UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS

Tecnura

https://revistas.udistrital.edu.co/index.php/Tecnura/issue/view/1326

DOI: https://doi.org/10.14483/22487638.23913

INVESTIGACIÓN

Image-Based Deep Learning Tool for Predicting Tomato Leaf Diseases

Herramienta para la predicción de enfermedades en cultivos de tomate mediante tecnologías de aprendizaje automático

José-Neftalí Torres-Marín [1], Edward-J. Marín-García [2] y Andrés Felipe Calvo Salcedo [3]

Fecha de Recepción: 15 de julio de 2025 **Fecha de Aceptación**: 30 de marzo de 2026

Cómo citar: J.N. Torres-Marín, E. J. Marín-García, y A.F. Calvo Salcedo, «Image-Based Deep Learning Tool For Predicting», *Tecnura*, vol. 30, n.° 87, mar. 2026. 3–24. https://doi.org/10.14483/22487638.23913

Abstract

Objective: Late blight disease, caused by Phytophthora infestans, represents a critical threat to tomato cultivation, significantly affecting its quality and productivity. This study aimed to develop a technological tool for the early and precise detection of tomato disease in a controlled agricultural environment using machine learning (ML) techniques and digital image processing..
Methodology: A total of 3560 foliar images encompassing three severity levels of *Phytophthora infestans* were acquired. The Segment Anything Model (SAM) was implemented as a preprocessing step to isolate the biological tissue and eliminate chromatic background noise. Three classifiers were evaluated using a unified Monte Carlo cross-validation framework (10 iterations, 70/30 split): SVM and KNN, based on global chromatic descriptors (HSV/YIQ and PCA), compared with a CNN that extracts local spatial topologies.
Results: Within the evaluated dataset, the convolutional neural network (CNN) and the local distance-based model (KNN) achieved classification metrics exceeding 98%, significantly outperforming the global color-based baseline (SVM, 79%). By evaluating the localized spatial morphology of the necrotic lesions, the CNN architecture effectively mitigated misclassifications in extreme severity phenotype. Consequently, these high classification metrics must be interpreted as an upper performance bound, strictly constrained to the ideal and controlled conditions of the greenhouse setup.
Conclusions: The developed tool demonstrates high potential for automated disease diagnosis in tomatoes, contributing to phytosanitary management through emerging technologies. This study highlights ML's role as a driver of innovation for more precise, efficient, and sustainable agriculture.
Keywords: Precision agriculture; Machine learning; Phytophthora infestans; Late blight.

Resumen

Objetivo: El tizón tardío, causado por Phytophthora infestans, representa una amenaza crítica para el cultivo de tomate, afectando su calidad y productividad. Ante esta problemática, el presente estudio tiene como objetivo desarrollar una herramienta

1 Universidad del Valle. Grupo de investigación GiiDEA. Cartago – Colombia. Email neftali.torres@correounivalle.edu.co
2 Universidad del Valle. Grupo de investigación GiiDEA. Cartago – Colombia. Email marin.edward@correounivalle.edu.co
3 Universidad Tecnológica de Pereira. Grupo de investigación GIE. Pereira - Colombia. Email afcalvo@utp.edu.co

tecnológica para la detección temprana y precisa de la enfermedad en hojas de tomate, utilizando técnicas de aprendizaje automático y procesamiento digital de imágenes en un entorno agrícola controlado.
Metodología: Se adquirieron 3560 imágenes foliares en tres niveles de severidad de *Phytophthora infestans*. El modelo SAM se implementó como preprocesamiento para aislar el tejido biológico y eliminar el ruido cromático del fondo. Mediante una Validación Cruzada de Monte Carlo unificada (10 iteraciones, 70/30), se evaluaron tres clasificadores: SVM y KNN, basados en descriptores cromáticos globales (HSV/YIQ y PCA), frente a una CNN que extrae topologías espaciales locales.
Resultados: Los resultados demuestran que la red neuronal (CNN) y el modelo local (KNN) acertaron en más del 98% de las clasificaciones, superando ampliamente al modelo de color (SVM, con 79%). La CNN logró eliminar las confusiones al identificar hojas gravemente enfermas porque evalúa la forma de las manchas. Estos altos aciertos representan un rendimiento máximo esperable, válido estrictamente bajo las condiciones ideales y controladas del invernadero.
Conclusiones: La herramienta desarrollada demuestra un alto potencial para el diagnóstico automatizado de enfermedades en tomate, aportando al manejo fitosanitario con base en tecnologías emergentes. Se destaca el papel del aprendizaje automático como motor de innovación para una agricultura más precisa, eficiente y sostenible.
Palabras clave: Agricultura de precisión; Aprendizaje automático; Phytophthora infestans; Tizón tardío

Introduction

Tomato (Lycopersicum esculentum Mill.) is a widely cultivated crop native to the Andean region. It gained popularity in Mexico during the 16th century before spreading across Europe, the Middle East, Africa, and the United States [1], [2], [3]. The tomato belongs to the Solanaceae family and grows as a shrub with alternatively arranged compound leaves along the stem.

For optimal growth, the tomato plant requires daytime temperatures between 20°C and 30°C, nighttime temperatures between 1°C and 17°C, and relative humidity levels ranging from 60% to 80%. It also needs ample light and thrives slightly acidic to alkaline soils with good drainage [4]. As a dicotyledonous, herbaceous perennial, tomato is cultivated annually for its fruit. The plant has a deep, moderately branched root system and a 2 to 4 cm thick herbaceous stem. Its pinnate compound leaves feature petiolated, lobed, and toothed leaflets, while its flowers form in cymose clusters containing 4-12 blooms. The tomato fruit, botanically classified as a berry, typically weighs approximately 600 g [4].

The plant's growth and production cycle consists of three phenological phases: initial, vegetative, and reproductive, spanning approximately 120 days [5].

However, the tomato plant is vulnerable to a range of biotic and abiotic diseases that can severely impact its yield and quality. Biotic threats include spider mites, whiteflies, aphids, thrips, leaf miners, and caterpillars. The plant is also susceptible to fungal diseases, such as gray rot, alternariosis, powdery mildew, anthracnose, and late blight. The latter, caused by Phytophthora infestans, is particularly damaging [6].

Late blight manifests as mottled, embossed brown spots on the fruit surface, along with lesions on stems and leaves that eventually turn brown and dry. Infection severity was measured as the percentage

of the affected area, ranging from 0% to 100%. The effective management of late blight involves preventing prolonged high-humidity conditions and applying appropriate treatments to control symptoms [6].

Research on disease detection in tomato crops has gained significant attention, as evidenced by the growing number of literature reviews. In Guatemala, pathogens responsible for purple tip disease in tomatoes and potatoes were investigated, and a high incidence of phytoplasmas BLTVA and *Candidatus Liberibacter solanacearum* was identified [7]. Their study highlights the critical role of molecular characterization in accurate disease diagnosis. Similarly, [8] reviewed five key attributes of *Phytophthora infestans*, a pathogen linked to late blight disease in both potato and tomato crops. Their work traces historical development, worldwide distribution, and disease progression, while also identifying existing knowledge gaps.

A previous study [9] focused on image processing and ML approaches and analyzed traditional disease detection techniques applied to various crops. Their evaluation confirmed the high accuracy of these models, yet they noted a significant challenge: many producers lack the expertise to interpret the model outputs and understand the underlying characteristics. A related study, [10] conducted a bibliographic review of diseases caused by fungi, viruses, and bacteria affecting crops in an open field south of Tamaulipas, Mexico, contributing further insight into regional phytopathology.

This work presents the main diseases affecting tomato crops, emphasizing symptoms, key morphological characteristics, economic impacts, and effective disease management strategies. Researchers from Universidad Francisco de Paula Santander conducted a systematic review of digital image processing techniques applied to pest and disease detection in Colombian agricultural crop. Framed within the Agriculture 4.0 paradigm, their study highlights emerging technologies, such as neural networks and deep learning to reduce economic losses caused by pests and diseases. Key contributions include identifying advanced image processing methods and discussing current challenges in their implementation, thereby providing a solid theoretical foundation for future research [11].

Previous studies have deepened the understanding of pathogen dynamics, particularly for late blight, while underscoring the importance of integrating interdisciplinary approaches—including molecular biology, image processing, and machine learning—to improve disease detection, prevention, and management. A convolutional neural network (CNN)-based system was designed [12] and implementedon low-cost devices such as the Raspberry Pi 4, offering an accessible and efficient solution for farmers in Mexico. This system utilized the PlantVillage dataset, comprising 18 160 RGB images, with 1910 depicting late blight symptoms, to train CNN models for tomato leaf disease classification. Transfer learning was applied using four architectures—MobileNetV2, NasNetMobile, Xception, and MobileNetV3—yielding accuracies of 97%, 100%, 99%, and 97%, respectively. Validation of 100 images from 12 diseased tomato crops under real-world conditions demonstrated visually consistent results with simulated data.

Similarly, conditional generative adversarial networks (C-GANs) have been employed in India to advance disease detection by generating synthetic images of tomato leaves [13]. The DenseNet121 model trained on both real and synthetic data obtained from the PlantVillage database achieved a precision of >97%.

Meanwhile, a tomato disease detection system combining image processing techniques with CNNs inspired by architectures such as AlexNet and YOLO was developed in Nepal [14]. The main contributions of these studies include the use of digital images to identify common tomato crop diseases. For example, one system used 121 images of late blight symptoms (100 for training and 21 for testing) and achieved a prediction accuracy of 98%. This system was implemented on low-cost devices, such as the Raspberry Pi, enhancing its accessibility for practical use in the agricultural sector.

In China, [15] developed a deep learning approach for identifying tomato leaf diseases, specifically late blight. Using CNNs and transfer learning, they evaluated architectures, including AlexNet, GoogLeNet, and ResNet, which were optimized via fine-tuning. The best ResNet model combined with SGD optimization achieved an impressive accuracy of 97.28%. Their dataset included 726 images of late blight, with 58 images for training and 145 for testing.

Similarly, researchers at the Vellore Institute of Technology in India proposed a system to detect diseases in tomato leaves, including late blight, based on residual convolutional neural networks (residual CNNs) enhanced with attention mechanisms. This method integrates residual connections with an attention mechanism that highlights important features across multiple network layers. The system achieved 98% accuracy in disease identification using an expanded PlantVillage dataset (22,000 images, with 17,600 and 4400 images for training and 4400 images for testing) [16].

A disease detection system based on CNNs with three convolutional layers, three max-pooling layers, and two fully connected layers was also presented in India [17] This model was optimized through hyperparameter tuning and trained on an augmented PlantVillage dataset balanced across classes (1,000 images for training and 750 for testing/validation). The model attained an average precision of 91.2% and was compared with pretrained models, highlighting its suitability for mobile applications due to its low memory and computational demands.

Gunadarma University in Indonesia [18] implemented a CNN-based model to classify five tomato leaf disease categories, including late blight. Their work automates the diagnostic process by eliminating the need for manual feature extraction and segmentation, making it practical for farmers. The accuracy of the model was 96.60% with 800 training images and 200 images for testing and validation.

[19] conducted a study at the Universidad Tecnológica de Pereira, Colombia, to develop a methodology for detecting late blight disease on tomato leaves using digital image processing. Their approach

employed YIQ and TSL color models to effectively segment and analyze diseased areas with high accuracy. The study utilized the PlantVillage database, consisting of 1909 images, which were split into training, testing, and validation subsets. Specifically, 1527 images (80%) were used for training, while the remaining 382 images (20%) were divided between testing and validation. The proposed method achieved detection accuracies of 98.60% for healthy pixels and 98.17% for diseased pixels, demonstrating robust performance. Additionally, the study compared its results with state-of-the-art techniques, highlighting superior precision and reduced processing times, underscoring the efficacy and efficiency of the proposed methodology.

Collectively, these studies represent significant advances in image processing and deep learning techniques for detecting tomato plant diseases, particularly late blight. The main characteristics and results of these studies are summarized in Table 1.

Table 1. *Late blight detection in tomato plants*

Authors	Algorithm used	Precision
(González et al., 2021)	CNN (MobileNetV2, NasNetMobile, Xception, and MobileNetV3)	97%, 100%, 99%, 97%
(Abbas et al., 2021)	C-GAN with DenseNet121	>97%
(Adhikari et al., 2018)	CNN (AlexNet and YOLO)	98%
(Zhang et al., 2018)	CNN (AlexNet, GoogLeNet, and ResNet with SGD)	97.28%
(Karthik et al., 2020)	Residual CNN with an Attention mechanism	98%
(Agarwal et al., 2020)	CNN (3 convolutional layers and max pooling)	91.2%
(Widiyanto et al., 2020)	CNN	96.60%
(Pamplona et al., 2020)	Color models (YIQ and TSL)	98.60% (healthy), 98.17% (diseased)

Source: Authors

The gap between algorithmic performance in controlled settings and operational deployment in agricultural environments remains a significant challenge. In situ images exhibit high variability due to occlusion, fluctuating illumination, and co-occurring stressors, which often degrade model generalization. This study establishes a highly controlled baseline to evaluate the intrinsic morphological and chromatic separability of Phytophthora infestans severity stages rather than directly addressing open-field variability. This study presents a supervised image-based classification framework that evaluates deep and classical ML approaches under strictly uniform conditions.

The main contributions of this work are as follows:

- A structured dataset of 3560 tomato leaf images acquired under uniform greenhouse illumination, labeled across three severity levels (healthy: 2,089; low severity: 960; extreme severity: 511).

- The integration of the Segment Anything Model (SAM) exclusively as an automated preprocessing step to isolate the foliar region of interest and eliminate nonbotanical background noise.

- A comparative evaluation of a custom Convolutional Neural Network (CNN) against classical baselines (SVM and KNN), executed under a unified and traceable Monte Carlo cross-validation protocol (10 iterations, 70/30 stratified split).

Observational evidence demonstrates that using purely chromatic descriptors, early-stage necrotic lesions are not globally linearly separable, justifying the need for spatial feature extraction (CNN) or local density estimations (KNN).

Methodology

The development of an image-based classification framework for categorizing Phytophthora infestans severity in tomato leaves requires a structured methodological design. This approach comprises a series of sequential stages (Figure 1) for evaluating the feature extraction and classification models under strictly controlled experimental conditions.

Data Acquisition and Environmental Setup

To construct the image dataset, tomato plants were cultivated in a controlled greenhouse facility. The phytopathological progression of Phytophthora infestans was systematically induced by regulating the internal microclimate, maintaining temperatures between 22°C and 26°C and a relative humidity exceeding 85%, simulating optimal conditions for pathogen emergence [20].

To capture the localized structural variance of the necrotic lesions, high-resolution foliar imaging was performed using a 50-megapixel primary sensor (f/1.8 aperture) operating in macrofocus mode. All photographs were acquired under uniform, diffuse illumination to ensure spatial and chromatic consistency across the dataset, strictly minimizing shadows or reflections that could introduce artifacts during algorithmic feature extraction.

To standardize the dataset before algorithmic segmentation, all acquired images were initially resized to a resolution of 1200 × 1600 pixels. The dataset was categorized into three severity classes reflecting the biological progression of the disease, as illustrated in Figure 2: “Healthy” leaves with no visible lesions; “Low severity” leaves with incipient necrotic spots; and “Extreme severity” leaves with extensive foliar necrosis. The final evaluated dataset comprises 3560 images, with a natural class distribution of Healthy (n = 2,089), Low severity (n = 960), and Extreme severity (n = 511).

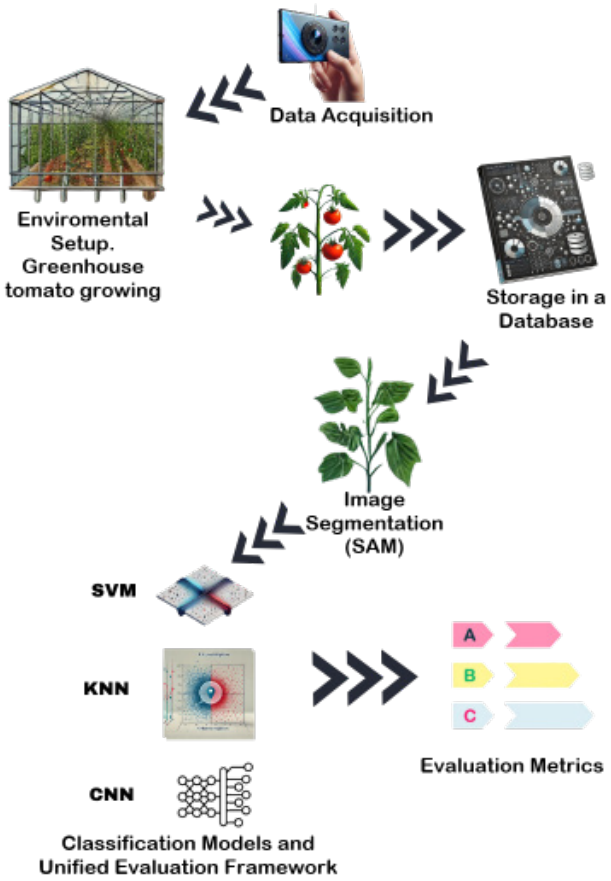

Figure 1. Methodological workflow for the image-based classification of severity of Phytophthora infestans.

Source: Authors

Figure 2. Leaf categories: healthy, low severity, and extreme severity.

Source: Authors

Synthetic resampling techniques were applied to prevent the majority-class bias due to the inherent class imbalance during model training. These balancing procedures were executed exclusively on the training partitions of each iteration (70%) during the cross-validation protocol to ensure methodological traceability and prevent data leakage. The resampling strategies were adapted to specific algorithmic architectures: classical models (SVM and KNN) utilized the Synthetic Minority Over-sampling Technique (SMOTE) within the reduced feature space, whereas the deep learning approach (CNN) employed stochastic spatial augmentation (affine rotations and Gaussian noise) directly on the image tensors. The specific augmentation parameters are detailed in the others section.

Image segmentation

This stage is structural preprocessing within the classification pipeline. Recent developments in computer vision have introduced foundation models for generalized image segmentation, notably the Segment Anything Model (SAM) [21]. SAM is based on a vision transformer (ViT-Huge) architecture and operates through a prompt-driven mechanism that isolates specific objects without requiring task-specific retraining. In this study, SAM functions exclusively as an automated tool to separate biological tissue from nonbotanical background elements.

As illustrated in Figure 3, the segmentation procedure was executed by applying color space transformations (BGR to RGB) and utilizing the spatial center coordinates of each image as the primary input prompt for the SAM predictor. The algorithm evaluated multiple generated masks based on their prediction scores and selected the most geometrically accurate representation of the leaf. The ROI was then isolated by assigning a scalar value of zero (pure black) to all non-mask pixels. This procedure strictly aims to reduce the chromatic noise introduced by the capture environment and standardizes the spatial representation of the data, avoiding any intervention in the clinical labeling of the disease severity.

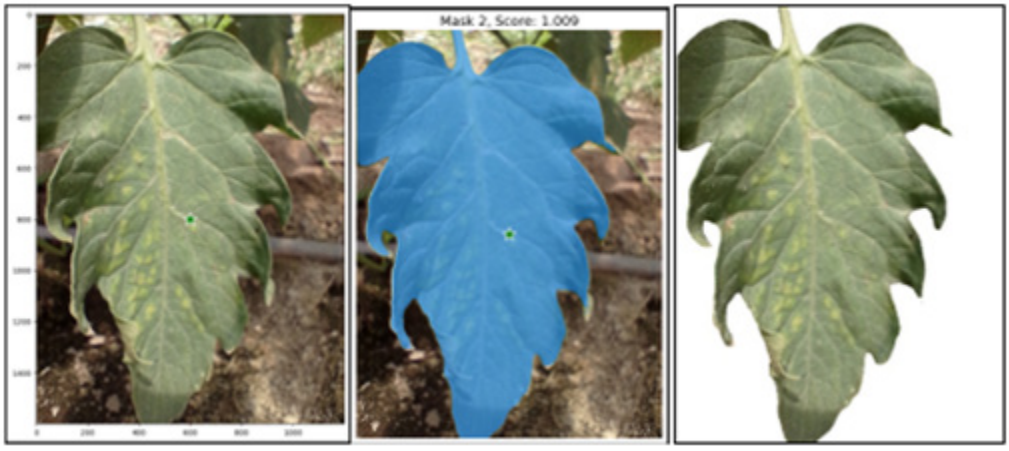

Figure 3. Automated isolation of the foliar region of interest using the Segment Anything Model (SAM).

Source: Authors

Classification Models and the Unified Evaluation Framework

A comparative analysis was conducted between classical machine learning algorithms (SVM and KNN) and a deep learning approach (CNN) to categorize the severity of P. infestans. All models used an identical base dataset of 3560 SAM-segmented images. A unified evaluation framework was established for all architectures to guarantee a fair methodological comparison and prevent data leakage: a Monte Carlo cross-validation protocol with 10 independent iterations, utilizing a strictly stratified 70/30 partition for training and testing, respectively.

While the validation protocol and the base image inputs remained constant, the feature extraction and synthetic balancing strategies were systematically adapted to each architectural paradigm's requirements. The classical models (SVM and KNN) relied on the manual extraction of global chromatic descriptors, subsequent dimensionality reduction via Principal Component Analysis (PCA), and the application of the Synthetic Minority Over-sampling Technique (SMOTE) exclusively within the numerical training feature space. Conversely, the deep learning model (CNN) operated directly on the standardized image tensors, exclusively during the training phase, using stochastic spatial data augmentation (rotations and Gaussian noise) to address class imbalance. The following subsections detail the specific architectural configurations, hyperparameters, and tensor transformations of each classifier.

- *Support Vector Machines (SVMs):* Grounded in statistical learning theory for margin maximization [22], the feature engineering process for the SVM classifier (Figure 4) relied exclusively on global chromatic descriptors, strictly discarding the null background pixels generated by the prior SAM segmentation. The original RGB images were transformed into perceptual and luminance-chrominance color spaces to mitigate illumination dependencies. For each resulting channel, normalized histograms were computed to generate a concatenated initial feature vector per image.

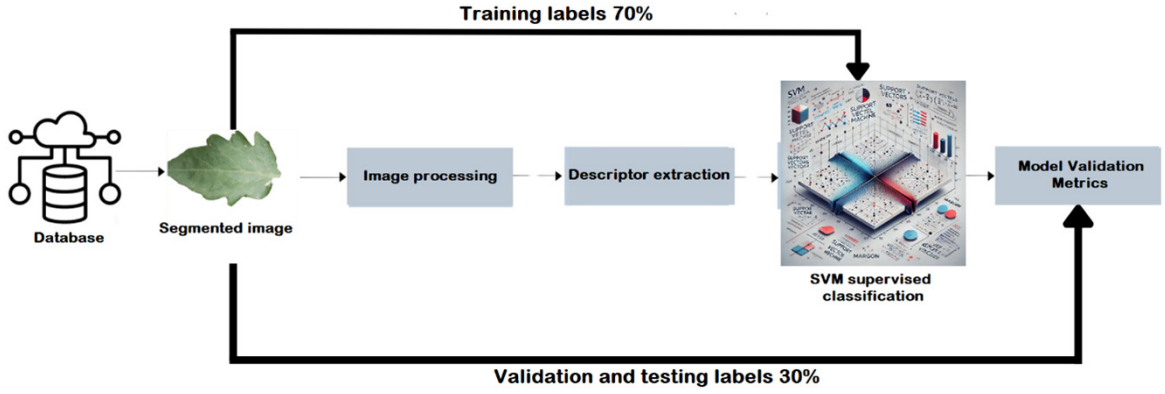

Figure 4. Algorithmic workflow for chromatic feature extraction and SVM classification.

Source: Authors

The feature vectors were standardized within the training partition of each Monte Carlo iteration (70% of the data). To reduce dimensionality and eliminate collinear noise, Principal Component Analysis (PCA) was applied. Subsequently, the synthetic minority over-sampling technique (SMOTE) was implemented to address the biological class imbalance within the training subset, generating synthetic vectors in the latent space to equalize the representation of the minority classes.

The classification step was performed using a nonlinear multiclass SVM. The independent test partitions (30%) were transformed and evaluated exclusively on their respective training folds using the scaling and PCA parameters. Table 2 presents the exact algorithmic configurations to avoid textual redundancy and ensure strict methodological reproducibility, including specific color spaces, histogram bins, PCA variance retention, SMOTE target distributions, and SVM hyperparameters.

Table 2. *Feature extraction and hyperparameter configuration of SVM baseline*

Procedure / Parameter	Configuration / Value
Color Spaces	HSV, YIQ (6 channels total)
Histogram Bins	10 per channel
Initial Dimensionality	60 features
Standardization	Z-score (StandardScaler)
Dimensionality Reduction (PCA)	95% retained variance (~21 components)
Balancing Strategy (Training only)	SMOTE (Target: 1,462 samples per class)
Kernel	Radial Basis Function (RBF)
Regularization Parameter (C)	1.0
Kernel Coefficient (Gamma)	'scale'
Multiclass Strategy	One-vs-Rest (OvR)

Source: Authors

K-Nearest Neighbors (KNN): Grounded in non-parametric local density estimation [23], the K-Nearest Neighbors (KNN) classifier (Figure 5) was implemented to evaluate the topological grouping of the phytopathological progression within the feature space. The data were standardized exclusively within the training partition (70%) of each unified Monte Carlo iteration on identically structured initial feature vectors derived from the global chromatic descriptors.

To mitigate the curse of dimensionality, which inherently degrades distance-based evaluation, Principal Component Analysis (PCA) was applied strictly to the training feature space. This orthogonal transformation ensured that the similar metrics were driven by structural variance rather than collinear noise. Subsequently, to address the biological class imbalance, the Synthetic Minority Over-sampling Technique (SMOTE) was executed within the latent PCA space, generating synthetic instances to equalize the minority classes without introducing structural artifacts.

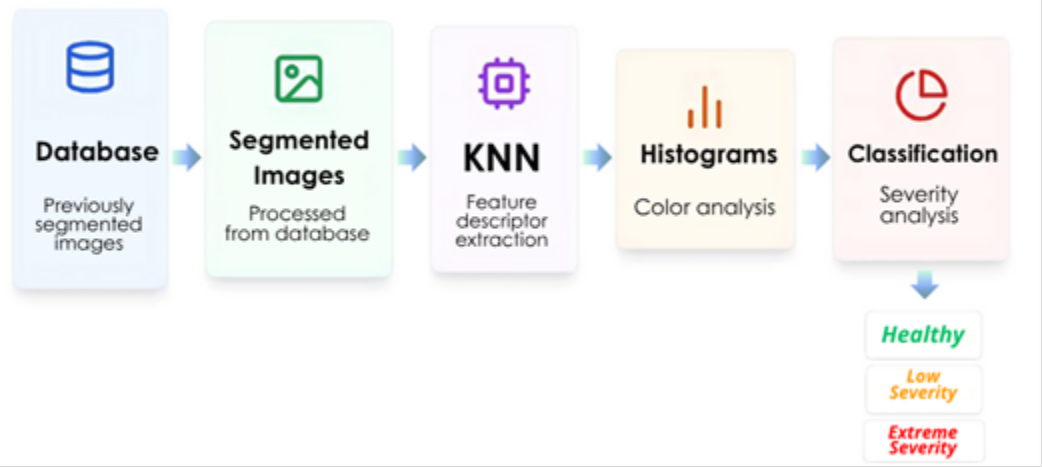

Figure 5. Algorithmic workflow for the KNN classification.

Source: Authors

The classification step was performed using an optimized local neighborhood topology based on standard spatial proximity. The independent test partitions (30%) were transformed and evaluated using the scaling and PCA parameters fitted strictly on their respective training folds. To avoid textual redundancy and maintain cross-architectural traceability, the exact configuration parameters including the specific number of neighbors (K), distance metrics, weighting mechanisms, and dimensionality retention are consolidated in Table 3.

Table 3. *Feature extraction and hyperparameter configuration of the KNN baseline*

Procedure / Parameter	Configuration / Value
Color Spaces	HSV, YIQ (6 channels total)
Initial Dimensionality	60 features (10-bin histograms per channel)
Standardization	Z-score (StandardScaler)
Dimensionality Reduction (PCA)	95% retained variance (~21 components)
Balancing Strategy (Training only)	SMOTE (Target: 1,462 samples per class)
Algorithm	K-Nearest Neighbors (KNN)
Number of Neighbors (K)	5
Distance Metric	Euclidean
Weighting Mechanism	Uniform

Source: Authors

- *Convolutional Neural Networks (CNNs):* A custom CNN (Figure 6) was implemented without relying on transfer learning architectures to automate the extraction of morphological and textural features directly from the raw image tensors. The hierarchical architecture and optimization pipeline were developed using TensorFlow 2 and Keras frameworks [24]. Unlike the classical baselines, which relied on manually engineered global chromatic descriptors, the CNN exploits the necrotic lesions' spatial topology and localized structural variance. The SAM-segmented images were spatially resized, cast to floating-point precision, and scalar normalized to ensure gradient stability during backpropagation.

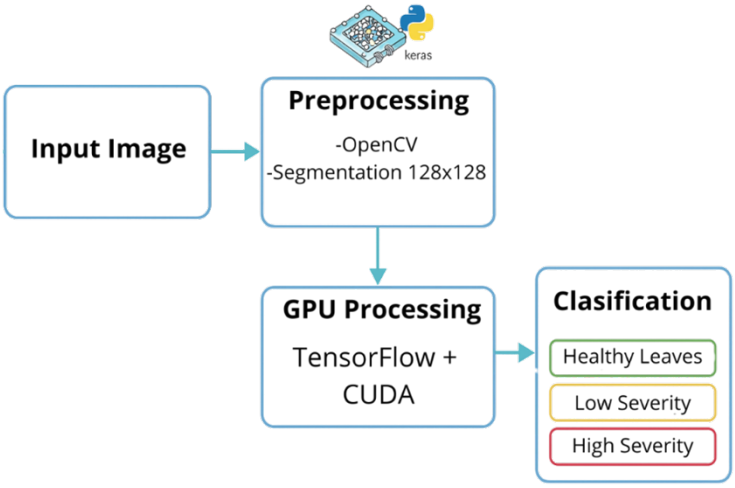

Figure 6. Architectural layout and processing workflow of the custom Convolutional Neural Network (CNN).

Source: Authors

A stochastic spatial data augmentation strategy was implemented to address the biological class imbalance and prevent overfitting. To prevent data leakage, these transformations were executed strictly within the training partition of each unified Monte Carlo iteration. Instead of interpolating vectors in a latent space like SMOTE, synthetic image tensors were generated on-the-fly using affine spatial rotations and additive Gaussian noise until class parity was achieved. This ensured that the model was exposed to high morphological variance without distorting the independent test partitions' original distributions.

The network topology follows a sequential feed-forward architecture, comprising three convolutional blocks for progressive feature abstraction, followed by fully connected layers for multiclass prediction. The independent test partitions were evaluated using the frozen weights of the trained network, strictly without any spatial augmentation. To avoid textual redundancy and maintain cross-architectural traceability, the exact structural configurations including tensor dimensions, precision formats, filter depths, kernel sizes, regularization mechanisms, and optimization hyperparameters are consolidated in Table 4.

Table 4. *Architectural specifications and hyperparameter configuration of the CNN*

Component /Layer	Specifications / Parameters	Function / Purpose
Input Resolution	128 x 128 x 3 (RGB)	Morphological feature preservation
Preprocessing	Min-Max Normalization [0,1]	Numerical stability and convergence
Data Balancing	Dynamic Augmentation (± 45°, σ – 0.03)	Bias mitigation and dataset expansion
Conv. Layers	3-Block (32, 64, 128 filters)	Hierarchical feature extraction
Kernel Size	3 x 3, ReLU Activation	Nonlinear spatial pattern recognition
Regularization	Batch Normalization & Dropout (0.2)	Prevention of overfitting
Downsampling	Max-Pooling (2 x 2)	Spatial dimensionality reduction
Dense Layer	1024 units	Global feature integration
Optimizer	Adam (Learning Rate = 0.001)	Adaptive weight optimization
Training Setup	100 Epochs, Batch Size = 32	Iterative model convergence

Source: Authors

Evaluation metrics

The predictions from the independent test partitions of all Monte Carlo iterations were aggregated to quantify and compare the classification performance across the distinct architectural paradigms. The quantitative assessment was based on macro-averaged metrics to ensure equitable evaluation across the severity classes: accuracy, precision, recall, and F1-Score. Additionally, a normalized confusion matrix was computed for each model to strictly evaluate specific interclass misclassifications and assess the sensitivity of the algorithms in identifying early-stage necrotic lesions

Results

This section presents the quantitative evaluation of the proposed classification framework. Following the unified Monte Carlo cross-validation protocol (10 iterations), the predictive performance of the classical baselines (SVM and KNN) and the custom deep learning architecture (CNN) is detailed. The findings are structured into two primary analyses: an evaluation of the topological separability of disease stages within the feature space and a comparative assessment of the classification metrics (accuracy, precision, recall, and F1-score) and interclass confusion matrices.

Feature Space Topology and Class Separability

Prior to evaluating the classification algorithms, the distribution of the chromatic descriptors was analyzed to understand the intrinsic separability of the phytopathological stages. Figure 7 shows the first two principal components (PC1 and PC2) extracted from the standardized HSV/YIQ feature space, which cumulatively explain 33.0% of the total geometric variance.

Visually, the projection demonstrates a continuous topological gradient along the primary axes, transitioning from healthy tissue to severe necrosis. While the "Healthy" and "Extreme severity" classes form relatively distinct clusters at opposite ends of the geometric space, a significant spatial overlap is observed between the "Healthy" and "Low severity" samples within the central region. This bidimensional chromatic overlap anticipates the intrinsic challenge of classifying early-stage pathogenic lesions using linear boundaries, justifying the projection of the data into higher-dimensional spaces via PCA (for KNN evaluation) or the automated extraction of localized spatial topologies (via CNN) to achieve optimal class discrimination.

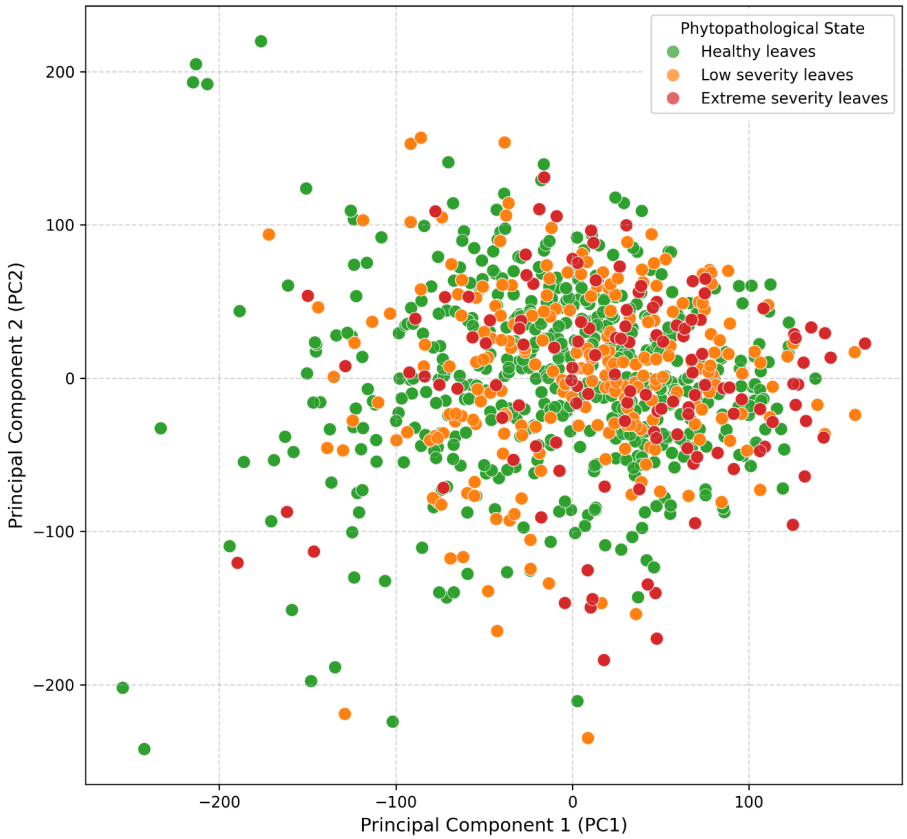

Figure 7. 2D PCA projection of the extracted chromatic features illustrating class distribution and topological overlap

Source: Authors

Comparative classification performance

Table 5 summarizes the macro-averaged classification metrics obtained using the unified 10-iteration Monte Carlo cross-validation framework to quantify the predictive capability of the evaluated algorithms.

Table 5. *Macro-averaged performance metrics across the evaluated classification models (10-iteration Monte Carlo Cross-Validation)*

Model / Architecture	Feature Extraction Strategy	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
SVM (RBF Kernel)	Manual (HSV/YIQ + PCA)	79.15 ± 2.20	78.40 ± 2.83	79.79 ± 2.33	78.88 ± 2.53
KNN (K=5)	Manual (HSV/YIQ + PCA)	98.74 ± 0.39	98.15 ± 0.66	98.15 ± 0.64	98.13 ± 0.57
CNN (Custom 3-Block)	Automated Spatial Tensors	98.54 ± 0.42	98.31 ± 0.45	98.48 ± 0.51	98.39 ± 0.48

Source: Authors

Relying on global hyperplanes, the SVM baseline achieved a macro-averaged Accuracy of 79.15% (±2.20%) and an F1-Score of 78.88% (±2.53%). This performance baseline mathematically confirms the geometric challenge anticipated in the prior topological analysis: the transition from healthy tissue to early-stage necrosis is not linearly separable within a global chromatic feature space.

In contrast, the distance-based KNN and hierarchical CNN models demonstrated significantly higher classification metrics, yielding F1-Scores of 98.13% (±0.57%) and 98.39% (±0.48%), respectively. The success of the KNN algorithm confirms that the pathogenic signatures form dense, localized micro-clusters within the reduced 21-dimensional PCA space, which are effectively resolved using Euclidean proximity metrics, whereas global hyperplanes fail.

Although the KNN and CNN exhibited statistically equivalent global accuracies (98.74% ± 0.39% and 98.54% ± 0.42%), the deep learning approach demonstrated a slightly superior balance across the precision and recall distributions. This quantitative variance suggests that the automated extraction of spatial morphological topologies (texture and lesion shape) via the CNN mitigates the residual chromatic biases of the manual global descriptors.

Interclass Confusion Analysis

The aggregated and row-normalized confusion matrices from the Monte Carlo cross-validation are analyzed below to evaluate the specific classification boundaries and identify each model's algorithmic limitations across the phytopathological progression.

The normalized matrix for the SVM baseline (Figure 8) reveals the structural limitations of relying exclusively on global chromatic descriptors. The classifier demonstrated its highest sensitivity in the "Extreme severity" category (87.92%), where advanced necrosis presents a distinct chromatic contrast against healthy tissue. However, the model struggled significantly with the transitional "Low severity" stage, accurately classifying only 70.49% of the instances. A critical 23.65% of "Low severity" leaves were misclassified as "Healthy," while 18.31% of "Healthy" samples were erroneously predicted as "Low severity." This symmetrical confusion numerically confirms that global color histograms are insufficient to establish a precise linear boundary between uninfected tissue and early-stage localized chlorosis.

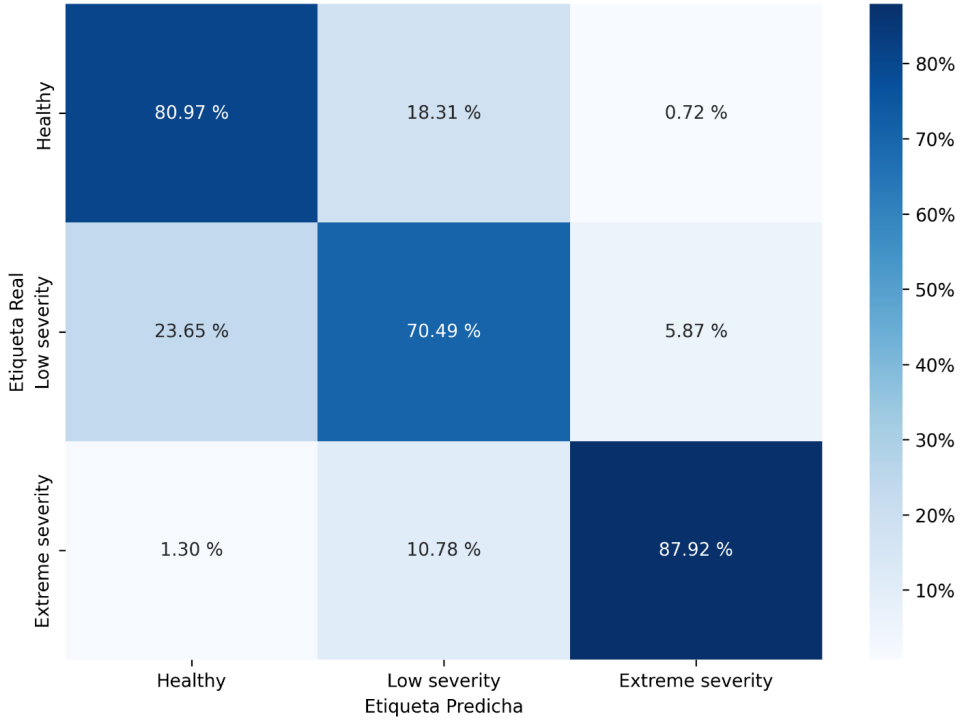

Figure 8. Normalized average confusion matrix for the SVM classifier

Source: Authors

Transitioning to the local density estimation approach, the KNN matrix (Figure 9) exhibits a substantial reduction in early-stage misclassifications. The distance-based evaluation within the PCA latent space successfully separated the initial chromatic overlap, correctly classifying the early infections. Nevertheless, a specific topological limitation was identified in the advanced disease stage: 3.57% of the "Extreme severity" samples were misclassified as "Low severity," limiting its accuracy in this class to 96.30%. Algorithmically, because the KNN evaluates the average chromatic vector, leaves with extensive necrosis that still retain significant portions of healthy green tissue are mathematically pulled toward the intermediate severity cluster in the Euclidean space.

The custom CNN architecture (Figure 10) effectively resolved the classical baseline residual misclassifications. By automating the extraction of localized spatial morphologies (such as the specific texture, shape, and distribution of necrotic lesions) rather than relying on global color averages, the deep learning model increased the classification accuracy of the "Extreme severity" class to 98.72%. Notably, the specific confusion rate between extreme and low severity was reduced to 0.00%. This matrix confirms that spatial tensor extraction mitigates the chromatic averaging bias, ensuring equitable predictive sensitivity across all stages of phytopathological progression without sacrificing early-stage detection accuracy.

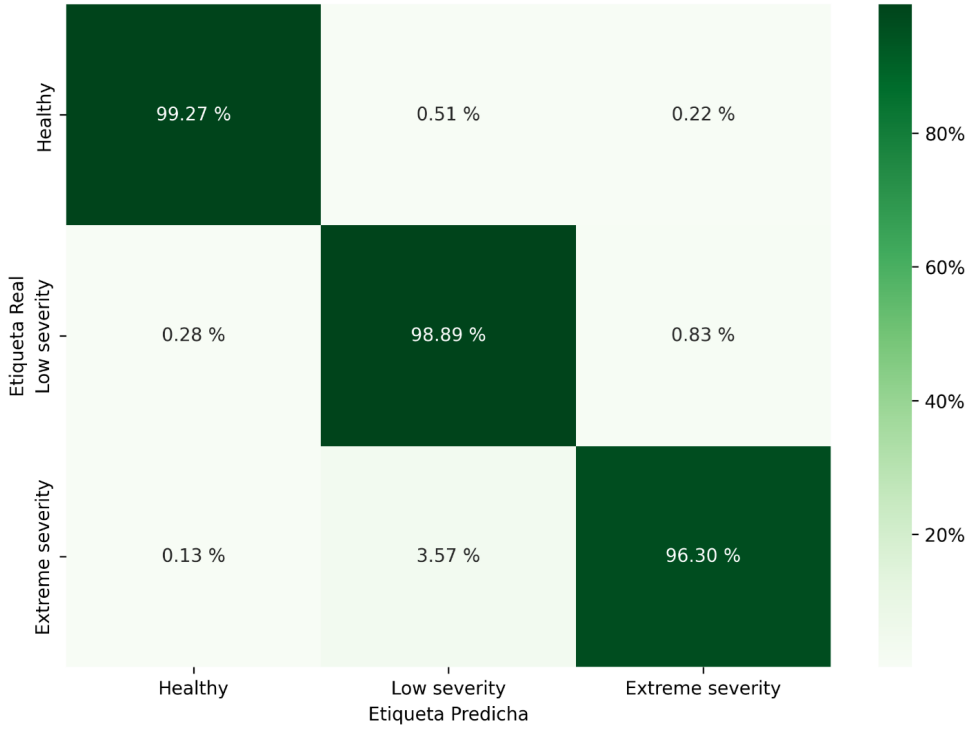

Figure 9. Normalized average confusion matrix for the KNN classifier (K=5)

Source: Authors

The quantitative results indicate that the SVM achieves a baseline macro-averaged F1-Score of 78.88%. As evidenced by the confusion matrices, color averaging alone is insufficient to reliably discriminate early pathological stages (low severity) from healthy tissue owing to their high chromatic similarity. Although the projection of these descriptors into a PCA latent space coupled with local density estimation (KNN) significantly improves the classification metric (F1-Score: 98.13%), it retains a 3.57% false-negative rate in extreme severity phenotypes due to the geometric limitations of averaging residual green tissue.

Within the evaluated dataset, the implementation of the custom Convolutional Neural Network addressed these morphological ambiguities. In contrast to the classical baselines that required manual feature engineering (HSV/YIQ transformations and deterministic PCA reduction), the deep learning architecture automated the extraction of spatial and textural hierarchies directly from the normalized pixel tensors. This architectural approach mitigated the chromatic biases observed in the probabilistic and geometric models, eliminating the false negatives in the extreme severity class and yielding a final F1-Score of 98.39%.

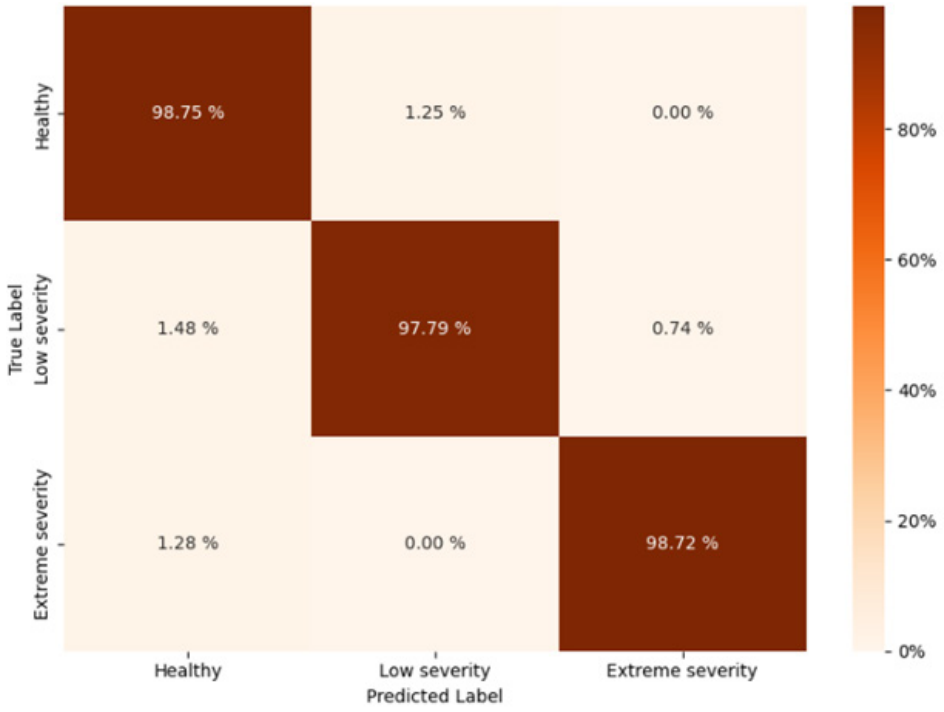

Figure 10. Normalized average confusion matrix for the CNN classifier

Source: Authors

Discussion

The classification performance across the evaluated architectures provides critical insights into the progression of Phytophthora infestans. The topological overlap observed in the initial PCA projection physically represents the disease's biological continuum. Early-stage localized chlorosis (Low severity) inherently shares significant chromatic variance with pristine tissue. This biological reality explains the structural limitations of the SVM baseline; the transition from a healthy leaf to early infection is not linearly separable, causing global hyperplanes to misclassify transitional stages. Conversely, the success of the distance-based KNN model confirms that pathological signatures form dense, localized microclusters in higher dimensional spaces, whereas global separation fails.

However, relying exclusively on global chromatic aggregation introduces a specific taxonomic vulnerability. As evidenced by the residual misclassifications in the advanced disease stage of the KNN, when its colors are averaged, a severely necrotic leaf that retains significant portions of healthy green tissue is mathematically pulled toward the intermediate severity cluster. The custom Convolutional Neural Network effectively resolved this artifact. By automating the extraction of localized spatial tensors and analyzing the specific texture, shape, and distribution of the necrotic spots rather than averaging their global color the deep learning architecture proved to be resilient to chromatic biases. This

demonstrates that spatial morphology provides an essential discriminative value that is lacking in pure color histograms.

Furthermore, the competitive performance of the classical models (particularly the KNN) highlights the fundamental impact of the preprocessing pipeline. The high geometric separability of the dataset is fundamentally driven by the rigorous isolation of the leaf region via SAM segmentation. Eliminating nonbotanical background noise drastically reduced the intraclass variance. This structural condition suggests that strict, automated segmentation can elevate the predictive capabilities of classical, low-resource ML algorithms to levels traditionally reserved for deep learning

Scope and Limitations

These findings must be strictly interpreted within the experimental design boundaries. The dataset was acquired inside a controlled greenhouse facility characterized by uniform diffuse illumination, regulated microclimates, and systematically induced disease progression. The reported metrics represent the upper bound of the algorithmic separability under ideal conditions. Uncontrolled variables, such as dynamic lighting, severe shadows, foliar occlusions, and co-occurring abiotic stressors, introduce substantial morphological heterogeneity in open-field agriculture. Therefore, while this framework successfully validates the theoretical separability of the disease stages using deep and classical architectures, its direct generalization to real-world edge devices remains a scope limitation. Future research must prioritize the validation of this pipeline against diverse, open-field datasets to calibrate the models against uncontrolled environmental noise.

Conclusions

This study evaluated a supervised classification framework for grading the severity of Phytophthora infestans in tomato leaves, contrasting classical machine learning baselines with a custom deep learning architecture. The results indicate that the rigorous isolation of the foliar region via the Segment Anything Model (SAM) is a determinative factor in algorithmic performance. By eliminating non-botanical background noise, intra-class chromatic variance was sufficiently reduced to allow local distance-based models (KNN) to achieve a macro-averaged F1-Score of 98.13% (± 0.57%), performing competitively alongside the deep learning approach within the controlled experimental setup.

While the KNN and the custom Convolutional Neural Network (CNN) demonstrated statistically equivalent global accuracies (98.74% ± 0.39% and 98.54% ± 0.42%, respectively), the deep learning architecture proved structurally superior in resolving specific taxonomic ambiguities. Relying on global chromatic descriptors in the classical baselines induced a false-negative rate of 3.57%, resulting in the misclassification of extreme-severity phenotypes into the low severity class due to residual green tissue

averaging. The CNN effectively mitigated this geometric limitation by automating the extraction of localized spatial topologies specifically the texture and morphology of the necrotic lesions eliminating the misclassification rate for extreme necrosis and yielding a final F1-Score of 98.39% (±0.48%).

Despite the stability of these classification metrics, the reported performance is strictly bound to the controlled greenhouse conditions (uniform diffuse illumination, regulated microclimate, and isolated backgrounds) under which the dataset was acquired. Consequently, the direct deployment of this framework to open-field agriculture remains a structural limitation. Future research must prioritize the physical expansion of the dataset through continuous monitoring in uncontrolled environments. The architectures must be validated against dynamic lighting, foliar occlusions, and co-occurring abiotic stressors to objectively calibrate the algorithms for edge-device integration and real-world phytosanitary monitoring.Future research will focus on two primary objectives to address the current limitations of the classification framework. First, the image dataset was expanded through continuous monitoring across diverse agricultural environments, aiming to capture a broader range of real-world phytopathological stages, dynamic lighting conditions, and foliar occlusions. Second, advanced transfer learning paradigms, specifically exploring attention-based architectures (such as Vision Transformers or hybrid CNN-Transformers). This study aims to systematically compare the proposed custom CNN's spatial feature extraction efficiency, capacity to model long-range morphological dependencies, and computational overhead. This expansion is strictly required to objectively calibrate the algorithms for deployment in open-field precision agriculture and edge-device integration.

Data Availability Statement

The image dataset generated and analyzed during the current study, including the SAM-segmented foliar images categorized by Phytophthora infestans severity and the source code for the implemented architectures, are publicly available in the GitHub repository: https://github.com/hk5ymt/ClasificadorHojasTomate

References

[1] DANE, "Boletín mensual insumos y factores asociados a la producción agropecuaria. Sistema de información de precios y abastecimientos del sector agropecuario," Bogotá, Colombia, 2019. [Online]. Available on: https://www.agronet.gov.co/Lists/Boletin/Attachments/2598/Bol_Insumos_mar_2019.pdf

[2] J. P. Gutiérrez Acuña, "Evaluación de variables productivas de injertos de tomate tipo grape (*Solanum lycopersicum*) bajo invernadero," trabajo de grado, Universidad Nacional, Buenos Aires, Puntarenas, Costa Rica, 2019. [Online]. Available on: http://hdl.handle.net/11056/18161

[3] M. S. Valenzuela Molina, "Caracterización económica de la cadena agroalimentaria del tomate de uso industrial," ODEPA, Santiago de Chile, 2019. [Online]. Available: https://www.odepa.gob.cl/publicaciones/caracterizacion-economica-de-la-cadena-agroalim

[4] A. Paredes Zambrano, *Manual del cultivo de tomate en invernadero*. Bogotá, Colombia: Corporación Colombiana de Investigación Agropecuaria – AGROSAVIA, 2009. [Online]. Available: http://hdl.handle.net/20.500.12324/2196

[5] G. R. Castillo Gavidia, "Evaluación de la rotación y periodos de aplicación de productos para el manejo de *Phytophthora infestans* en tomate (*Solanum lycopersicum*) a campo abierto," trabajo de grado, Escuela Superior Politécnica de Chimborazo, Riobamba, Ecuador, 2022. [Online]. Available: http://dspace.espoch.edu.ec/handle/123456789/17217

[6] Agronet, "Plagas y Enfermedades del Tomate". Bogotá: Ministerio de Agricultura – MinAgricultura, 2020. [Online]. Available: https://www.agronet.gov.co/Noticias/Paginas/Plagas-y-Enfermedades-del-Tomate.aspx

[7] J. E. Berdúo-Sandoval, J. A. Ruiz-Chután, L. Méndez, L. Mejía, D. Maxwell, y A. Sánchez-Pérez, "Detección de patógenos asociados a la enfermedad punta morada en los cultivos de papa y tomate en Guatemala," *Cienc. Tecnol. Salud*, vol. 7, núm. 2, pp. 205–217, 2020. https://doi.org/10.36829/63CTS.v7i2.794

[8] W. E. Fry *et al.*, "Five reasons to consider *Phytophthora infestans* a reemerging pathogen," *Phytopathology*, vol. 105, núm. 7, pp. 966–981, 2015. https://doi.org/10.1094/PHYTO-01-15-0005-FI

[9] D.-D. Leal Lara, J. Barón Velandia, y C. E. Rocha Calderón, "Interpretabilidad en el campo de la detección de enfermedades en las plantas: Una revisión," *Rev. Fac. Ing.*, vol. 30, núm. 58, art. e13495, 2021. https://doi.org/10.19053/01211129.v30.n58.2021.13495

[10] A. L. Salas Gómez, E. Osorio Hernández, C. A. Espinoza Ahumada, R. Rodríguez Herrera, M. T. de J. Segura Martínez, y E. Neri Ramírez, "Principales enfermedades del cultivo de tomate (*Solanum lycopersicum* L.) en condiciones de campo," *Cienc. Lat. Rev. Cient. Multidiscip.*, vol. 6, núm. 1, pp. 4190–4210, 2022. https://doi.org/10.37811/cl_rcm.v6i1.1793

[11] A. Gómez Camperos, H. Y. Jaramillo, y G. Guerrero Gómez, "Digital image processing techniques for detection of pests and diseases in crops: A review," *Ing. Compet.*, vol. 24, núm. 1, 2022. https://doi.org/10.25100/iyc.v24i1.10973

[12] V. Gonzalez Huitron, J. A. León Borges, A. E. Rodriguez Mata, L. E. Amabilis Sosa, B. Ramírez Pereda, y H. Rodriguez, "Disease detection in tomato leaves via CNN with lightweight architectures implemented in Raspberry Pi 4," *Comput. Electron. Agric.*, vol. 181, art. 105951, 2021. https://doi.org/10.1016/j.compag.2020.105951

[13] A. Abbas, S. Jain, M. Gour, y S. Vankudothu, "Tomato plant disease detection using transfer learning with C-GAN synthetic images," *Comput. Electron. Agric.*, vol. 187, art. 106279, 2021. https://doi.org/10.1016/j.compag.2021.106279

[14] S. Adhikari, B. Shrestha, S. K. K.C., y B. Baiju, "Tomato plant diseases detection system using image processing," en Proc. *1st KEC Conf.*, Kantipur Engineering College, Dhapakhel, Lalitpur, Nepal, 2018, ISBN: 978-9937-0-4872-9. [Online]. Available: https://www.researchgate.net/publication/327930730

[15] K. Zhang, Q. Wu, A. Liu, y X. Meng, "Can deep learning identify tomato leaf disease?" *Adv. Multimed.*, vol. 2018, art. 6710865, pp. 1–10, 2018. https://doi.org/10.1155/2018/6710865

[16] R. Karthik, M. Hariharan, S. Anand, P. Mathikshara, A. Johnson, y R. Menaka, "Attention embedded residual CNN for disease detection in tomato leaves," *Appl. Soft Comput.*, vol. 86, art. 105933, 2020. https://doi.org/10.1016/j.asoc.2019.105933

[17] M. Agarwal, A. Singh, S. Arjaria, A. Sinha, y S. Gupta, "ToLeD: Tomato leaf disease detection using convolution neural network," *Procedia Comput. Sci.*, vol. 167, pp. 293–301, 2020. https://doi.org/10.1016/j.procs.2020.03.225

[18] S. Widiyanto, R. Fitrianto, y D. T. Wardani, "Implementation of convolutional neural network method for classification of diseases in tomato leaves," en *Proc. IEEE Int. Conf. Informatics and Computing (ICIC)*, 2019. https://doi.org/10.1109/ICIC47613.2019.8985909

[19] L. E. Pamplona Berón, A. F. Calvo Salcedo, y A. Bejarano Martínez, "Detection of foliar diseases using image processing techniques," *Rev. Ceres*, vol. 67, núm. 2, pp. 100–110, 2020. https://doi.org/10.1590/0034-737X202067020002

[20] E. Marín-García, J.-N. Torres-Marín, y A. Chaverra-Lasso, "Smart greenhouse and agriculture 4.0," *Rev. Cient.*, vol. 46, núm. 1, pp. 37–50, 2023. https://doi.org/10.14483/23448350.1981

[21] G. Boesch, "Segment Anything Model (SAM) explained," Viso.ai, dic. 2024. [Online]. Disponible en: https://viso.ai/deep-learning/segment-anything-model-sam-explained/

[22] V. N. Vapnik, *The Nature of Statistical Learning Theory*, 2.ª ed. Berlin, Germany: Springer, 2000. https://doi.org/10.1007/978-1-4757-3264-1

[23] N. S. Altman, "An introduction to kernel and nearest-neighbor nonparametric regression," *Am. Stat.*, vol. 46, núm. 3, pp. 175–185, 1992. https://doi.org/10.2307/2685209

[24] J. T. Viñals, *Python Deep Learning: introducción práctica con Keras y TensorFlow 2*. Barcelona, España: Editorial Marcombo, 2020. [Online]. Available: https://books.google.com.co/books?id=5vpmzQEACAAJ