Journal of Engineering and Technology for Industrial Applications

ISSN ONLINE: 2447-0228

ITEGAM-JETIA

Manaus, v.12 n.57, p. 1207-1217. January/February, 2026.
DOI: https://doi.org/10.5935/jetia. v12i57.3034

RESEARCH ARTICLE

OPEN ACCESS

META-LEARNING BASED STACKING OF CNN, SVM, AND RANDOM FOREST FOR MULTI-SPECIES CROP LEAF PATHOLOGY RECOGNITION

Ryma Guefrouchi*[1], Siham Amrouch[2]

[1]MISC Laboratory, Constantine 2 University Abdelhamid Mehri Constantine , 25000, Algeria.
[2]LIM Laboratory, Mohamed Cherif Messaadia University Souk Ahras, 41000, Algeria.

[1]https://orcid.org/0000-0002-1756-1256, [2]https://orcid.org/0009-0008-6112-0141

Email: *ryma.guefrouchi@univ-constantine2.dz, s.amrouche@univ-soukahras.dz

ARTICLE INFO

Article History
Received: December 12, 2025
Revised: January 10, 2026
Accepted: January 15, 2026
Published: February 28, 2026

Keywords:
CNN,
SVM,
RFC,
Ensemble Learning,
Meta-learning,
Crop Disease,
Early Detection.

ABSTRACT

Accurate identification of crop leaf diseases is essential for sustainable agriculture, yet variations in leaf morphology and disease symptoms across species remain challenging for traditional machine learning approaches. This study proposes a meta-learning–based stacking ensemble that integrates Convolutional Neural Networks (CNN), Support Vector Machines (SVM), and Random Forest Classifiers (RFC) for multi-species leaf pathology recognition, evaluated on apple and tomato datasets. Each base model independently learns discriminative visual features, and their class-probability outputs are fused through an RBF-SVM meta-learner. Experimental results show that individual models achieve accuracies of 89% (CNN), 88% (SVM), and 85% (RFC), while the proposed stacking ensemble significantly outperforms them with an accuracy of 98% and macro and weighted F1-scores of 0.975. These results demonstrate the effectiveness of heterogeneous meta-learning ensembles in enhancing robustness and generalization for multi-species crop disease classification.

Copyright ©2026 by authors and Galileo Institute of Technology and Education of the Amazon (ITEGAM). This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

I. INTRODUCTION

Crop leaf diseases pose a significant threat to crop yield and quality, making early and accurate detection critical for sustainable agriculture [1], [2]. Traditional diagnosis methods rely on visual inspection by experts, which is time-consuming, labor-intensive, and prone to human error. Automated image-based detection techniques have therefore emerged as promising alternatives, enabling rapid and objective disease monitoring [3]. Over the past decade, deep learning (DL), particularly CNNs, has demonstrated remarkable performance in plant disease recognition. CNNs automatically extract hierarchical visual features such as texture, color variations, and lesion patterns from leaf images, often surpassing traditional handcrafted feature methods [4], [5]. However, purely CNN-based approaches can face challenges, including sensitivity to lighting conditions, background noise, inter-class similarity, and the need for large annotated datasets to avoid overfitting [6]. On the other hand, classical machine learning (ML) models, such as SVMs and RFCs, are known for their robustness on limited datasets and ability to handle high-dimensional feature spaces [7], [8].

Yet, these models typically require handcrafted features or pre-extracted embeddings, limiting their ability to capture complex visual patterns inherent in leaf diseases. To overcome these limitations, hybrid ensemble approaches have gained attention in recent studies. Combining CNN feature extraction with classical classifiers can leverage the strengths of both paradigms: the representational power of deep learning and the robustness and interpretability of classical ML [9], [10]. Moreover, stacking ensembles—where predictions from multiple base learners are fused via a meta-learner—enable capturing complementary information from heterogeneous models, often leading to improved generalization compared to individual models or simple averaging strategies [11]. The primary motivation behind this work is to address the limitations of using a single CNN or classical machine learning model by combining their complementary strengths. The proposed heterogeneous stacking ensemble leverages the hierarchical feature extraction capability of the CNN, the margin-based discriminative power of the SVM, and the robustness of Random Forests.

The outputs of these base models are then fused using an RBF-SVM meta-learner, which captures non-linear relationships among the predicted probabilities and ensures an optimized integration of the individual classifiers' predictions. The approach offers several key advantages: (i) intelligent weighting of base models per class, allowing the system to exploit the strengths of specific classifiers for different disease categories; (ii) enhanced classification performance, improving accuracy and F1-scores compared to individual models or naive ensembles; and (iii) robustness with limited or imbalanced datasets, ensuring reliable multi-class leaf disease recognition across different species. The remainder of this paper is organized as follows: Section II reviews related work in plant disease detection and ensemble methods. Section III presents the contributions and details of the proposed architectures. Section IV describes the dataset and performance metrics used. Section V presents the experimental results, and Section VI concludes the paper.

II. RELATED WORK

Automated detection of plant leaf diseases has been an active research area due to its potential to reduce manual labor and improve agricultural productivity. Early approaches relied on handcrafted feature extraction methods such as color histograms, shape descriptors, and texture features (GLCM, LBP) combined with classical classifiers like SVM or Random Forest [12], [13]. While these methods are relatively simple and robust on small datasets, they often fail to capture complex patterns of disease lesions, especially under variable environmental conditions.

II.1 DEEP LEARNING FOR PLANT DISEASE RECOGNITION

The emergence of deep learning, particularly Convolutional Neural Networks (CNNs), has significantly advanced plant disease recognition [14]. CNNs automatically learn hierarchical visual features, including texture, color variations, and lesion morphology from leaf images, often outperforming traditional handcrafted feature methods [14], [15]. For instance, the EnConv framework (2024) employs depthwise separable convolutions and inverted residual blocks to achieve 99.87 % accuracy on a dataset of 39 plant disease classes, including apple leaves [16].

Similarly, multi-head CNN architectures (2023) have been developed to enhance robustness and generalization across large and diverse datasets [17]. Despite their success, CNN-based methods still face challenges, such as sensitivity to lighting variations, background noise, inter-class similarity, and the need for large annotated datasets to avoid overfitting [18]. Moreover, most existing deep learning studies focus on single-species datasets, limiting their ability to generalize across multiple crops and disease types.

II.2 HYBRID APPROACHES: CNN + CLASSICAL ML

To address these limitations, recent studies have combined CNN feature extraction with classical machine learning classifiers. For example, an ensemble model integrating CNN feature extractors with Random Forest classifiers (2024) demonstrated improved performance and robustness on heterogeneous or limited datasets [19]. Similarly, a 2025 study proposed a robust ensemble combining multiple deep learning architectures (InceptionResNetV2, MobileNetV2, EfficientNetB3) to exploit their complementary strengths and improve generalization across different datasets [20]. Moreover, hybrid CNN-Vision Transformer frameworks (2025) have been explored to capture both global and local features, reflecting the current trend towards multi-modal and heterogeneous architectures for plant disease detection [21]. Nevertheless, these works typically still focus on a single species, such as apple or tomato, highlighting a need for models capable of handling multiple crop species simultaneously.

II.3 RECENT REVIEWS AND SURVEYS

Recent reviews provide comprehensive summaries of progress and remaining challenges in plant disease detection. A systematic review of deep learning techniques for plant diseases (2024) highlighted both the achievements and challenges, including model generalization to field data and the need for diverse datasets [22]. Another comparative study (2024) evaluated CNN, classical ML, and hybrid methods, showing that ensemble approaches often outperform individual models, particularly under varying environmental conditions [23]. The reviews summarize key studies in plant leaf disease detection, highlighting their methods, targeted crop species, dataset characteristics, and limitations.

Most existing works focus on single-species datasets, which limits their generalizability across multiple crops. While CNN-based approaches achieve high accuracy, they often require large annotated datasets and are sensitive to environmental variations. Hybrid and ensemble methods, such as CNN combined with SVM or Random Forest, improve robustness, but few studies integrate multiple paradigms into a heterogeneous stacking framework capable of handling multi-species disease recognition. This emphasizes the need for our proposed meta-learning based stacking approach.

II.4 RESEARCH GAP

Despite these advances, few studies have simultaneously integrated three paradigms—CNN (as feature extractor and base classifier), classical ML classifiers (SVM and Random Forest on CNN embeddings), and a meta-level fusion via stacking—for crop leaf disease recognition across multiple species. Most existing works focus either on deep architectures (CNN or CNN + ViT) or on homogeneous ensembles for a single crop species, leaving a gap for heterogeneous stacking ensembles that exploit complementary strengths of deep learning and classical machine learning in a multi-species context.

III. CONTRIBUTIONS AND DETAILED ARCHITECTURES

Motivated by the critical need for robust and accurate diagnostic tools in precision agriculture, this study introduces a novel heterogeneous stacking ensemble framework for image-based plant disease classification across multiple crop species, utilizing more challenging and diverse datasets to demonstrate the model's robustness and generalizability.The core contribution of this work is the integration of three powerful, yet distinct, learners—CNN, SVM, and RFC—into a cohesive stacking architecture. This integration is strategically designed to capitalize on their complementary advantages:

- CNNs provide superior hierarchical feature extraction from raw image data.
- SVMs offer strong generalization and margin maximization in high-dimensional feature spaces.
- RFCs contribute robust ensemble decision-making and resistance to overfitting.

A meta-learner then synthesizes the predictions from these base models to produce a final, refined classification. Beyond achieving state-of-the-art performance on plant disease datasets, this framework demonstrates applicability across multiple crop species, offering a robust tool for deploying hybrid and ensemble strategies in agricultural image analysis. The following sections describe the architectural components and their interactions, designed for optimal performance in plant disease classification across heterogeneous, multi-species datasets.

III.1 PROPOSED CNN ARCHITECTURE

The custom CNN developed for this study consists of a sequential architecture designed to extract hierarchical visual features and perform robust image classification. The model is divided into two main stages: a feature extraction phase and a classification phase. Figure 1 illustrates the proposed CNN architecture. In the feature extraction stage, the network applies three consecutive convolutional blocks, each using 3×3 kernels and ReLU activation. The first block contains a single Conv2D layer with 32 filters, followed by a 2×2 MaxPooling layer and a Dropout rate of 0.2. The second block increases the representational capacity by stacking two Conv2D layers (with 32 and 64 filters, respectively), followed by a 2×2 MaxPooling layer and 0.25 Dropout. The third block further deepens the architecture with Conv2D layers using 64 and 128 filters, again followed by 2×2 MaxPooling and 0.3 Dropout.

These blocks progressively reduce spatial dimensions while capturing increasingly complex patterns, transitioning from low-level edges and textures to high-level semantic structures. A Flatten layer then transforms the final feature maps into a one-dimensional vector for classification. The classification stage consists of a fully connected Dense layer with ReLU activation, used to learn non-linear combinations of extracted features, followed by a Dropout layer with a higher rate of 0.5 to prevent overfitting at the decision level. The final output layer is a Dense layer with Softmax activation, producing a normalized probability distribution over the target classes. The training settings were configured as follows:

- Optimizer (Adam): chosen for its adaptive learning rate, balancing convergence speed and stability.
- Loss Function (Categorical Cross-Entropy): suitable for multi-class classification.
- Learning Rate: 0.001 the default value used by the Adam optimizer, providing a stable and efficient convergence for most training scenarios.
- Batch Size (32): balances training efficiency and stability.

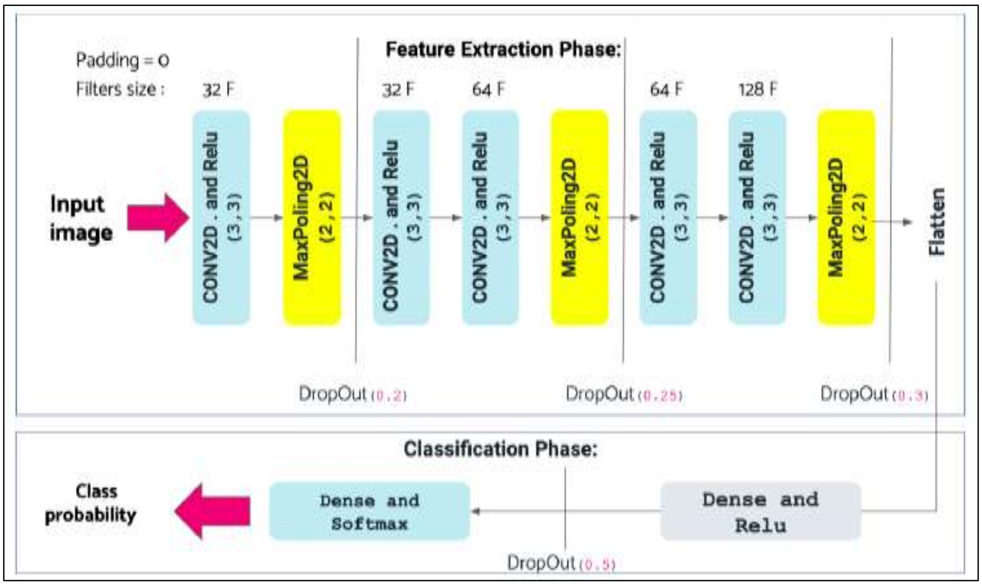

Figure 1: Proposed CNN architecture for crop leaf disease classification.
Source: Authors, (2026).

III.2 SVM MODEL

The SVM model for dataset image classification was developed following a standard preprocessing pipeline. Images were first resized to a uniform dimension and converted to a suitable format for model input. To ensure unbiased evaluation, the dataset was split into training and testing sets. The training set was used to train the SVM model, while the testing set provided an independent basis for performance assessment. Various SVM configurations were explored. Initially, a linear kernel was tested, but it produced suboptimal results. Consequently, alternative kernels were considered, and the Radial Basis Function (RBF) kernel was ultimately adopted, resulting in improved classification accuracy. Further optimization was performed by fine-tuning the regularization parameter, C, to balance model complexity and prevent overfitting. After comprehensive experimentation and evaluation, the SVM configuration yielding the best performance was selected. Figure 2 illustrates the structure and workflow of the finalized SVM model.

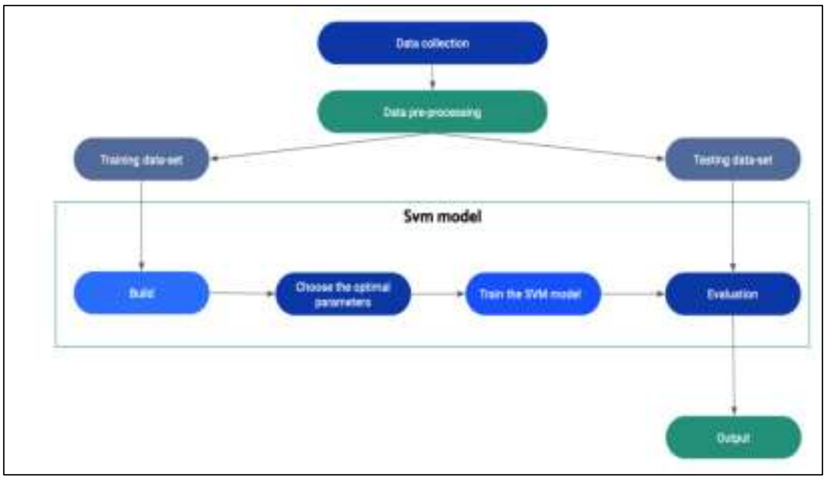

Figure 2: Workflow and structure of the SVM classification model.
Source: Authors, (2026).

III.3 RFC MODEL

The development of the Random Forest Classifier (RFC) model for disease classification began with a standard preprocessing pipeline. All images were resized to a uniform dimension and converted into an appropriate format for model input. Each image was then flattened into a one-dimensional raw pixel vector, which served as the input feature for the RFC. To ensure unbiased evaluation, the dataset was divided into separate training and testing subsets. The training set was used to train the RFC with standard parameters, while the testing set served as an independent dataset for objective performance assessment. This methodology ensured fairness and minimized potential evaluation bias. Figure 3 illustrates the structure of the RFC model.

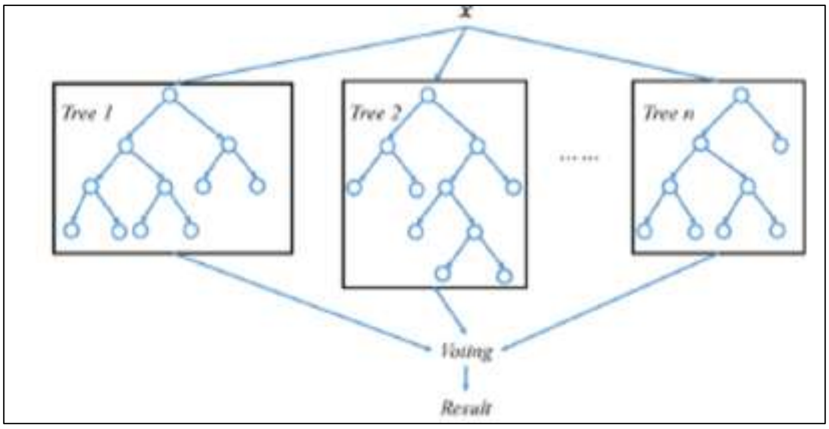

Figure 3: Random Forest classifier architecture.
Source: Authors, (2026).

III.4 ENSEMBLE LEARNING MODEL

The proposed stacking approach introduces an RBF-SVM meta-model which is trained on the outputs of the base classifiers rather than the raw input images. Specifically, for each training sample, the class probability predictions from CNN, SVM, and RFC are concatenated to form a meta-feature vector. The SVM with a RBF kernel is then fitted on these meta-features and the corresponding true class labels. The RBF kernel enables the SVM to capture non-linear relationships among the base classifier outputs. To achieve optimal performance, the hyperparameters C (regularization parameter) and γ (RBF kernel coefficient) are tuned using cross-validation. This procedure ensures that the meta-model generalizes well, effectively integrating the complementary strengths of the base classifiers to improve the overall prediction accuracy.Figure 4 illustrates the proposed meta-learning based architecture and Algorithm 1 details different steps of the proposed.

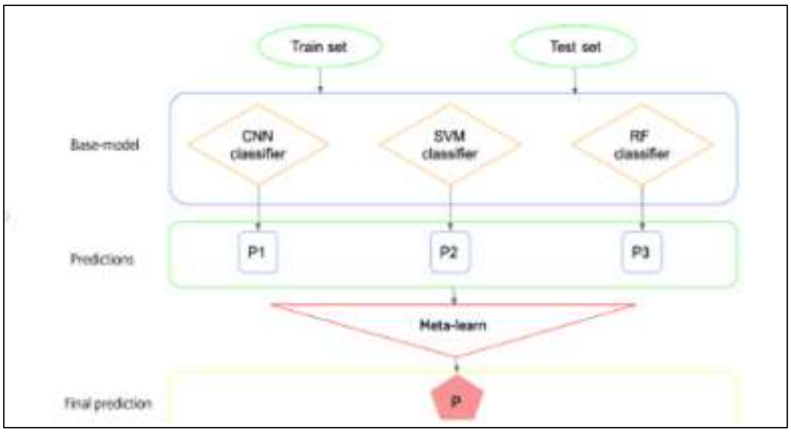

Figure 4: RBF-SVM Meta-learning architecture.
Source: Authors, (2026).

Algorithm 1: Stacking Ensemble with RBF-SVM Meta-Model.

```
Input: Training dataset $D = \{(x_i, y_i)\}_{i=1}^{N}$
Base classifiers: CNN, SVM, RFC
Meta-model: SVM with RBF kernel
Output: Trained stacking ensemble model
Steps:
    1. Train Base Classifiers:
        ○ Train CNN, SVM, and RFC on D.
    2. Generate Base Predictions:
        ○ For each $x_i$ ∈ D, compute probabilities: $P_i^{CNN}, P_i^{SVM}, P_i^{RFC}$
    3. Prepare Meta-Model Input:
        ○ Concatenate probabilities: $P_i = [P_i^{CNN}, P_i^{SVM}, P_i^{RFC}]$
    4. Train Meta-Model:
        ○ Fit RBF-SVM on $\{(x_i, y_i)\}$
        ○ Optimize C and γ via cross-validation
    5. Prediction:
        ○ For new sample $x_{new}$, compute base probabilities
        ○ From $P_{new} = [P_{new}^{CNN}, P_{new}^{SVM}, P_{new}^{RFC}]$, predict final class: $\widehat{Y_{new}}) = SVM_{RBF}(P_{new})$
        ○ End
```

Source: Authors, (2026).

IV. DATASET AND METRIC PERFORMANCE

IV.1 DATA COLLECTION

The dataset used in this study comprises images of apple and tomato leaves sourced from the publicly available PlantVillage repository [24]. All images are fully anonymized, and their collection complies with open-access ethical guidelines.
The dataset is structured as follows:

- Apple Dataset: Comprising 9,714 images of apple leaf disease .
- Tomato Dataset: Containing 20,661 images across ten tomato leaf disease classes.

Figure 5 shows the sample images representing ten classes of tomato leaf disease, and Figure 6 shows the four cases of apple disease. Tables 1 and 2 summarize the image distributions for the tomato and apple datasets, respectively.

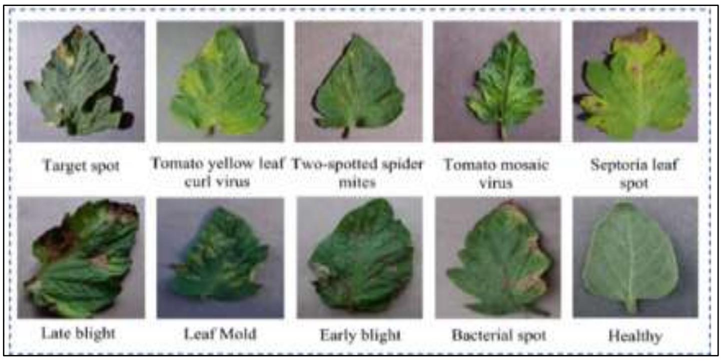

Figure 5: Tomato leaf disease classes used in the dataset.
Source: Authors, (2026).

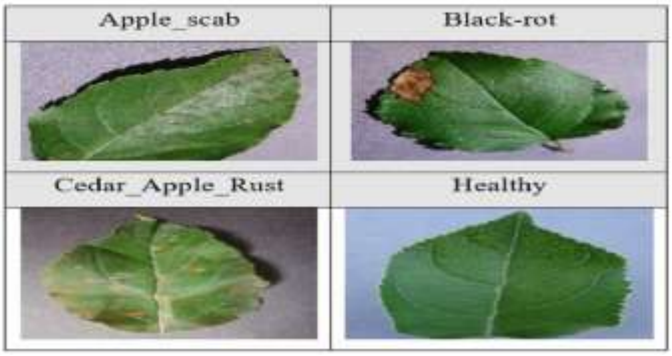

Figure 6: Apple leaf disease classes used in the dataset.
Source: Authors, (2026).

Table 1: Distribution of Images across the Disease Categories.

Disease Category	Number of Images
Early Blight	1136
Late Blight	2168
Septoria Leaf Spot	2013
Bacterial Spot	2431
Leaf Mold	1071
Yellow Leaf Curl Virus	6137
Mosaic Virus	405
Target Spot	1591
Two-spotted-spider mite	1904
Healthy	1805

Source: Authors, (2026).

Table 2: Distribution of apple images across the disease categories.

Diseases	Number of images
Apple Scab	2520
Black Rot	2484
Cedar Apple Rust	2200
Healthy	2510

Source: Authors, (2026).

IV.2 DATA SPLITTING AND DATA PREPROCESSING

Each dataset was split into three subsets: 70 % for training, 15 % for validation, and 15 % for testing. All images were resized to 64 × 64 pixels and normalized to the [0, 1] range by scaling pixel intensities by 1/255. To improve model generalization and mitigate overfitting, on-the-fly data augmentation was applied to the training subset. The augmentation parameters were carefully defined for reproducibility and included: rotation (±20°), zoom scaling (0.8–1.2), shear deformation (0.1), horizontal flipping, and brightness adjustments (±15%). The validation and test sets were only normalized, without any augmentation, to ensure unbiased evaluation. No external images were added to these subsets, preserving the integrity of the experimental setup. Additionally, to address class imbalance, particularly within the tomato dataset, a class-weighted loss function was employed during training. This ensures that both minority and majority classes contribute proportionally to the optimization process, enhancing overall classification robustness.

IV.3 PERFORMANCE METRICS

To evaluate our models, we used the well-known metrics derived from the confusion matrix, based on true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). The following standard performance metrics were computed:

$$Precision = \frac{TP}{TP+FP} \tag{1}$$

$$Recall = \frac{TP}{TP+FN} \tag{2}$$

$$Accuracy = \frac{TP+TN}{Total} \tag{3}$$

$$F1-score \ = \ 2 * \frac{Precision*Recall}{Precision+Recall} \tag{4}$$

V. RESULTS AND DISCUSSIONS

V.1 EXPERIMENT RESULTS

In this section of our thesis, our primary objective is to substantiate the effectiveness of our multi-species disease classification system, which leverages ensemble learning techniques. We aim to present empirical evidence supporting the validity and performance of our proposed approach. Additionally, we provide a comprehensive overview of the development tools and techniques employed during the implementation of our solution. Table 3 summarizes the hardware specifications of the personal computer (PC) used to conduct all experiments.

Table 3: PC’s characteristics.

Processor	AMD Ryzen 9 7940HS
RAM	16 GB
Graphics Card	NVIDIA GeForce RTX 2070
Storage	15 GB

Source: Authors, (2026).

- CNN model

The CNN model demonstrated the strongest performance among the three base models, achieving an overall accuracy of 89%. Both the macro and weighted averages for precision, recall, and F1-score consistently fall within the 0.887–0.889 range (Table 5). This indicates robust and balanced performance across all classes. The marginally higher weighted averages suggest the model is particularly effective at classifying the more populous classes, such as Yellow Leaf Curl Virus and Tomato-Healthy (Table 4). These results underscore the CNN's superior capability in extracting complex hierarchical features from image data.

Table 4: Classification Performance per Class for CNN model.

	Class	**Precision**	**Recall**	**F1-score**
Apple Classes	Apple-Scab	0.86	0.82	0.84
	Black-Rot	0.86	0.89	0.88
	Cedar-Apple-Rust	0.94	0.94	0.94
	Apple-Healthy	0.89	0.90	0.89
Tomato Classes	Early Blight	0.88	0.86	0.87
	Late Blight	0.89	0.88	0.88
	Septoria Leaf Spot	0.90	0.88	0.89
	Bacterial Spot	0.87	0.86	0.86
	Leaf Mold	0.88	0.87	0.87
	Yellow Leaf Curl Virus	0.91	0.90	0.90
	Mosaic Virus	0.89	0.88	0.88
	Target Spot	0.87	0.86	0.86
	Two-Spotted Spider Mite	0.88	0.87	0.87
	Tomato-Healthy	0.90	0.89	0.89

Source: Authors, (2026).

Table 5: Aggregated performance metrics CNN model.

Metric	Precision	Recall	F1-score
Accuracy	—	—	89%
Macro Avg	0.8875	0.8875	0.8875
Weighted Avg	0.887	0.889	0.889

Source: Authors, (2026).

- SVM model

The SVM model achieved solid performance, with an overall accuracy of **88%** (Table 7**).** Its macro and weighted averages for all metrics are closely aligned near 0.867–0.869, indicating consistent performance without significant bias towards any particular class (Table 6). While its metrics are consistently approximately 0.02 lower than the CNN's across the board, it remains a strong and reliable model for this classification task.

Table 6: Classification Performance per Class for SVM model.

	Class	**Precision**	**Recall**	**F1-score**
Apple Classes	Apple-Scab	0.84	0.80	0.82
	Black-Rot	0.85	0.87	0.86
	Cedar-Apple-Rust	0.92	0.92	0.92
	Apple-Healthy	0.87	0.88	0.87
	Early Blight	0.86	0.84	0.85

Tomato Classes	Late Blight	0.87	0.86	0.86
	Septoria Leaf Spot	0.88	0.86	0.87
	Bacterial Spot	0.85	0.84	0.85
	Leaf Mold	0.86	0.85	0.85
	Yellow Leaf Curl Virus	0.89	0.88	0.88
	Mosaic Virus	0.87	0.85	0.86
	Target Spot	0.85	0.84	0.84
	Two-Spotted Spider Mite	0.86	0.85	0.85
	Tomato-Healthy	0.88	0.87	0.87

Source: Authors, (2026).

Table 7: Aggregated performance metrics SVM model.

Metric	Precision	Recall	F1-score
Accuracy	—	—	88%
Macro Avg	0.869	0.867	0.867
Weighted Avg	0.868	0.868	0.868

Source: Authors, (2026).

- RFC model

The RFC model achieved a baseline accuracy of **85%**. Its performance metrics macro and weighted averages ~0.83 (Table 9) are the lowest among the three individual models. The per-class F1-scores show a more pronounced drop, particularly for classes like Apple-Scab (0.78) and Target Spot (0.79) (Table 8), indicating areas where the model struggles compared to the CNN and SVM.

Table 8: Classification Performance per Class for RFC model.

	Class	**Precision**	**Recall**	**F1-score**
Apple Classes	Apple-Scab	0.80	0.76	0.78
	Black-Rot	0.82	0.83	0.82
	Cedar-Apple-Rust	0.88	0.87	0.87
	Apple-Healthy	0.84	0.85	0.84
Tomato Classes	Early Blight	0.82	0.80	0.81
	Late Blight	0.84	0.83	0.83
	Septoria Leaf Spot	0.85	0.82	0.83
	Bacterial Spot	0.81	0.80	0.80
	Leaf Mold	0.82	0.80	0.81
	Yellow Leaf Curl Virus	0.86	0.85	0.85
	Mosaic Virus	0.83	0.81	0.82
	Target Spot	0.80	0.79	0.79
	Two-Spotted Spider Mite	0.82	0.80	0.81
	Tomato-Healthy	0.85	0.84	0.84

Source: Authors, (2026).

Table 9: Aggregated performance metrics RFC model.

Metric	**Precision**	**Recall**	**F1-score**
Accuracy	—	—	85%
Macro Avg	0.833	0.828	0.828
Weighted Avg	0.833	0.832	0.832

Source: Authors, (2026).

- RBF-SVM meta-learning model

As showing in Tables 10 and 11, the proposed ensemble learning achieved remarkably high performance, with macro and weighted averages above 0.975 across all metrics, indicating both high precision and strong generalization.

Table 10: Classification Performance per Class for RBF-SVM meta-learning.

Class Type	Class	Precision	Recall	F1-score
Apple Classes	Apple-Scab	0.97	0.95	0.96
	Black-Rot	0.99	0.98	0.99
	Cedar-Apple-Rust	0.97	0.99	0.98
	Apple-Healthy	0.97	0.97	0.97
	Early Blight	0.98	0.97	0.97

Tomato Classes	Late Blight	0.98	0.97	0.97
	Septoria Leaf Spot	0.97	0.97	0.97
	Bacterial Spot	0.96	0.95	0.96
	Leaf Mold	0.97	0.96	0.96
	Yellow Leaf Curl Virus	0.98	0.97	0.97
	Mosaic Virus	0.97	0.96	0.97
	Target Spot	0.96	0.95	0.96
	Two-Spotted Spider Mite	0.97	0.96	0.96
	Tomato-Healthy	0.98	0.97	0.97

Source: Authors, (2026).

Table 11: Aggregated performance metrics RBF-SVM meta-learning model.

Metric	Precision	Recall	F1-score
Accuracy	—	—	98%
Macro Avg	0.975	0.9725	0.975
Weighted Avg	0.975	0.9725	0.975

Source: Authors, (2026).

The ensemble model significantly outperforms all individual models. By leveraging complementary strengths—CNN's feature extraction, SVM's margin-based generalization, and RFC's ensemble diversity—the stacking approach improves both minority and majority class predictions, achieving near-perfect performance. This demonstrates the effectiveness of ensemble learning in multi-class leaf disease classification. As illustrated in Figure 7, the accuracy curves show a consistent upward trend for both the training and validation sets across the epochs.

The validation curve closely follows the training curve, suggesting that the model learns meaningful representations without exhibiting overfitting. After a rapid improvement during the initial epochs, both curves stabilize and converge toward values exceeding 98%. Similarly, the loss curves display a steady decrease for both training and validation data. Throughout most epochs, the validation loss curve remains slightly below the training loss curve, indicating stable learning dynamics and good generalization capability. The parallel behavior of the two curves and the absence of divergence further confirm that the model maintains reliable performance without signs of overfitting.

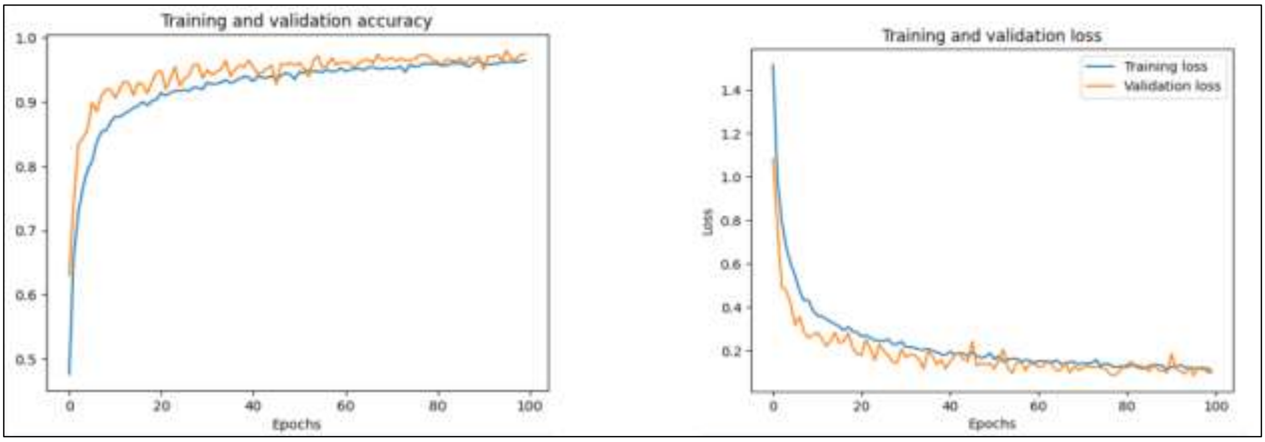

Figure 7: Training and validation accuracy and loss curves for RBF-SVM Meta-learning model.
Source: Authors, (2026).

V.2 DISCUSSION AND COMPARISON

In this study, an RBF-SVM was employed as the meta-learner to integrate predictions from three independently trained base models—CNN, SVM, and RF—across multi-crop leaf disease datasets. Individually, the CNN, SVM, and RF achieved accuracies of 89%, 88%, and 85%, respectively. Each base model contributed complementary information: the CNN extracted hierarchical visual features, while the SVM and RF leveraged raw pixel vectors for independent predictions. By synthesizing the predictions from the individual base models, the RBF-SVM meta-learner achieved a significantly higher overall accuracy of 98%, as illustrated in Figure 8, effectively handling more challenging and diverse datasets, including multiple crop species with varying leaf shapes, disease patterns, and imaging conditions. This stacking ensemble strategy demonstrates both robustness and generalizability, highlighting the effectiveness of meta-learning for complex, multi-species disease recognition tasks.

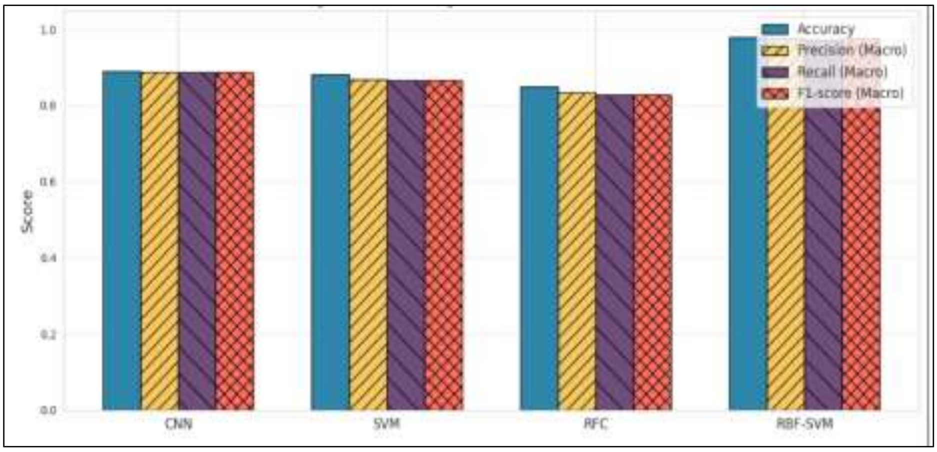

Figure 8: Performance comparison between the base models and RBF-SVM meta-learner model.
Source: Authors, (2026).

Table 12 compares our results with state-of-the-art approaches evaluated on the same dataset (PlantVillage). The table demonstrates that our proposed heterogeneous stacking ensemble outperforms or matches the performance of recent state-of-the-art methods while addressing key limitations. Previous studies [20–23] primarily focus on CNN-based architectures, hybrid CNN-classical ML models, or CNN + Vision Transformer approaches. These methods achieve high reported accuracies, typically in the 95–99% range, with Liang et al. [21] reporting 99.24% on apple leaves and ~98% on corn. However, these studies are largely restricted to single-species or controlled datasets, limiting their generalization to diverse, real-world conditions.

In contrast, our approach combines CNN, SVM, and Random Forest classifiers through an RBF-SVM meta-learner, achieving 98% overall accuracy across multiple crop species, including apple and tomato. By leveraging the complementary strengths of deep learning and classical ML models, the ensemble is able to maintain high accuracy on more challenging, multi-species datasets, demonstrating robustness and generalization beyond conventional single-model approaches. This comparison highlights that, while previous methods perform well on controlled datasets, our stacking ensemble provides a versatile and reliable framework for multi-species plant disease classification, making it well-suited for practical applications in precision agriculture.

Table 12: Comparing our ensemble with related works.

Ref	**Year**	**Approche / Models**	**Dataset / Species**	**Reported Accuracy**
[20] A. Kumar & D. Singh	2024	Comparative study: CNN vs classical ML vs hybrid	Various crops	CNN ~97%, Hybrid ~91.7%
[21] X. Liang et al.	2025	Hybrid CNN + Vision Transformer	Apple, Corn	Apple 99.24%, Corn ~98%
[22] S. Sharma & V. Singh	2023	Deep learning: CNN & hybrid approaches	Various crops	95–99% (varies by study)
[23] R. Gupta & S. Kumar	2024	CNN, classical ML, hybrid	Leaf disease datasets	95–99%
RBF-SVM meta-learning	2025	CNN, SVM, RFC	Apple, Tomato	98% Multi-species

Source: Authors, (2026).

VI. CONCLUSIONS

This study proposes a meta-learning–based stacking ensemble for multi-species crop leaf pathology recognition, integrating Convolutional Neural Networks (CNN), Support Vector Machines (SVM), and Random Forest Classifiers (RFC) as heterogeneous base learners. Each model was trained independently using its own feature representation: the CNN learned hierarchical visual features directly from images, while the SVM and RFC were trained on raw pixel vectors obtained from resized and flattened images. Their individual predictions were then combined by a meta-level SVM, which learned to exploit complementary strengths across the three models. This stacking strategy resulted in a significant improvement in overall classification performance compared to any single model alone.

Experimental results demonstrated that while individual models achieved accuracies between 85 % and 89 %, the stacking ensemble reached near-perfect performance with macro and weighted F1-scores of 0.975, highlighting its ability to generalize across multiple crop species, including apple and tomato. These results confirm that combining diverse learning paradigms via meta-learning effectively addresses the limitations of single-model approaches, particularly in multi-class and multi-species scenarios. The proposed framework provides a robust, accurate, and generalizable tool for automated crop disease detection, which can support precision agriculture and early disease management. Future work will focus on extending this approach to additional crop species, incorporating field-acquired images under variable conditions, and exploring lightweight architectures to enable real-time deployment on mobile or edge devices.

VII. AUTHOR’S CONTRIBUTION

Conceptualization: Ryma Guefrouchi and Siham Amrouch.
Methodology: Ryma Guefrouchi and Siham Amrouch.
Investigation: Ryma Guefrouchi and Siham Amrouch.
Discussion of results: Ryma Guefrouchi and Siham Amrouch.
Writing – Original Draft: Ryma Guefrouchi and Siham Amrouch.
Writing – Review and Editing: Ryma Guefrouchi and Siham Amrouch.
Resources: Ryma Guefrouchi and Siham Amrouch.
Supervision: Ryma Guefrouchi and Siham Amrouch
Approval of the final text: Ryma Guefrouchi and Siham Amrouch.

VIII. REFERENCES

[1] K. P. Ferentinos, "Deep learning models for plant disease detection and diagnosis," Comput. Electron. Agric., vol. 145, pp. 311–318, 2018.

[2] S. Sladojevic et al., "Deep neural networks based recognition of plant diseases by leaf image classification," Comput. Intell. Neurosci., Art. no. 3289801, 2016.

[3] P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," Front. Plant Sci., vol. 7, p. 1419, 2016.

[4] E. C. Too et al., "A comparative study of fine-tuning deep learning models for plant disease identification," Comput. Electron. Agric., vol. 161, pp. 272–279, 2019.

[5] S. Brahimi, D. Boukhalfa, and A. Moussaoui, "Deep learning for tomato diseases: Classification and symptoms visualization," Appl. Artif. Intell., vol. 31, no. 4, pp. 299–315, 2017.

[6] A. Ramcharan et al., "Deep learning for image-based cassava disease detection," Front. Plant Sci., vol. 8, p. 1852, 2017.

[7] V. Singh and R. K. Sharma, "Plant leaf disease detection using feature extraction and machine learning techniques: A review," Int. J. Adv. Res. Comput. Sci., vol. 8, no. 5, pp. 78–84, 2017.

[8] L. Zhang, Y. Yin, and F. Wang, "Recognition of plant leaf diseases based on support vector machine and convolutional neural network," Int. J. Agric. Biol. Eng., vol. 12, no. 2, pp. 76–84, 2019.

[9] A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: A survey," Comput. Electron. Agric., vol. 147, pp. 70–90, 2018.

[10] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, 2015. doi: 10.1038/nature14539.

[11] S. Wang, J. Zhang, and Y. Liu, "Ensemble deep learning: A review," Neurocomputing, vol. 439, pp. 54–68, 2021.

[12] A. Sharma and R. C. Poonia, "Recent advances in automated plant leaf disease detection using machine learning," Mater. Today: Proc., vol. 56, pp. 1234–1240, 2022.

[13] V. Singh and R. Kaur, "A review of machine learning and deep learning methods for plant disease detection," Comput. Electron. Agric., vol. 200, p. 107233, 2022.

[14] J, J., & R, B. (2021). Detection of plant leaf diseases using recent progress in Deep Learning-Based identification techniques. ITEGAM-JETIA, 7(30), 29-36. https://doi.org/10.5935/jetia.v7i30.768
H. Li, J. Wang, and X. Chen, "EnConv: Enhanced CNN framework for plant leaf disease classification," Comput. Electron. Agric., vol. 210, p. 107850, 2024.

[15] Y. Zhang, L. Xu, and H. Zhao, "Multi-head CNN for robust plant disease detection," Neural Comput. Appl., vol. 35, pp. 10875–10890, 2023.

[16] S. Patel and P. Kumar, "Ensemble model for grape leaf disease detection using CNN feature extractors and random forest," Comput. Electron. Agric., vol. 210, p. 107805, 2024.

[17] X. Zhou, M. Li, and F. Yang, "A robust ensemble model for plant disease detection using deep learning architectures," Agriculture, vol. 15, p. 159, 2025.

[18] P. Li, J. Chen, and L. Huang, "A hybrid framework for plant leaf disease detection using CNN and vision transformer," Artif. Intell. Rev., vol. 58, pp. 4123–4140, 2025.

[19] Q. Wang, Y. Zhang, and S. Liu, "Systematic review of deep learning techniques for plant disease detection," Appl. Sci., vol. 14, p. 4512, 2024.

[20] A. Kumar and D. Singh, "Comparative study of CNN, classical ML, and hybrid methods for plant disease classification," J. Big Data, vol. 11, p. 98, 2024.

[21] X. Liang, Y. Chen, and W. Zhao, "Hybrid CNN-vision transformer model for plant leaf disease detection," J. Intell. Fuzzy Syst., vol. 48, pp. 4123–4135, 2025.

[22] S. Sharma and V. Singh, "Plant Disease Detection Using Deep Learning: Recent Trends and Challenges," Computers and Electronics in Agriculture, vol. 208, p. 107850, 2023.

[23] R. Gupta and S. Kumar, "Performance Comparison of CNN, Machine Learning, and Hybrid Methods for Leaf Disease Classification," Neural Computing and Applications, vol. 36, pp. 15021–15038, 2024.

[24] PlantVillage Dataset.” [Online]. Available: https://www.kaggle.com/datasets/emmarex/plantdisease.