(1/\beta)) \cdot \exp[\, -(|x-\mu| / \alpha_R)^{\beta}], \; x \geq \mu \qquad (9)$$

Here α_L and α_R control the spread on each side of μ, β controls the shape of the peak, Γ is the gamma function, and x is the gray level of a pixel of the detail image.

2.5 Parameter estimation by the method of moments

The parameters α_L, α_R, and β are estimated by the method of moments [17]:

$$\alpha_L = [\, \Gamma(1/\beta)\, \Sigma_{x_i<\mu} x_i^2 / (N_L\, \Gamma(3/\beta)) \,]^{1/2} \qquad (10)$$

$$\alpha_R = [\, \Gamma(1/\beta)\, \Sigma_{x_i\geq\mu} x_i^2 / (N_R\, \Gamma(3/\beta)) \,]^{1/2} \qquad (11)$$

in which μ is the average gray level over the n pixels of the detail image,

$$\mu = (1/n)\, \Sigma_{i=1}^{n} x_i \quad (12)$$

and N_L and N_R are the numbers of pixels lying below and above this mean:

$$N_L = card\{ x_i : x_i < \mu \}, \quad N_R = card\{ x_i : x_i \geq \mu \} \qquad (13)$$

The asymmetry of the histogram is measured by the ratio γ:

$$\gamma = [(1/N_L) \Sigma_{x_i<\mu} x_i^2 / ((1/N_R) \Sigma_{x_i \geq \mu} x_i^2)]^{1/2} \quad (14)$$

and β is obtained by inverting the generalized Gaussian ratio function GGR:

$$\beta = GGR^{-1}[(1+\gamma)(1+\gamma^3)(\Sigma|x_i|)^2 / (N (1+\gamma^2)^2 \Sigma x_i^2)] \quad (15)$$

For each detail image, β is retained as the texture marker. For healthy leaves, the diagonal, horizontal, and vertical textures are described by β_{sd}, β_{sh}, and β_{sv}. For infected leaves, they are described by β_{id}, β_{ih}, and β_{iv}.

2.6 SVM classification

The marker pair (β_{sh}, β_{sv}) identifies healthy leaves, and (β_{id}, β_{ih}) identifies infected ones. A support vector machine (SVM), which replaces the k-nearest-neighbor classifier used in [8], then separates the two classes. We score the model on a held-out test set of 60 images, 30 per class, with four standard measures: accuracy, precision, recall, and F-score. Infected leaves are the positive class.

The SVM relies on a Gaussian radial basis function (RBF) kernel. Its hyperparameters were set by the optimization routine of MATLAB R2024b, which chose a box constraint C = 10 and a kernel scale of 0.1. The markers are standardized before training, and the model is validated by 5-fold cross-validation.

3 RESULTS AND DISCUSSION

3.1 Discriminative power of the markers

Figure 4 plots the selected markers for both classes. Infected leaves sit at low values of the horizontal-axis marker (β_{sh}, β_{id}) and climb to high values of the vertical-axis marker (β_{sv}, β_{ih}). Healthy leaves spread along the horizontal axis and stay low on the vertical one. The two classes separate well, with some overlap near the origin, so the selected markers give the classifier a sound basis for telling healthy from infected leaves.

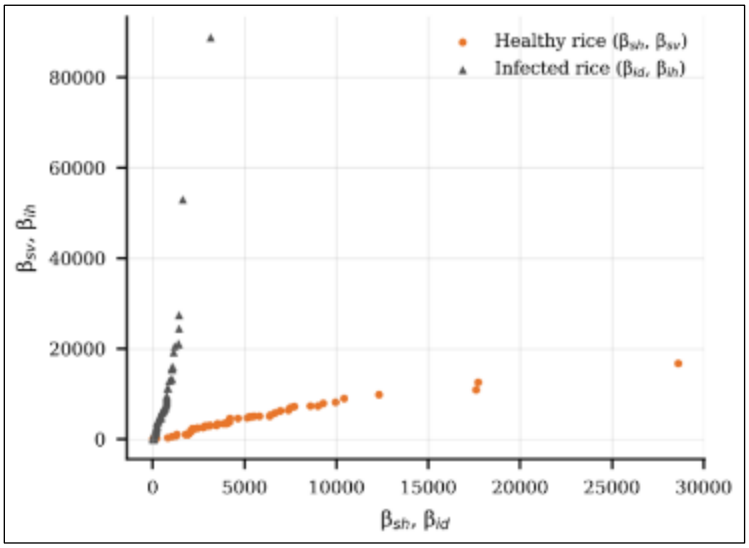


Figure 4: Separation of healthy and infected rice leaves in the marker space.

3.2 Performance of AGGD-MM-SVM

Table 2 shows the confusion matrix on the held-out test set of 60 images. All 30 infected leaves are detected, while one of the 30 healthy leaves is wrongly flagged as infected. Table 3 gives the resulting measures: 98.33% accuracy (59 of 60 images), 96.77% precision, 100% recall and a 98.36% F-score. Five-fold cross-validation yielded 95.00% accuracy, that is, a validation error of 5.00%.

Table 2: Confusion matrix on the test set

	Pred. infected	Pred. healthy
True infected	30	0
True healthy	1	29

Table 3: Performance of the AGGD-MM-SVM model

Accuracy	Precision	Recall	F-score
98.33%	96.77%	100%	98.36%

3.3 Comparison with existing models

Table 4: Comparison of model performance (%)

Model	Accuracy	Precision	Recall	F-score
RES50	90.98	90.55	90.33	90.37
VGG-16	89.02	89.98	89.58	89.70
VGG-19	90.87	90.90	90.25	90.98
MN-1	88.98	88.86	88.40	88.79
MN-2	87.98	88.98	88.78	86.69
INC	89.79	89.23	89.93	89.56
XP	90.01	90.89	91.90	92.86
DN	88.39	88.25	88.28	88.93
Prototypical Networks	97.60	97.60	97.30	97.60
AGGD-MM-SVM	98.33	96.77	100	98.36

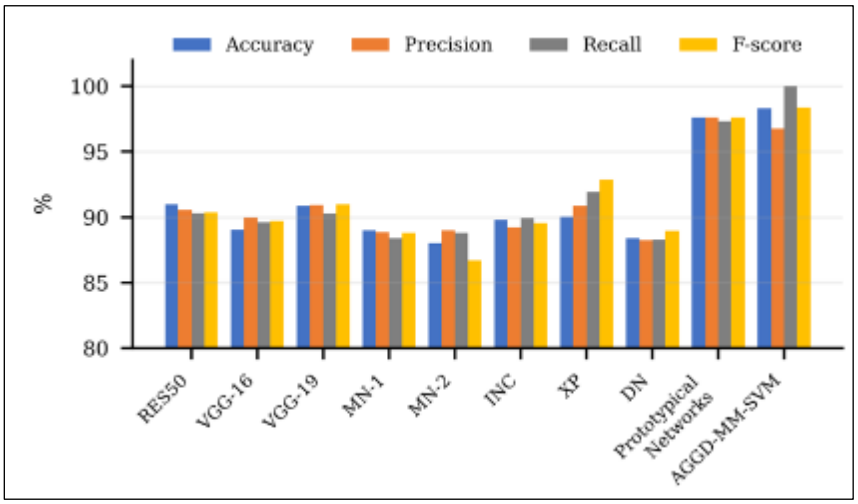


Figure 5: Comparison of model performance.

Table 4 and Figure 5 compare AGGD-MM-SVM with the models reported in [18]. Our model has the highest accuracy (98.33%), recall (100%) and F-score (98.36%), while its precision (96.77%) is slightly lower than that of Prototypical Networks (97.60%). The remaining models reach between 87.98% and 90.98% accuracy. Because the test set holds

only 60 images, a single error shifts the accuracy by 1.7 points, so the margin over Prototypical Networks should be interpreted with caution.

We attribute this gain to the high-frequency components of the wavelet decomposition, which respond to the slight texture changes that disease causes in the leaf. This sensitivity could allow detection before symptoms become clearly visible, although our dataset does not test this directly.

4 CONCLUSION

We have presented AGGD-MM-SVM, a method for the early detection of rice diseases that models the distribution of wavelet coefficients and extends our earlier tomato study [8] to rice. The method fits an AGGD with a non-zero mean to the histograms of the detail images, estimates its parameters by the method of moments and classifies the resulting markers with an SVM. On our test set of healthy and infected rice leaves, it outperforms Prototypical Networks in accuracy. We used the Haar wavelet here, but other wavelet families could be substituted. AGGD-MM-SVM thus appears to be a competitive alternative to Prototypical Networks for image-based early detection; larger datasets and images taken in the field are still needed to confirm this

STATEMENTS ON COMPLIANCE WITH ETHICAL STANDARDS

Acknowledgments

The authors used artificial intelligence (AI) tools to support the preparation of this manuscript, specifically for grammar and language editing, for formatting the text, equations, and tables to the journal template, and for formatting and redrawing figures. All scientific content, including the method, experiments, results, and conclusions, was designed, verified, and approved by the authors, who take full responsibility for the paper.

Disclosure of conflict of interest

The authors declare that they have no conflict of interest.

REFERENCES

[1] Fang Y, Ramasamy RP. Current and prospective methods for plant disease detection. Biosensors (Basel). 2015;5(3):537-61. doi:10.3390/bios5030537.

[2] Tiwari V, Joshi RC, Dutta MK. Dense convolutional neural networks based multiclass plant disease detection and classification using leaf images. Ecol Inform. 2021;63:101289. doi:10.1016/j.ecoinf.2021.101289.

[3] Mohanty SP, Hughes DP, Salathé M. Using deep learning for image-based plant disease detection. Front Plant Sci. 2016;7:1419. doi:10.3389/fpls.2016.01419.

[4] Hasan MJ, Mahbub S, Alom MS, Nasim MA. Rice disease identification and classification by integrating support vector machine with deep convolutional neural network. In: Proceedings of the 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT); Dhaka, Bangladesh. IEEE; 2019. p. 1-6. doi:10.1109/ICASERT.2019.8934568.

[5] Kamilaris A, Prenafeta-Boldú FX. Deep learning in agriculture: a survey. Comput Electron Agric. 2018;147:70-90. doi:10.1016/j.compag.2018.02.016.

[6] Ferentinos KP. Deep learning models for plant disease detection and diagnosis. Comput Electron Agric. 2018;145:311-8. doi:10.1016/j.compag.2018.01.009.

[7] Shah JP, Prajapati HB, Dabhi VK. Rice leaf diseases [dataset]. UCI Machine Learning Repository; 2017. doi:10.24432/C5R013.

[8] Azoma A, Zannou B, Didavi KBA, Alihonou CA, Toupouvougui JO, Assogba MK. Early detection of tomato diseases based on a novel hybrid method combining wavelet decomposition, asymmetric generalized Gaussian distribution, and KNN classification. In: Doctoral Workshop on Artificial Intelligence and Advanced Digital Technologies in Smart Cities, Healthcare, and Banking Applications. CEUR Workshop Proceedings. 2026;4232:154-64.

[9] Krishnamoorthy N, Narasimha Prasad LV, Pavan Kumar CS, Subedi B, Abraha HB, Sathishkumar VE. Rice leaf diseases prediction using deep neural networks with transfer learning. Environ Res. 2021;198:111275. doi:10.1016/j.envres.2021.111275.

[10] Ritharson PI, Raimond K, Mary XA, Robert JE, J A. DeepRice: a deep learning and deep feature based classification of rice leaf disease subtypes. Artif Intell Agric. 2024;11:34-49. doi:10.1016/j.aiia.2023.11.001.

[11] Thai-Nghe N, Tri NT, Hoa NH. Deep learning for rice leaf disease detection in smart agriculture. In: Proceedings of the International Conference on Artificial Intelligence and Big Data in Digital Era. Cham: Springer; 2021.

[12] Dubey RK, Choubey DK. An efficient adaptive feature selection with deep learning model-based paddy plant leaf disease classification. Multimed Tools Appl. 2024;83(8):22639-61. doi:10.1007/s11042-023-16247-3.

[13] Sowmiyan L, Vaidya S, Karpagam GR. An explainable AI (XAI)-based framework for detecting diseases in paddy crops. In: Proceedings of the International Conference on Data Science and Applications (ICDSA 2023). Singapore: Springer Nature; 2023. doi:10.1007/978-981-99-7817-5_31.

[14] Grossmann A, Morlet J. Decomposition of Hardy functions into square integrable wavelets of constant shape. SIAM J Math Anal. 1984;15(4):723-36.

[15] Van de Wouwer G, Scheunders P, Van Dyck D. Statistical texture characterization from discrete wavelet representations. IEEE Trans Image Process. 1999;8(4):592-8.

[16] Lefebvre A. Contribution de la texture pour l'analyse d'images à très haute résolution spatiale : application à la détection de changement en milieu périurbain [dissertation]. Rennes (France): Université Rennes 2; 2011.

[17] Lasmar N. Modélisation stochastique pour l'analyse d'images texturées : approches bayésiennes pour la caractérisation dans le domaine des transformées [dissertation]. Bordeaux (France): Université Bordeaux I; 2012.

[18] El-Behery H, Attia AF, Rezk NG. An explainable deep learning framework for few shot crop disease detection in rice and sugarcane using CNN based feature extraction. Sci Rep. 2026;16:8272. doi:10.1038/s41598-026-37501-2.