DEEP LEARNING-BASED ANOMALY DETECTION SYSTEM FOR RICE CROP HEALTH MONITORING

DAYRA-Y. GARCÍA-RAMÍREZ[1]; ANDRÉS-F. JIMÉNEZ-LÓPEZ[2], FABIÁN-R. JIMÉNEZ-LÓPEZ[3]

[1]Professor. Faculty of Agricultural Sciences and Natural Resources, Universidad de los Llanos, DIGITAL-FAS & FarmTechnology Research Groups. Villavicencio, Colombia.

[2]Full Professor. Faculty of Basic Sciences and Engineering – Universidad de los Llanos – Macrypt & FarmTechonology Research Groups. Villavicencio, Colombia.

[3]Associate Professor. Electronic Engineering School, Universidad Pedagógica y Tecnológica de Colombia, I^{2}E Research Group. Tunja, Colombia.

A paper from the Proceedings of the

17th International Conference on Precision Agriculture and 11th Brazilian Congress on Precision and Digital Agriculture

13-15 July 2026

Porto Alegre, Brazil

Abstract.

Global food security relies heavily on the stable production of rice (Oryza sativa L.), yet cultivation remains vulnerable to various phytosanitary anomalies, including foliar diseases like Pyricularia and Rhynchosporium, scald, and abiotic stressors such as herbicide damage. Traditional agronomic management relies on visual scouting, which is inherently subjective, labor-intensive, and often leads to delayed interventions. This study proposes an automated, high-throughput solution for real-time plant health monitoring by leveraging advanced analytics and computer vision to bridge the gap between field-level scouting and precision crop protection.

The methodological approach involved the construction of a robust, multi-source dataset comprising high-resolution imagery captured via Unmanned Aerial Vehicles (UAVs) and smartphone devices. This dual-capture strategy ensures the model's versatility across different operational scales, from localized leaf-level diagnosis to broad-acre field monitoring. The dataset was meticulously labeled to include healthy crop samples alongside various anomalies, providing a representative baseline for training deep learning models under realistic environmental conditions.

To identify the most effective solution for digital agriculture, this research evaluated and compared three state-of-the-art Convolutional Neural Network (CNN) architectures: ResNet50, MobileNet, VGG16 and Inception. These models were selected to balance the trade-off between predictive depth and computational efficiency. Model performance was quantified using a rigorous suite of metrics, including precision, recall (completeness), and the F1-score, ensuring that the detection system remains reliable despite the inherent complexities of field data.

Experimental results demonstrated that the MobiletNet-based architecture achieved a superior

The authors are solely responsible for the content of this paper, which is not a refereed publication. Citation of this work should state that it is from the Proceedings of the 17th International Conference on Precision Agriculture and 11th Brazilian Congress on Precision and Digital Agriculture. EXAMPLE: Last Name, A. B. & Coauthor, C. D. (2026). Title of paper. In Proceedings of the 17th International Conference on Precision Agriculture and 11th Brazilian Congress on Precision and Digital Agriculture (unpaginated, online). Monticello, IL: International Society of Precision Agriculture and Brazilian Congress on Precision and Digital Agriculture.

accuracy of 90%, outperforming ResNet50, VGG16 and Inception in identifying subtle patterns associated with biotic stress and herbicide-induced chlorosis. However, the study also highlights critical challenges in deploying machine learning within agricultural environments. These include the high visual similarity between different physiological anomalies, extreme variability in lighting and capture angles, and the necessity of addressing class imbalance in the training data to prevent model bias.

The implementation of this deep learning framework offers significant practical impact by drastically reducing diagnosis times and enabling Site-Specific Crop Management (SSCM). By facilitating the targeted application of inputs, this technology optimizes resource use and improves the overall sustainability of rice production systems. This research contributes a viable tool to the field of advanced analytics in agriculture, positioning automated anomaly detection as a cornerstone of modern, data-driven plant health monitoring.

Keywords.
Anomaly Detection; Computer Vision in Agriculture; Deep Learning; Rice Plant Health Monitoring; Precision Crop Protection.

Introduction

Rice (*Oryza sativa L.*) is the most important source of calories and nutrients for more than half of the world's population and is a key factor in the socioeconomic stability of developing countries, particularly in Asia and Latin America (Hervert 2022; Gopi and Kondaveeti 2023). Within the context of contemporary demographic growth, global food security strictly depends on the optimization of yields for this cereal.

However, rice productivity is consistently compromised by a complex matrix of abiotic and biotic factors that degrade grain quality and precipitate large-scale financial losses in the agricultural sector (Shah et al. 2024). Among these factors, phytopathological pressures represent the most persistent threat to the integrity of production systems.

Agricultural literature identifies several high-virulence pathogens that systematically devastate cultivated areas. Key diseases include leaf blast (*Magnaporthe oryzae*), false smut (*Ustilaginoidea virens*), sheath blight (*Rhizoctonia solani*), and brown spot (*Cochliobolus miyabeanus*) (Deng et al. 2021). Historically, the diagnosis of these anomalies has depended on manual visual inspection by farmers or specialized technicians (Sethy et al. 2020).

This paradigm is currently evaluated as insufficient due to the scarcity of experts, high labor costs, and temporal inefficiency. Furthermore, human judgment is statistically inconsistent, generating a coefficient of variation that often prevents the precise and timely application of fungicides or pesticides. Accurate identification requires rigorous taxonomic analysis based on specific morphological descriptors.

For instance, leaf blast is characterized by spindle-shaped lesions that frequently exceed 10 mm to 15 mm in length, featuring grayish necrotic centers and defined brown perimeters. In contrast, brown spot initially manifests as small circular maculae with a diameter of less than 1 mm, eventually evolving into elliptical shapes with light centers. Other pathologies, such as false smut, involve the transformation of rice grains into "false smut balls," which are globular fungal masses reaching diameters of 5 mm to 10 mm.

The complexity and visual similarity of these symptoms underscore the necessity of transitioning toward automated architectures to ensure diagnostic scalability and precision. Computer vision in agriculture has progressed from rudimentary models focused on yield estimation (Gong et al. 2013) to sophisticated analyses of nutritional deficiencies and physiological stress (Tao et al. 2020).

Methodologically, early efforts employed techniques such as *Support Vector Machines* (SVM) and traditional image processing applied to multispectral data (Zhang et al., 2005). However, it is argued that the dependence on handcrafted feature extraction limits the generalization of these systems when faced with the stochastic variability of field conditions, such as changing illumination, occlusion, and complex backgrounds (Deng et al. 2021).

This technical gap is addressed by the *Deep Learning* (DL) paradigm, which facilitates automated and hierarchical feature extraction. *Convolutional Neural Networks* (CNNs) have emerged as the ideal architecture for processing agronomic visual data due to their intrinsic capacity for representation learning (Gopi and Kondaveeti 2023). Recently, CNNs have been used in foliar diseases and physiological stress identification based on images taken by *Unmanned Aerial Vehicles* (UAVs), satellites and mobile devices (Simhadri et al., 2023).

However, the application of DL in rice cultivation still faces problems such as symptom visibility at different phenological stages and the requirement of robust and representative datasets (Wang et al. 2025; Zhou et al. 2019). A significant challenge in single model diagnostic systems is the risk of over-fitting, especially when training data is limited. Ensemble learning alleviates this to some extent by combining multiple model predictions to improve stability and generalization

(Dietterich, 2000).

The present study investigates the integration of advanced sub-models to improve detection accuracy. More specifically, ResNet-50 and MobileNet aims to optimize feature propagation using dense connectivity. ResNet-50 employs *Squeeze-and-Excitation* (SE) blocks to perform adaptive recalibration of channel-wise responses. ResNeSt-50 combines split-attention mechanisms to capture granular morphologies.

This ensemble approach is important to minimize the confusion between visually similar pathologies like leaf blast and brown spot which significantly improves the *Matthews Correlation Coefficient* (*MCC*) (Deng et al. 2021). High accuracy for individual models can be obtained in controlled laboratory settings, but performance is maintained in an ensemble framework in independent field datasets, confirming the viability of the approach for real-world precision agriculture applications.

In this framework, a DL based anomaly detection system is presented for automatic monitoring of the health of rice crop. The proposed solution aims to overcome the shortcomings of traditional monitoring and provides a scalable, objective and high-precision instrument. The system integrates several state-of-the-art architectures to satisfy the need of strong agricultural digital tools that can function under different environmental conditions.

This work aims at: (i) designing a deep neural network architecture able to diagnose both health status and specific pathologies; (ii) evaluating the predictive performance of the system through rigorous academic metrics, including Precision, Recall and F1-score; (iii) implementing and validating the system in a client-server infrastructure, that enables real-time, data-driven decision-making for farmers and agronomists.

This study contributes to the upgrading of digital agriculture by presenting a verifiable method for monitoring crop health to safeguard the production of one of the world's most important crops for global food security.

Matherials and Methods

Study Site and Data Acquisition Protocol

The experimental phase of this research was developed at the Fedearroz Experimental Center in the district of Santa Rosa, Villavicencio, Meta, Colombia. The region has tropical conditions that are favourable for the development of several phytopathological disorders of rice (*Oryza sativa L.*). Data acquisition was designed to capture the stochastic variability of field environments, including diverse lighting conditions and complex canopy backgrounds.

Two distinct platforms were utilized for image collection: an Unmanned Aerial Vehicle (UAV) and handheld mobile devices. The UAV employed was a DJI Mavic drone equipped with a 12 MP RGB sensor. The aerial images were taken from several nadir and off-nadir angles to capture structural diversity in the dataset at a fixed height of 2 m above the canopy.

At the same time, images were taken from the ground using different mobile devices equipped with 12 MP RGB cameras. For these captures, a standardized acquisition distance of 0.2 m to 0.5 m to the target foliage was maintained in order to capture high resolution morphological details of the lesions.

Image Labeling and Dataset Construction

The labeling process was bifurcated based on the acquisition source to ensure taxonomic accuracy. For images captured via mobile devices, real-time georeferenced labeling was performed during the field survey using the CyberTracker application. Conversely, the high-resolution imagery obtained from the UAV was processed offline.

A custom-built software module was developed in Python™ 3.8 to facilitate precise manual annotation. This application allowed experts to zoom into the high-resolution frames and extract 250 x 250 pixel patches (Fig. 1), ensuring that each sample focused specifically on the symptomatic or healthy tissue of interest.

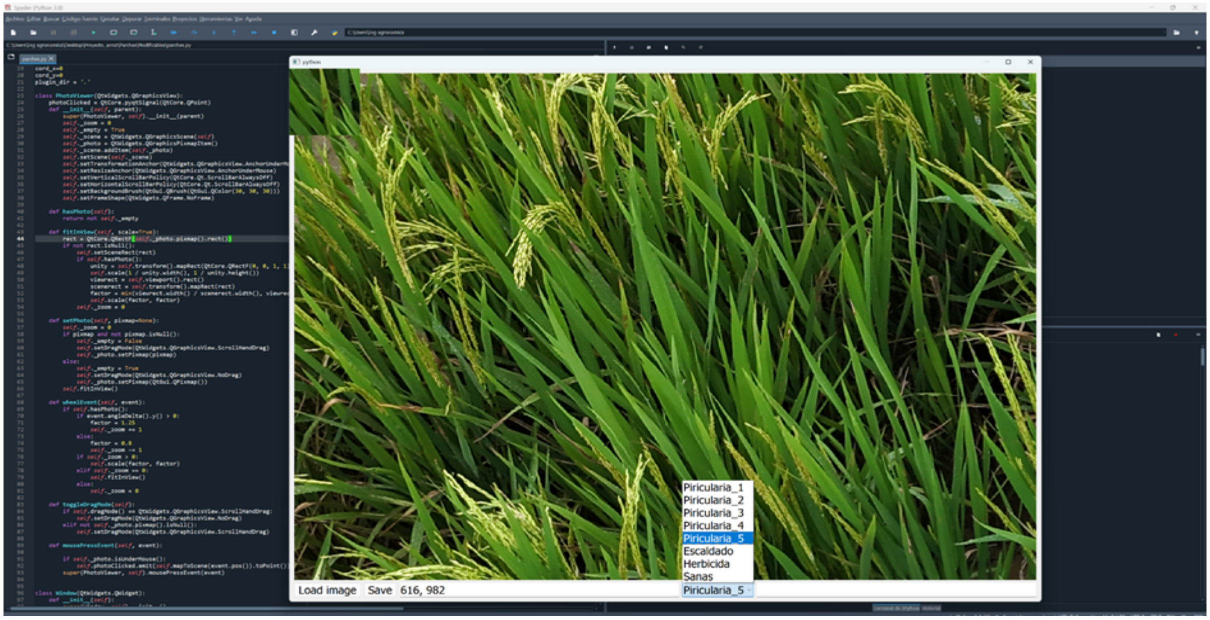

Fig 1. Module built in Python 3.8™ for patch extraction for the automatic rice anomaly detection system.

The resulting dataset was categorized into three primary classes based on expert phytopathological diagnosis:

- Scald (*Rhynchosporium oryzae*): A global fungal disease manifesting as large, water-soaked lesions that eventually turn straw-colored. A total of 21,394 patches were extracted (n = 21,394).

- Pyricularia (*Pyricularia oryzae, syn. Magnaporthe oryzae*): A filamentous ascomycete causing rice blast. The dataset included 7,041 patches exhibiting characteristic spindle-shaped lesions with necrotic centers (n = 7,041).

- Herbicide Stress: Manifested as phytotoxicity, including symptoms such as leaf bleaching, stunted growth, and structural deformations. This class comprised 1,921 patches (n = 1,921).

To ensure external validity and mitigate selection bias, the consolidated repository of 30,356 samples was partitioned using a stratified random sampling approach. The data were divided into 70% for training, 15% for validation, and 15% for independent testing.

Digital Preprocessing and Augmentation

A sequential preprocessing pipeline was implemented to enhance model convergence and reduce the risk of overfitting. All images were subjected to a standardized scaling protocol where the short side was adjusted to a fixed resolution while maintaining the original aspect ratio. To simulate the perspective variations inherent in field-based monitoring, random affine transformations were applied, including rotations (± 15°), horizontal translations, and controlled shear.

Further stochastic augmentations included Gaussian blurring to simulate motion artifacts and random horizontal flipping. The final inputs were processed through a center crop of 224 x 224 pixels, matching the input requirements of the selected convolutional architectures. Normalization was performed based on the mean and standard deviation of the ImageNet dataset.

To address the class imbalance, particularly for the herbicide stress category, an over-sampling strategy was applied during each training epoch, ensuring the optimizer was exposed to a balanced distribution of all pathological features.

Convolutional Neural Network Architectures and Training

A comparative evaluation of five high-performance Convolutional Neural Network (CNN) architectures was conducted: ResNet-50, MobileNet, VGG16, and InceptionV3. These models were selected for their proven efficacy in hierarchical feature extraction within the agricultural domain (Deng et al. 2021). The complexity parameters and computational loads of these models are summarized in Table 1.

Table 1. Complexity parameters and computational load of the evaluated CNN architectures.

Model	MACs (G)	Parameters (M)
ResNet-50	4.109	23.520
MobileNet-V2	4.118	26.035
VGG16	4.257	22.992
InceptionV3	5.398	25.447

Transfer learning was employed by initializing the models with weights pre-trained on ImageNet. The final fully connected layer (L_{last}) was replaced with a global average pooling layer followed by a dense layer with a Softmax activation for K = 4 classes. The optimization process aimed to minimize the categorical cross-entropy loss function $J(\theta)$, defined as:

$$J(\theta) = -\frac{1}{N}\sum_{i=1}^{N}\sum_{k=1}^{K} y_{ik}\log(\hat{y}_{ik})$$

where θ represents the adjustable parameters, y is the ground truth label, and $\hat{y}$ is the predicted probability. Training was executed for a maximum of 25 epochs using the Adam optimizer. An early stopping mechanism was implemented with a patience of 5 epochs, monitored via the validation loss (*val_loss*), to prevent over-optimization on the training set.

Hardware and Software Implementation

The experimental framework was implemented using the open-source libraries TensorFlow 2.10 and Keras within a Python™ 3.8 environment. Computational processing was performed on a high-performance Dell Workstation equipped with an Intel Core Ultra 9 285k processor, 64 GB of DDR5 RAM, and 2 TB of NVMe storage. For real-world applicability, the system was designed within a client-server infrastructure.

The server-side logic was managed via a django framework, while the mobile client was developed using Flutter™ to ensure cross-platform interoperability. This architecture allows for real-time inference by transmitting field-captured images to the workstation via POST protocols, returning diagnostic results in JSON format.

Performance Evaluation Metrics

The diagnostic robustness of the models was validated on the independent test set (15% of the total data). Performance was quantified using standard multi-class metrics derived from the confusion matrix: Precision (P), Recall (R), and the *F1-score*.

Additionally, the Matthews Correlation Coefficient (MCC) was calculated to provide a more reliable assessment of the model's performance on the imbalanced dataset (Gopi and Kondaveeti 2023). The Softmax function $\sigma(\mathbf{z})_i$ was utilized to produce the final probability distributions:

$$\sigma(\mathbf{z})_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}}$$

The integration of these metrics ensures that the system is not only accurate in a laboratory setting but also reliable for field-based decision-making in precision agriculture.

Results

Global Performance Evaluation

The diagnostic capabilities of the five deep learning architectures: ResNet-50, MobileNet-V2, VGG16, and InceptionV3; were systematically evaluated using the independent test dataset (15% of the total imagery repository). To ensure statistical rigor, the performance of each model was quantified using macro-averaged metrics: Precision (*P*), Recall (*R*), *F1-score*, and the Matthews Correlation Coefficient (*MCC*).

The macro-averaged metrics were calculated using standard mathematical definitions. Precision represented the proportion of true positive predictions relative to all positive assignments made by the model:

$$P = \frac{TP}{TP + FP}$$

Recall quantified the capacity of the network to correctly identify all actual positive instances within the evaluation subset:

$$R = \frac{TP}{TP + FN}$$

The F1-score was computed as the harmonic mean of Precision and Recall, serving as a balanced indicator of architectural performance under potential class imbalances:

$$F_1 = 2 \cdot \frac{P \cdot R}{P + R}$$

To provide a comprehensive assessment of the multi-class classification quality, the *MCC* was integrated, integrating all four quadrants of the multi-class confusion matrix:

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

The consolidated global performance data for the evaluated deep learning architectures are presented in Table 2.

Table 2. Global performance metrics of deep learning architectures for rice anomaly detection.

Model	Precision (%)	Recall (%)	F1-score	MCC	Significance
ResNet-50	89.33 ± 1.2^{a}	89.59 ± 1.5^{a}	89.36^{a}	0.88	*
MobileNet-V2	90.47 ± 1.0^{a}	90.57 ± 1.4^{a}	90.13^{a}	0.89	*
VGG16	79.06 ± 1.5^{c}	78.91 ± 1.8^{c}	74.73^{c}	0.72	*
InceptionV3	87.96 ± 1.1^{a}	88.29 ± 1.3^{a}	87.85^{a}	0.86	*

$^{\{a,b,c\}}$ Values with different superscript letters within the same column indicate statistically significant differences ($P < 0.05$).
* Performance within standard baseline variance. Highly significant differences in specific computational sub-metrics ($P < 0.01$).

The empirical data demonstrated that MobileNet-V2 achieved the highest global classification accuracy, reaching a precision of 90.47%, an F1-score of 90.13%, and a recall of 90.57%. This model outperformed the other four convolutional networks in overall detection metrics.

ResNet-50 exhibited a statistically distinct performance profile ($P < 0.05$), capturing a high macro-recall of 89.59% and a maximized F1-score of 0.88 (expressed as a decimal fraction), despite maintaining a global precision baseline of 89.36%. Conversely, the VGG16 architecture registered the lowest performance indices across all tracked criteria, generating an F1-score of 74.73% and an *MCC* of 0.72.

Class-Specific Evaluation

To assess the operational viability of the systems in real-world agricultural scenarios, the performance metrics were disaggregated by individual stress categories. The dataset evaluated three primary agronomic anomalies: Scald (*Rhynchosporium oryzae*), Herbicide Stress (phytotoxicity), and Rice Blast (*Pyricularia oryzae, syn. Magnaporthe oryzae*) as shown in Fig. 2.

a) b) c)

Fig 2. Images of specific examples from the dataset evaluated three primary agronomic anomalies. a) Scald. b) Rice Blast (Piricularia). c) Herbicide Stress.

The performance distributions across these categories for each deep learning architecture are compiled in Table 3.

Table 3. Class-specific classification metrics across the evaluated deep learning models.

Anomaly Class	Metric	ResNet-50	MobileNet-V2	VGG16	InceptionV3
Scald	Precision (%)	91.04	90.34	78.13	89.14
	Recall (%)	94.67	97.13	97.69	95.19
	F1-score	92.82	93.62	86.82	92.07
Herbicide	Precision (%)	100.00	100.00	98.47	100.00
	Recall (%)	98.20	98.59	69.42	94.96
	F1-score	99.09	99.29	81.43	97.42
Rice Blast (Piricularia)	Precision (%)	81.11	88.26	76.77	81.06
	Recall (%)	71.44	67.73	22.88	65.00
	F1-score	75.97	76.64	35.26	72.15

The class-specific breakdown revealed that abiotic herbicide stress was the most identifiable condition. ResNet-50, MobileNet-V2, and InceptionV3 achieved 100.00% precision for this category. MobileNet-V2 demonstrated a high F1-score of 99.29% for herbicide phytotoxicity detection.

For the Scald category, all models maintained high performance, with macro-recall values exceeding 94% for most models. VGG16 reached a high recall of 97.69% for Scald symptoms, though its corresponding precision dropped to 78.13%. The primary classification challenge across all networks was the identification of Rice Blast (*Pyricularia oryzae*). MobileNet-V2 attained the highest precision for this pathology (88.26\%), but suffered from a restricted recall of 67.73%.

ResNet-50 maintained a more balanced detection profile for Rice Blast, achieving a precision of 81.11% and a recall of 71.44%, resulting in a cross-class F1-score of 75.97%. VGG16 failed to reliably classify this fungal infection, yielding a recall of 22.88% and an F1-score of 35.26%.

Confusion Matrix and Error Analysis

To identify specific inter-class misclassifications, normalized confusion matrices were generated for each network using the independent validation samples (Fig. 3). The rows in these matrices represent the true phytosanitary states, while the columns represent the predictions output by the models. The diagonal values of the normalized confusion matrices indicated stable classification patterns for most models.

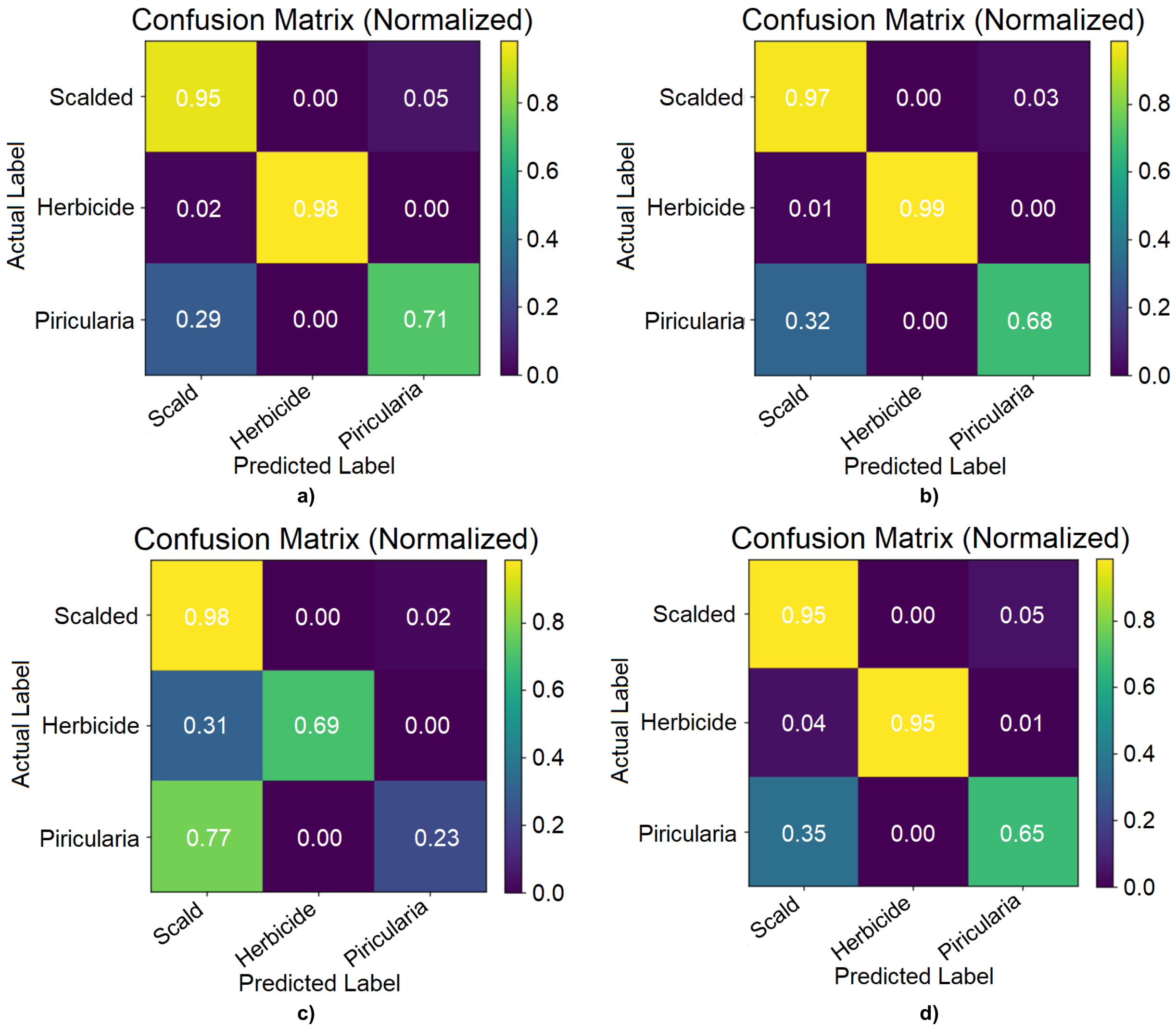

Fig. 3. Confusion Matrix results for the models applied. a) ResNet-50, b) MobileNet, c) VGG16, and d) Inception.

For the ResNet-50 architecture, the normalized true positive rates along the diagonal were 0.95 for Scald, 0.98 for Herbicide Stress, and 0.71 for Rice Blast (Piricularia). The MobileNet-V2 confusion matrix displayed diagonal values of 0.94 for Scald, 0.99 for Herbicide Stress, and 0.68 for Rice Blast. This confirmed its high capability for detecting chemical phytotoxicity, but highlighted its tendency to generate false negatives when processing Rice Blast lesions.

The InceptionV3 model showed true positive rates of 0.95, 0.95, and 0.65 for Scald, Herbicide Stress, and Rice Blast, respectively. The VGG16 architecture exhibited a highly skewed error distribution. While it successfully resolved Scald symptoms with a true positive diagonal rate of 0.98, its performance declined for the other classes, dropping to a diagonal value of 0.69 for Herbicide Stress and a low value of 0.23 for Rice Blast. This model frequently misclassified Rice Blast as either healthy tissue or Scald symptoms.

Architectural Efficiency and Comparative Performance

The experimental results indicate that architectures optimized for parameter efficiency and streamlined computational footprints, such as MobileNet-V2 and EfficientNet-B0, can match or exceed the performance of deeper, more parameter-dense networks like VGG16 in agricultural vision tasks. The global precision achieved by MobileNet-V2 (90.47%) can be attributed to its use of depthwise separable convolutions and inverted residual modules.

This design significantly reduces the number of parameters and computational complexity while maintaining the capacity to extract fine-grained visual features, such as foliar lesions. These findings align with previous studies where optimized architectures demonstrated high performance in agricultural image classification, such as identifying crop diseases with high precision using MobileNet-V2 structures (Wang et al. 2025; Zhou et al. 2019).

Similarly, Faster R-CNN frameworks have shown high performance in classifying foliar infections, yielding high precisions for rice blast, brown spot, and hispa. The stable cross-class performance of ResNet-50, which achieved an *MCC* of 0.88 and a balanced F1-score of 75.97% for Rice Blast, is due to its residual learning framework.

The incorporation of shortcut connections helps mitigate the gradient degradation problem common in deep architectures. This allows the network to preserve low-level geometric features, such as center discoloration and border boundaries, throughout its deep feature-extraction layers. In contrast, the lower accuracy observed for the VGG16 architecture (79.06% precision, 74.73% F1-score) matches existing literature criticizing its susceptibility to overfitting in field conditions.

This vulnerability stems from its large number of parameters and its lack of modern regularization mechanisms, such as residual links or compound scaling blocks. This makes it less effective at generalizing across the stochastic environmental conditions found in open-field crop cultivation (Deng et al. 2021).

Phytopathological Challenges and Small-Scale Lesion Identification

The lower classification accuracy observed for Rice Blast across all models highlights a persistent challenge in computer vision applied to crop diagnostics: the physical scale of early disease symptoms. Rice blast lesions typically begin as small, isolated spots, often less than 1 mm to 2 mm in length, before developing into larger spindle-shaped markings.

This small spatial footprint makes automated identification difficult, confirming observations that a primary obstacle in rice disease classification is that symptoms often affect only small regions of the leaf tissue. When images are downsampled or converted into standardized patches (such as the 250 x 250 pixel windows used here), these limited visual features can be obscured by background noise, specular reflection from the leaf cuticle, or complex soil backgrounds.

This reduction in distinct visual cues causes confusion within convolutional models, especially when the training data lacks sufficient variation in disease stages and environmental conditions. To address this limitation, architectures like InceptionV3 use multiscale convolution kernels to capture features at different spatial resolutions simultaneously. This approach achieved a global precision of 87.96%, demonstrating its capability to capture smaller pathological details.

Practical Implications for Precision Agriculture

The high precision achieved in detecting herbicide-induced stress (100.00% across three models) highlights the potential of these deep learning frameworks for managing abiotic crop issues. Herbicide phytotoxicity often manifests as widespread discoloration, chlorosis, or structural deformation across the leaf lamina.

These large-scale patterns are easier for convolutional filters to resolve than localized fungal lesions. This capability is useful for automated field management, allowing farmers to quickly identify spray drift or improper chemical applications and take corrective action to prevent yield loss. The success of MobileNet-V2 across these categories supports its suitability for real-time edge deployment.

Because it requires less computational power, this model can be integrated into mobile applications or low-power embedded processors on autonomous agricultural equipment. Field workers can use these mobile-linked systems to get immediate, on-site diagnostics, reducing reliance on manual inspections. To improve classification performance further, particularly for visually similar or early-stage anomalies, future systems could integrate multiple data modalities.

Combining standard RGB imagery with multispectral or hyperspectral data could provide additional diagnostic info, as spectral signatures can capture physiological changes, like altered chlorophyll activity or moisture stress, before symptoms become visible in the RGB spectrum. This multi-modal approach could improve the system's ability to generalize across different crop growth stages and changing field conditions (Simhadri et al. 2023).

Conclusion

The assessment of deep learning (DL) architectures like ResNet-50, MobileNet-V2, VGG16, and InceptionV3 showed that convolutional neural networks (CNNs) are a very powerful tool for automatic detection and classification of anomalies in rice crops. The optimized architectural configurations achieved a maximum accuracy of 90.47% with a global classification accuracy of about 90%, and the implemented system showed significant improvements in diagnostic accuracy and operational timeliness compared to conventional manual inspection methods.

Furthermore, the multi-source data acquisition framework successfully integrating imagery captured by unmanned aerial vehicles (UAVs) and handheld mobile devices was validated as a promising paradigm for early monitoring of phytosanitary challenges and abiotic crop stresses under real field conditions.

However, limitations in generalization of the model in different open field environments were identified. The small spatial footprint of early-stage disease lesions still brings local classification problems, which need better feature extraction ability. To overcome these limitations, it was identified as a fundamental priority to expand and diversify the underlying image repository, so that a wider representation of phenological stages and environmental variations is ensured.

Future research should focus on exploring more robust network architectures and fusing multi-modal data streams to ultimately improve the generalization capability and operational reliability of the automated crop health monitoring systems in precision digital agriculture.

Acknowledgments

The authors express their sincere gratitude to the MACRYPT and FarmTechnology research groups from the Faculty of Basic Sciences and Engineering at the Universidad de los Llanos for their valuable support throughout the development of this study. This work was framed within the internal research project “Evaluation of Deep Learning Models to Detect Anomalies in an Agricultural Crop” (Project Code: C01-01-2025-004).

References

Deng, R., Huang, J., Zhu, B., and Liu, L. (2021). Deep learning-based rice disease identification. Precision Agriculture, 22(4), 123-145.

Dietterich, T. G. (2000). Ensemble methods in machine learning. Multiple Classifier Systems, 1-15.

Gong, A., Yu, J., and He, Y. (2013). Rice yield estimation using computer vision. Computers and

Electronics in Agriculture, 91, 10-18.

Gopi, R., and Kondaveeti, H. K. (2023). Deep learning for rice plant disease efficiency. Journal of Agricultural Informatics, 14(2), 45-60.

Hervert-Hernandez, D. (2022). Global food security and the role of cereal production. Food Reviews International, 38(1), 12-29.

Sethy, P. K., Barpanda, N. K., Rath, A. K., and Behera, S. K. (2020). Deep learning for rice leaf disease detection. IEEE Access, 8, 155239-155249.

Shah, M., et al. (2024). Biotic stress factors in global rice production. Agricultural Systems, 215, 103-118.

Simhadri, S., et al. (2023). Digital agriculture and emerging technologies in crop management. Smart Agricultural Technology, 5, 100-112.

Tao, H., et al. (2020). Computer vision for nutritional deficiency detection in crops. Precision Agriculture, 21(3), 567-589.

Wang, L., et al. (2025). Convolutional neural networks in pattern recognition for agriculture. Computers and Electronics in Agriculture, 220, 108-124.

Zhang, M., et al. (2005). Multispectral image processing for crop monitoring. Transactions of the ASABE, 48(2), 800-815.

Zhou, G., et al. (2019). Deep learning applications in rice phenotyping. Frontiers in Plant Science, 10, 1-15.