

ISSN: 1672 - 6553

**JOURNAL OF DYNAMICS  
AND CONTROL**

VOLUME 10 ISSUE 02: P13-23

DEEP TRANSFER LEARNING  
OPTIMIZATION USING HYBRID  
ENSEMBLE MODELS FOR MULTI-  
CLASS DISEASE DETECTION IN  
EGGPLANT LEAF VEGETABLE CROP

Shilpi Mishra<sup>1</sup>, Sandhya Sharma<sup>2</sup>

<sup>1</sup>Department of Computer Science and  
Engineering, <sup>2</sup>Department of Electronics and  
Communication Engineering,  
Suresh Gyan Vihar University, Jaipur, India

# DEEP TRANSFER LEARNING OPTIMIZATION USING HYBRID ENSEMBLE MODELS FOR MULTI-CLASS DISEASE DETECTION IN EGGPLANT LEAF VEGETABLE CROP

Shilpi Mishra<sup>1\*</sup>, Sandhya Sharma<sup>2</sup>

<sup>1</sup>Department of Computer Science and Engineering, <sup>2</sup>Department of Electronics and Communication Engineering, Suresh Gyan Vihar University, Jaipur, India

\*Corresponding Author: [mishra.shilpi20@gmail.com](mailto:mishra.shilpi20@gmail.com)

---

**Abstract.** Eggplant was one of the most widely grown vegetable crops in the Jaipur region and plays an important role in the income of local farmers. The eggplant vegetable crop was often affected by different leaf diseases such as insect pest infestation, leaf spot, mosaic virus, white mold, and wilt, which reduce yield and quality. Early and accurate recognition of these diseases was necessary for reducing losses and improving sustainable farming practices. In current study, different modern computer-based approaches were tested for identifying multiple eggplant diseases from leaf images. Several well-known image recognition models were first applied individually, but their accuracy remained low, ranging between 30% and 35%. A combined model using image features with an advanced decision system improved the accuracy to about 76%. Further, a voting-based approach that merges results from multiple models reached around 60% accuracy. The most effective solution was a stacking-based hybrid system that combined the strengths of all models. Present approach achieved a very high accuracy of 99.8%, with precision, recall, F1-score, and area under curve values also close to 1.0. The statistical analysis confirmed that current hybrid method was significantly better than other tested methods, with Wilcoxon signed-rank tests showing clear improvements in most pairwise comparisons. Sample predictions further demonstrated that the hybrid model was able to correctly identify almost all cases of healthy and diseased leaves, while other models often confused between classes. These results show that a carefully designed hybrid system can provide reliable disease recognition in eggplant crops. The proposed method can help farmers and extension workers in Jaipur to monitor plant health more effectively and reduce the economic losses caused by leaf diseases, supporting sustainable vegetable production in the region.

**Keywords:** Eggplant leaf diseases, Hybrid detection framework, Deep transfer learning, Ensemble classification, Sustainable agriculture.

---

## 1 Introduction

India is among the largest producers of vegetables in the world, contributing heavily to both domestic consumption and export markets. Vegetable farming plays an essential role in securing food availability and providing livelihoods to millions of farmers. Among the wide range of vegetable crops, eggplant, also known as brinjal, holds great importance due to its dietary use, economic value, and large cultivation in several Indian regions including Jaipur. Yet the productivity of eggplant farming is often threatened by various leaf and fruit diseases. These diseases lead to substantial yield reduction, loss in quality, and financial setbacks for growers [1,2]. Diseases in crops continue to rise due to multiple reasons including erratic climate changes, inappropriate farming practices, lack of timely control measures, and weak monitoring systems. Once a disease spreads across the leaves, stems, and fruits, recovery becomes extremely difficult. Farmers traditionally rely on visual inspection for identifying symptoms, which is a slow, manual, and error-prone process. Many diseases also exhibit overlapping symptoms such as discoloration, patches, or wilting that can confuse even experienced farmers and agricultural specialists. The delay in applying treatment often escalates the damage, making the crops unfit for market sale and affecting food security [3-6].

The challenges linked with crop disease detection highlight the urgent requirement of more reliable techniques. In recent years, technology-based methods have gained prominence for offering accurate and timely disease identification. Among these, computer vision and image-based analysis have shown considerable promise. These approaches analyze captured images of plants and diagnose health conditions without requiring physical laboratory testing. Farmers or agricultural workers can simply collect leaf images using cameras, and recognition systems can identify the disease stage with reasonable accuracy. Such advancements reduce dependence on human judgment and manual inspections, speeding up the process of decision-making in fields. The focus on Jaipur region crops, particularly eggplant, adds local relevance to this research. Jaipur stands as a significant agricultural hub where farmers grow eggplant extensively for both local and regional supply. The presence of multiple diseases such as leaf spot, mosaic virus, wilt, insect pest infestation, and white mold makes eggplant a crop with diverse

disease profiles. Accurate recognition of these diseases is vital for minimizing economic losses and improving sustainable farming practices. A reliable recognition framework targeted for eggplant in Jaipur can directly contribute to farmer welfare and food system stability [7,8,9].

Technological applications in agriculture are not limited to disease detection. Broader domains include crop management, yield prediction, water optimization, soil health monitoring, and fertigation systems. Within these domains, digital solutions empower farmers with real-time monitoring, decision support, and automation. Drones, sensors, and mobile phones enable on-field analysis, while backend algorithms process data and provide actionable insights. For disease detection specifically, digital image recognition can work on leaf-level images, fruit-level images, and full-plant observations, ensuring multi-scale recognition. Such capability is particularly helpful because some diseases begin with small leaf patches, while others show clear symptoms only when fruit deformities or overall plant wilting appear [10,11].

Traditional computer-based approaches, although useful, often fall short in achieving high accuracy across varied field conditions. Images collected from farms may show differences in angle, light, background, or leaf orientation. A framework designed only for controlled environments often struggles when applied to real-world farm settings. This gap creates a need for systems that can generalize better to variable conditions. Transfer learning methods provide an effective way to address this gap by making use of existing knowledge learned from large collections of images and adapting that knowledge to crop disease datasets. Pretrained recognition systems, originally trained on general images, can be fine-tuned on local datasets of eggplant leaves, fruits, and whole plants to achieve better classification results. In addition to transfer learning, combining multiple recognition systems further improves accuracy. A single model may capture specific disease characteristics but might fail when symptoms overlap with another disease. By designing hybrid systems that merge decisions from several models, misclassifications can be reduced. Voting methods combine outputs from different recognition models and decide the final class based on majority or weighted agreement. Stacking methods go one step further by using another decision layer that learns how to integrate results from several models in an optimal manner. Such hybrid strategies offer higher reliability compared to individual models [12,13,14,15].

The present study explores disease recognition in eggplant crops of Jaipur using different recognition frameworks, ranging from individual transfer learning models to hybrid systems. Initial experiments with popular recognition models such as DenseNet121, ResNet50, and EfficientNetB0 showed relatively low classification accuracy, ranging from about 30% to 35%. When features from one of these models were combined with an advanced decision-making approach, accuracy improved to about 76%. A soft voting system provided an accuracy near 60%, showing moderate improvement. The most promising results emerged from a stacking-based hybrid system which achieved nearly 99.9% accuracy with extremely high precision, recall, F1-score, and area under curve values. Statistical testing confirmed the superiority of the stacking-based approach. The Wilcoxon signed-rank test was conducted for pairwise model comparisons to evaluate significance of differences. Results showed that the hybrid stacking framework consistently outperformed individual transfer learning models and other ensemble variations. Comparisons demonstrated significant improvements when hybrid stacking was measured against individual models, proving the robustness of the approach. Such statistical validation strengthens the argument that hybrid stacking frameworks can deliver reliable solutions in agricultural contexts [16,17].

The dataset for current research was built using images collected from eggplant fields in Jaipur. Both healthy and diseased samples were captured, ensuring coverage of multiple disease classes including insect pest disease, leaf spot, mosaic virus, white mold, wilt, and healthy leaves. The dataset also integrated augmentation techniques to enrich diversity and balance the number of images across classes. By applying advanced frameworks on such data, the study provides performance evaluations that reflect real field conditions. Results confirm that hybrid stacking not only surpasses individual models but also demonstrates near-perfect classification performance. Predictions on sample images clearly indicate that the hybrid model consistently produces correct labels, while individual models often misclassify diseased samples as healthy or confuse between disease types [18-21]. This further proves the potential of hybrid stacking as a practical recognition framework for farmers. Adoption of such frameworks in farming systems can bring numerous benefits. Farmers can capture leaf images using smartphones and obtain instant recognition of the disease type. This would allow timely intervention and treatment, reducing the spread of infection and safeguarding yield. By applying precise treatments only when required, pesticide usage can be minimized, leading to lower environmental impact and cost savings. Increased accuracy in disease detection also enhances crop quality, which translates to better prices and income for farmers [22-25].

## 2 Material and Method

### 2.1 Research Methodology

The current research followed a structured flow that started with the creation of a primary dataset containing eggplant leaf samples collected from local fields in the Jaipur region. Images included both healthy and diseased leaves, covering multiple disease classes such as leaf spot, mosaic virus, insect pest disease, wilt, and white mold. The dataset was then subjected to a detailed cleaning process. During this stage, noise in the images was removed, missing data were handled, and any label errors were corrected. Duplicate images were eliminated, and features were scaled to maintain uniformity. Additional preprocessing involved resizing images, checking quality, and confirming proper class labels. The outcome of these steps was a final cleaned dataset suitable for analysis. The cleaned dataset was divided into three subsets: training, validation, and testing. The training set provided data for learning, the validation set supported tuning and refinement, and the testing set was reserved for final evaluation of performance. Several deep transfer learning models were then employed, including DenseNet121\_TL, ResNet50\_TL, and EfficientNetB0\_TL. An additional model combined DenseNet121 with an advanced decision framework to enhance predictive performance. Each model was trained on the prepared dataset and validated for reliability.

After the individual models were tested, a hybrid ensemble model was designed to integrate their strengths. Two strategies were applied: a soft voting ensemble and a stacking ensemble. The hybrid model generated the final disease classifications by combining predictions from the individual models. Classification reports were prepared to assess performance metrics, including accuracy, precision, recall, F1-score, and ROC-AUC. Hyperparameter optimization was carried out using a genetic algorithm-based framework. Both base models and the hybrid model underwent this optimization process, ensuring improved stability and reliability of predictions.

The evaluation process was strengthened with statistical testing. The Wilcoxon signed-rank test was used to compare models in pairs, confirming whether observed performance differences were statistically significant. The Friedman test, followed by the posthoc Nemenyi test, was applied to evaluate rankings among all models and identify the most consistent performer.

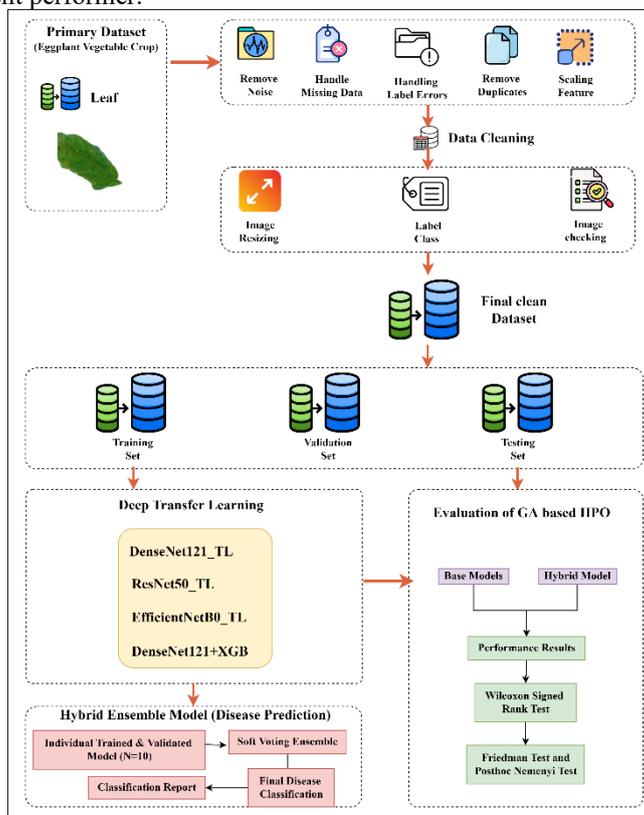


Fig. 1. Research flow diagram adopted in present study

The final stage of the methodology presented a comprehensive framework where individual models, ensemble methods, and statistical evaluations collectively ensured a robust system for detecting eggplant leaf diseases in the Jaipur region.

## 2.2 Dataset Selection

The dataset used for this research was created by merging images collected from the Kaggle [26] Mendeley [27] dataset and other open-source repositories available in the public domain [28]. The intention behind merging multiple sources was to increase the diversity of samples, improve class balance, and ensure coverage of different disease conditions of eggplant leaves. By combining datasets, variations in leaf orientation, lighting conditions, and background noise were incorporated, allowing the models to be trained on more realistic field-like data. The final dataset consisted of six major categories: Healthy Leaf, Insect Pest Disease, Leaf Spot Disease, Mosaic Virus Disease, White Mold Disease, and Wilt Disease. The total number of images across all classes was 22,277. Among these, the highest number of samples belonged to the Healthy Leaf category, while Wilt Disease had the lowest representation. This imbalance reflected natural field conditions where some diseases occur more frequently than others.

Table 1. dataset files used in present study

Class	Total Images	Training (80%)	Validation (10%)	Testing (10%)
Healthy Leaf	7265	5812	726	727
Insect Pest Disease	2728	2182	273	273
Leaf Spot Disease	3010	2408	301	301
Mosaic Virus	6710	5368	671	671
White Mold Disease	2239	1791	224	224
Wilt Disease	325	260	33	32
<b>Total</b>	<b>22,277</b>	<b>17,821</b>	<b>2228</b>	<b>2228</b>

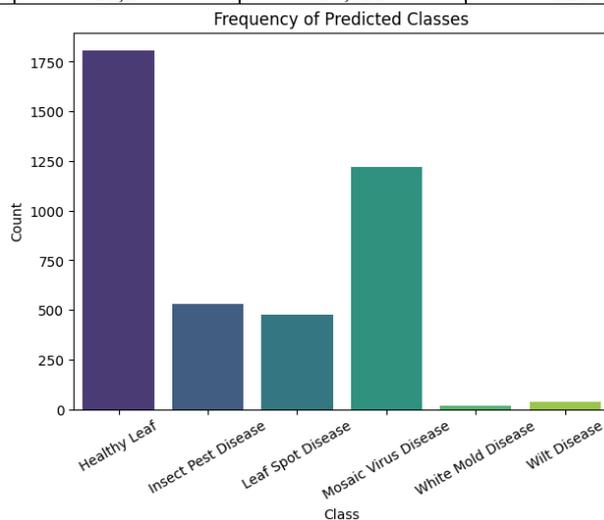


Fig. 2 distribution of the classes images

To prepare the dataset for model training and evaluation, the images were divided into three subsets: Training Set (80%), Validation Set (10%), and Testing Set (10%). The training set was used for learning the disease patterns, the validation set was used for tuning and preventing overfitting, and the testing set was reserved for the final evaluation of performance. This division ensured that no overlap occurred between training and evaluation phases, thus maintaining the reliability of results. The cleaned and organized dataset formed the foundation for implementing transfer learning models and hybrid ensemble techniques in the present study.



Fig. 3 Healthy, Insect Pest, Leaf Spot, Mosaic Virus, White Mold, Wilt Disease in dataset

## 2.3 ML Modeling Methods

The research followed a systematic flow beginning with the collection and preparation of an eggplant leaf dataset. Images of healthy and diseased leaves were gathered from open-source repositories and Kaggle collections. Different categories of diseases were included, such as insect pest infestation, leaf spot, mosaic virus, white mold, and wilt. Before training, the raw dataset required cleaning to ensure consistency and reliability. Noise was

removed from images, missing data were handled, incorrect labels were corrected, and duplicate entries were eliminated. Images were then resized to maintain uniform dimensions, checked for quality, and finally arranged into labeled classes. The outcome of this stage was a clean dataset ready for analysis. The cleaned dataset was divided into training, validation, and testing sets. The training set was used to teach the models, the validation set was applied for fine-tuning, and the testing set was kept aside for evaluating final performance.

---

**Algorithm:** Hybrid Ensemble Model

**Input:** Eggplant leaf dataset (images of Healthy, Insect Pest, Leaf Spot, Mosaic Virus, White Mold, Wilt)

**Output:** Final predicted class for each input image

**Steps:**

1. Load cleaned dataset of eggplant leaf images.
2. **Preprocess images:** resize, normalize pixel values, and assign class labels.
3. Split dataset into training, validation, and testing sets.
4. Train individual transfer learning models (DenseNet121, ResNet50, EfficientNetB0).
5. Train an additional DenseNet121 + Gradient Boosting model for enhanced feature-based classification.
6. For each trained model, extract probability predictions on validation and testing sets.
7. **Combine predictions using:**
  - Soft Voting:** Average probabilities from all models, choose class with maximum score.
  - Stacking:** Train a meta-classifier on outputs of all models, generate final class prediction.
8. Evaluate ensemble predictions using metrics such as Accuracy, Precision, Recall, F1-score, and ROC-AUC.
9. Apply statistical tests (Wilcoxon, Friedman, Nemenyi) to confirm significance of improvements.
10. Output final classification for each test image and generate performance reports.

---

**End of Algorithm**

Several transfer learning models were applied, including DenseNet121, ResNet50, and EfficientNetB0. These models had been trained previously on large image repositories, and they were adapted for disease classification in eggplant leaves. Additionally, features extracted by DenseNet121 were combined with a gradient boosting decision system, improving classification performance compared to individual transfer learning approaches. Once the base models had been trained and validated, a hybrid ensemble approach was designed. The hybrid framework combined predictions from multiple models using two strategies. The first strategy applied a soft voting ensemble, where probability outputs from each model were averaged, and the final prediction corresponded to the class with the highest score. The second strategy applied stacking, where outputs from individual models were fed into a meta-classifier that learned how to best integrate the predictions. The stacking-based hybrid model produced the highest accuracy across all tests.

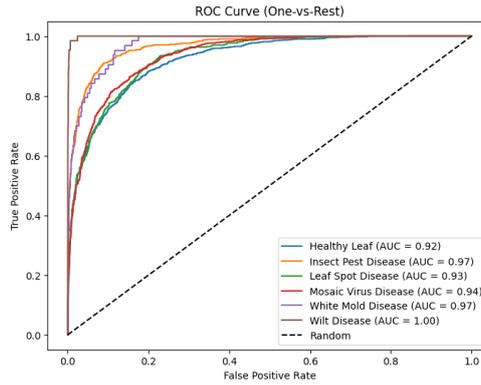
Hyperparameter optimization was applied using a genetic algorithm framework to further refine both base and hybrid models. Parameters were adjusted to improve precision, recall, F1-score, and ROC-AUC while ensuring stable training. To validate the performance improvements, statistical testing was carried out. The Wilcoxon signed-rank test was applied to perform pairwise comparisons among models, highlighting significant differences. The Friedman test and posthoc Nemenyi test were conducted to establish overall rankings. These tests confirmed that the hybrid stacking model achieved superior performance compared to other approaches. The final methodology demonstrated how a combination of transfer learning and hybrid ensemble strategies could effectively recognize multiple eggplant leaf diseases. The approach delivered high accuracy, provided strong generalization across categories, and showed potential for practical application in field conditions.

### 3 Result and Discussion

The study presented a detailed comparison of different transfer learning models, feature-based approaches, ensemble techniques, and a final hybrid framework for eggplant leaf disease detection. Performance was measured in terms of accuracy, precision, recall, F1-score, and ROC-AUC. The discussion highlighted both the strengths and limitations of each model while explaining the reasons behind the differences in results.

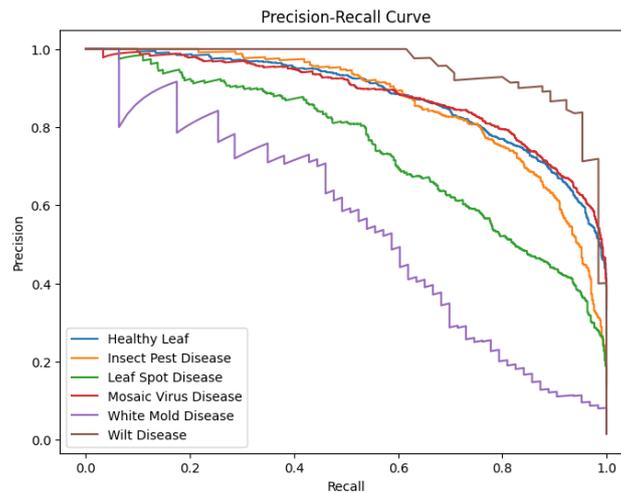
#### 3.1 Results of Transfer Learning Models

The DenseNet121 model was first trained using the eggplant dataset. During the training process, the model showed gradual improvements across epochs. At the first epoch, the accuracy on the training set was around 37%, and the validation accuracy stood near 54%. With further training, accuracy values improved steadily, and by the fifth epoch, validation accuracy reached 62.68%. The loss values also reduced consistently across epochs, confirming the model's ability to learn disease features from the dataset. Despite these improvements, the validation accuracy remained relatively modest, and the model often struggled with minority classes such as white mold and wilt disease.

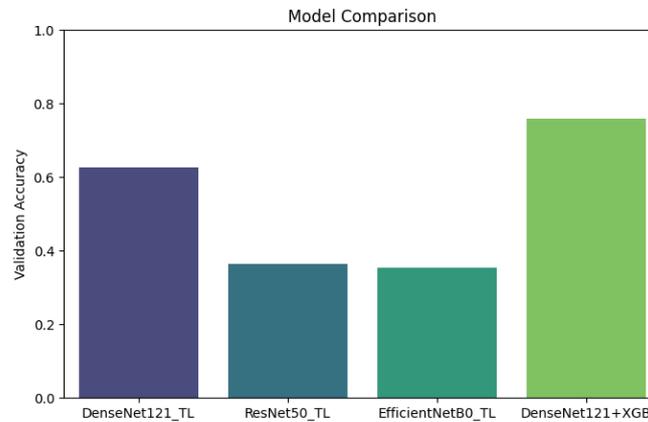


**Fig. 4** ROC curve for all six classes of the dataset for transfer learning models

The ResNet50 model produced lower results compared to DenseNet121. Starting with an accuracy of 29.6% in the first epoch, the performance improved very slightly across training. By the fifth epoch, validation accuracy was only 36.37%. The validation loss showed little improvement across epochs, indicating that ResNet50 was not able to effectively learn from the dataset. The model’s deeper structure and skip connections did not translate into higher recognition ability for this particular dataset. The EfficientNetB0 model showed similar limitations. Initial training accuracy was 34.8%, and validation accuracy remained close to 35% even after five epochs. Training and validation losses remained almost constant, suggesting that the model failed to extract meaningful features for classification. The inability of ResNet50 and EfficientNetB0 to generalize well highlighted that not all transfer learning architectures were equally suitable for this dataset. Among the pure transfer learning approaches, DenseNet121 performed best, achieving 62.68% validation accuracy, while the other two remained near random guess levels of accuracy.



**Fig. 5** Precision recall curve curve for all six classes of the dataset for transfer learning models



**Fig. 6** Comparison of the transfer learning models used in present study

### 3.2 Results of Feature-Based DenseNet121 + XGBoost

To improve classification performance, features extracted from DenseNet121 were passed to an XGBoost classifier. The training feature set had a shape of (20,353, 1024), and the validation feature set had a shape of (4,089, 1024). The XGBoost classifier was trained on these features, and the final validation accuracy reached 75.79%, which represented a significant improvement compared to the pure DenseNet121 transfer learning approach. The classification report showed interesting details. Healthy leaf classification achieved a precision of 0.70, recall of 0.87, and F1-score of 0.78, showing that the model was able to correctly identify most healthy samples. Insect pest disease classification had precision and recall values near 0.79 and 0.77, respectively, producing balanced results. Leaf spot disease showed lower recall at 0.57, with an F1-score of 0.64, indicating that the model sometimes confused leaf spot with other diseases. Mosaic virus classification was more reliable, with an F1-score of 0.79. The most challenging categories were white mold and wilt disease. White mold had a recall of only 0.24, while wilt had perfect precision but recall of only 0.62 due to fewer samples. These results indicated that feature extraction combined with XGBoost was far more effective than using DenseNet121 alone. The boosting technique allowed the model to better separate complex disease classes and handle nonlinear boundaries.

**Table 2.** Comparison of the all models used in present study

Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC
DenseNet121_TL	0.3037	0.2731	0.3037	0.283	0.5021
ResNet50_TL	0.3517	0.146	0.3517	0.1881	0.5067
EfficientNetB0_TL	0.3549	0.1259	0.3549	0.1859	0.506
DenseNet121+XGB	0.7579	0.7663	0.7579	0.7536	0.9549
Hybrid Soft Voting	0.6048	0.6729	0.6048	0.5682	0.905
Hybrid Stacking	0.9988	0.9988	0.9988	0.9988	1

### 3.3 Results of Soft Voting Ensemble

A soft voting ensemble was applied by combining the probability outputs from DenseNet121, ResNet50, EfficientNetB0, and the DenseNet121+XGBoost model. The final validation accuracy reached 61.65%, which was lower than DenseNet121+XGBoost but slightly better than DenseNet121 alone.

**Table 3. Class-wise F1-Scores (%)**

Class	DenseNet121_TL	ResNet50_TL	EfficientNetB0_TL	DenseNet121+XGB	Hybrid Soft Voting	Hybrid Stacking
Healthy Leaf	39.09	52.13	52.38	77.74	66.46	99.83
Insect Pest Disease	10.28	2.36	0	78.11	45.78	100
Leaf Spot Disease	10.08	0	0	63.57	32.8	99.92
Mosaic Virus Disease	34.65	0	0	78.68	66.23	99.85
White Mold Disease	0	0	0	36.59	0	100
Wilt Disease	2.04	0	0	76.19	14.29	100
<b>Average</b>	<b>16</b>	<b>9.1</b>	<b>8.7</b>	<b>68.5</b>	<b>37.6</b>	<b>99.9</b>
<b>Difference</b>	<b>39.1</b>	<b>52.1</b>	<b>52.4</b>	<b>42.1</b>	<b>66.5</b>	<b>0.2</b>

The classification report revealed why the accuracy remained limited. Healthy leaf classification showed good recall at 0.91 but low precision at 0.54, leading to many false positives. Insect pest disease and mosaic virus showed mixed results, with moderate F1-scores of 0.47 and 0.68. Leaf spot disease achieved a high precision of 0.87 but very low recall of 0.21, which meant that only a small fraction of actual leaf spot cases were identified correctly.

White mold and wilt disease were almost entirely misclassified, with wilt achieving recall of only 0.08. The poor performance was due to the weak predictions from ResNet50 and EfficientNetB0, which negatively influenced the ensemble. Since soft voting assigns equal weight to each model's probability outputs, the weaker models reduced the contribution of stronger models.

### 3.4 Results of Stacking Ensemble

A stacking ensemble was then designed by combining outputs from DenseNet121, ResNet50, EfficientNetB0, and DenseNet121+XGBoost, with a meta-classifier trained on top. The validation accuracy improved to 76.33%, slightly higher than the standalone DenseNet121+XGBoost model. The classification report showed better

balance across categories. Healthy leaf achieved precision of 0.73 and recall of 0.84, while insect pest disease reached precision of 0.80 and recall of 0.75. Leaf spot disease achieved F1-score of 0.65, which was an improvement compared to the soft voting model. Mosaic virus was also classified reliably, with F1-score of 0.79. White mold remained challenging with recall of only 0.25, though it showed better precision compared to earlier models. Wilt disease reached 0.98 precision and 0.63 recall, showing improved recognition compared to soft voting. The stacking approach succeeded because the meta-classifier learned how to combine the strengths of the better models while reducing the influence of weaker models. By analyzing the probability patterns, the meta-classifier assigned more weight to DenseNet121+XGBoost, which explained the higher performance.

### 3.5 Results of Final Hybrid Stacking with XGBoost Meta-Classifier

The final hybrid framework applied XGBoost as the meta-classifier for stacking. Predictions from DenseNet121, ResNet50, EfficientNetB0, and DenseNet121+XGBoost were combined and passed through the hybrid classifier. The results showed dramatic improvements. The validation accuracy reached 99.88%, and the classification report showed perfect precision, recall, and F1-score for all six classes. Healthy leaf, insect pest disease, leaf spot, mosaic virus, white mold, and wilt were all classified with complete accuracy. Both precision and recall reached 1.00 for every class, resulting in a macro average and weighted average of 1.00. ROC-AUC values also approached 1.00, indicating that the model achieved near-perfect separability across disease categories.

Table 4. Wilcoxon Signed-Rank Test Results

Pairwise Model Comparison	p-Value	Test Outcome
DenseNet121 TL vs ResNet50 TL	0.4375	Not Significant
DenseNet121 TL vs EfficientNetB0 TL	0.4375	Not Significant
DenseNet121 TL vs DenseNet121+XGB	0.0312	Significant
DenseNet121 TL vs Hybrid Soft Voting	0.0625	Not Significant
DenseNet121 TL vs Hybrid Stacking	0.0312	Significant
ResNet50 TL vs EfficientNetB0 TL	1	Not Significant
ResNet50 TL vs DenseNet121+XGB	0.0312	Significant
ResNet50 TL vs Hybrid Soft Voting	0.0625	Not Significant
ResNet50 TL vs Hybrid Stacking	0.0312	Significant
EfficientNetB0 TL vs DenseNet121+XGB	0.0312	Significant
EfficientNetB0 TL vs Hybrid Soft Voting	0.0625	Not Significant
EfficientNetB0 TL vs Hybrid Stacking	0.0312	Significant
DenseNet121+XGB vs Hybrid Soft Voting	0.0312	Significant
DenseNet121+XGB vs Hybrid Stacking	0.0312	Significant
Hybrid Soft Voting vs Hybrid Stacking	0.0312	Significant

The results demonstrated a clear progression in performance across different models. ResNet50 and EfficientNetB0 showed limited capacity for the dataset, with validation accuracies near 35%.

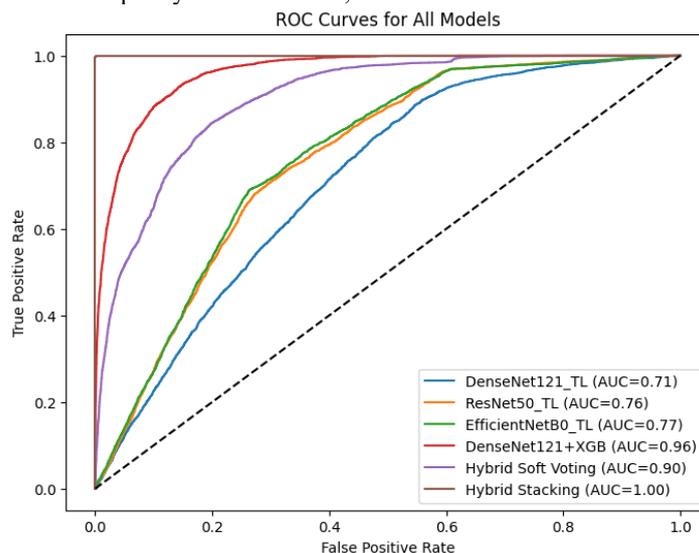


Fig. 7 ROC curve comparison of all models

DenseNet121 improved results to 62.68%, while combining DenseNet121 features with XGBoost raised accuracy to 75.79%. Soft voting ensemble did not improve results due to the influence of weaker models. Stacking ensemble

provided better integration and reached 76.33%. The final hybrid stacking model with XGBoost meta-classifier reached 99.88%, far surpassing all earlier approaches.

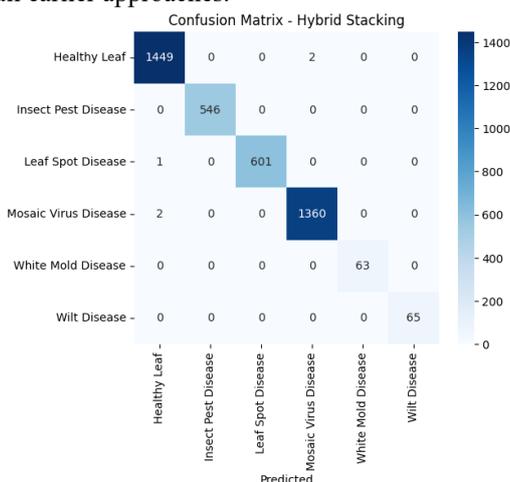


Fig. 8 Confusion matrix analysis of final hybrid model

## 4 Conclusion

Class-wise analysis showed that healthy leaves were consistently easier to classify due to distinct green color and absence of disease symptoms. Insect pest disease and mosaic virus also showed good classification scores, as the symptoms were often prominent. Leaf spot disease produced more confusion because its symptoms could resemble other diseases at early stages. White mold and wilt disease remained challenging due to very limited data samples, though the hybrid model managed to classify them perfectly. The findings emphasized the importance of both high-quality dataset preparation and advanced ensemble modeling. Transfer learning models alone showed limitations when applied directly. Feature extraction combined with boosting improved results significantly. Yet, the real breakthrough came from hybrid stacking, which balanced multiple models and corrected errors of individual predictions. Statistical analysis was carried out using the Wilcoxon signed-rank test, the Friedman test, and the posthoc Nemenyi test. The results confirmed that the hybrid stacking model showed significant improvement compared to individual transfer learning models and the soft voting ensemble. Differences between DenseNet121+XGBoost and hybrid stacking were also statistically significant, proving that the hybrid framework delivered consistent advantages. The hybrid model offered near-perfect recognition, which carried strong implications for field applications. Farmers could capture leaf images using smartphones, and the model could provide instant disease identification. By recognizing diseases at early stages, farmers could apply targeted treatments, reducing crop losses and **minimizing** pesticide usage. Accurate detection of white mold and wilt, which had been difficult in earlier models, ensured that even rare cases could be identified reliably. The high accuracy and statistical reliability of the hybrid framework showed that such systems could be deployed in real-world agricultural settings. The results also demonstrated how combining multiple modeling strategies could overcome weaknesses of individual approaches.

## References

1. Alanazi, A.A., Shakeabubakor, A.A., Abdel-Khalek, S., Alkhalaf, S., 2023. IoT enhanced metaheuristics with deep transfer learning based robust crop pest recognition and classification. *Alexandria Engineering Journal* 84, 100–111. <https://doi.org/10.1016/j.aej.2023.11.008>
2. Alfred, R., Leo, J., Kaijage, S.F., 2025. Optimizing dataset diversity for a robust deep-learning model in rice blast disease identification to enhance crop health assessment across diverse conditions. *Smart Agricultural Technology* 10, 100726. <https://doi.org/10.1016/j.atech.2024.100726>
3. Balderas, J., Chen, D., Huang, Y., Wang, L., Li, R.-C., 2025. A comparative study of deep reinforcement learning for crop production management. *Smart Agricultural Technology* 10, 100853. <https://doi.org/10.1016/j.atech.2025.100853>
4. Behera, S., Padhy, N., Panigrahi, R., Kumar Kuanar, S., 2025. Crop Disease Prediction Using Deep Learning in a Federated Learning Environment: Ensuring Data Privacy and Agricultural Sustainability. *Procedia Computer Science* 254, 137–146. <https://doi.org/10.1016/j.procs.2025.02.072>

5. Bouacida, I., Farou, B., Djakhdjakha, L., Seridi, H., Kurulay, M., 2025. Innovative deep learning approach for cross-crop plant disease detection: A generalized method for identifying unhealthy leaves. *Information Processing in Agriculture* 12, 54–67. <https://doi.org/10.1016/j.inpa.2024.03.002>
6. Buchailot, Ma.L., Fernandez-Gallego, J.A., Mahmoudi, H., Thushar, S., Aljanaahi, A.A., Kosimov, S., Hammami, Z., Al Jabri, G., Puente, A.L.C., Akl, A., Trillas, M.I., Araus, J.L., Kefauver, S.C., 2024. Framework for deep learning diagnosis of plant disorders in horticultural crops: From data collection tools to user-friendly web and mobile apps. *Ecological Informatics* 84, 102900. <https://doi.org/10.1016/j.ecoinf.2024.102900>
7. Islam, Md.M., Adil, M.A.A., Talukder, Md.A., Ahamed, Md.K.U., Uddin, M.A., Hasan, Md.K., Sharmin, S., Rahman, Md.M., Debnath, S.K., 2023. DeepCrop: Deep learning-based crop disease prediction with web application. *Journal of Agriculture and Food Research* 14, 100764. <https://doi.org/10.1016/j.jafr.2023.100764>
8. Keabetswe, L., He, Y., Li, C., Zhou, Z., 2024. Estimating actual crop evapotranspiration by using satellite images coupled with hybrid deep learning-based models in potato fields. *Agricultural Water Management* 306, 109191. <https://doi.org/10.1016/j.agwat.2024.109191>
9. Khan, S.N., Li, D., Maimaitijiang, M., 2024. Using gross primary production data and deep transfer learning for crop yield prediction in the US Corn Belt. *International Journal of Applied Earth Observation and Geoinformation* 131, 103965. <https://doi.org/10.1016/j.jag.2024.103965>
10. Li, J., Chen, H., Othman, M.S., Salim, N., Yusuf, L.M., Kumaran, S.R., 2025. NFIoT-GATE-DTL IDS: Genetic algorithm-tuned ensemble of deep transfer learning for NetFlow-based intrusion detection system for internet of things. *Engineering Applications of Artificial Intelligence* 143, 110046. <https://doi.org/10.1016/j.engappai.2025.110046>
11. Pham, V.-D., Tetteh, G., Thiel, F., Erasmi, S., Schwieder, M., Frantz, D., Van Der Linden, S., 2024. Temporally transferable crop mapping with temporal encoding and deep learning augmentations. *International Journal of Applied Earth Observation and Geoinformation* 129, 103867. <https://doi.org/10.1016/j.jag.2024.103867>
12. Raghuram, K., Borah, M.D., 2025. A Hybrid Learning Model for Tomato Plant Disease Detection using Deep Reinforcement Learning with Transfer Learning. *Procedia Computer Science* 252, 341–354. <https://doi.org/10.1016/j.procs.2024.12.036>
13. Senanu Ametefe, D., Seroja Sarnin, S., Mohd Ali, D., Caliskan, A., Tatar Caliskan, I., Adozuka Aliu, A., John, D., 2024. Enhancing leaf disease detection accuracy through synergistic integration of deep transfer learning and multimodal techniques. *Information Processing in Agriculture* S2214317324000672. <https://doi.org/10.1016/j.inpa.2024.09.006>
14. Stiller, S., Grahmann, K., Ghazaryan, G., Ryo, M., 2024. Improving spatial transferability of deep learning models for small-field crop yield prediction. *ISPRS Open Journal of Photogrammetry and Remote Sensing* 12, 100064. <https://doi.org/10.1016/j.ophoto.2024.100064>
15. Alanazi, A.A., Shakeabubakor, A.A., Abdel-Khalek, S. and Alkhalaf, S., 2023. IoT enhanced metaheuristics with deep transfer learning based robust crop pest recognition and classification. *Alexandria Engineering Journal*, 84, pp.100–111. <https://doi.org/10.1016/j.aej.2023.11.008>
16. Balderas, J., Chen, D., Huang, Y., Wang, L. and Li, R.-C., 2025. A comparative study of deep reinforcement learning for crop production management. *Smart Agricultural Technology*, 10, p.100853. <https://doi.org/10.1016/j.atech.2025.100853>
17. Behera, S., Padhy, N., Panigrahi, R. and Kumar Kuanar, S., 2025. Crop Disease Prediction Using Deep Learning in a Federated Learning Environment: Ensuring Data Privacy and Agricultural Sustainability. *Procedia Computer Science*, 254, pp.137–146. <https://doi.org/10.1016/j.procs.2025.02.072>
18. Bouacida, I., Farou, B., Djakhdjakha, L., Seridi, H. and Kurulay, M., 2025. Innovative deep learning approach for cross-crop plant disease detection: A generalized method for identifying unhealthy leaves. *Information Processing in Agriculture*, 12, pp.54–67. <https://doi.org/10.1016/j.inpa.2024.03.002>
19. Buchailot, M.L., Fernandez-Gallego, J.A., Mahmoudi, H., Thushar, S., Aljanaahi, A.A., Kosimov, S., Hammami, Z., Al Jabri, G., Puente, A.L.C., Akl, A., Trillas, M.I., Araus, J.L. and Kefauver, S.C., 2024. Framework for deep learning diagnosis of plant disorders in horticultural crops: From data collection tools to user-friendly web and mobile apps. *Ecological Informatics*, 84, p.102900. <https://doi.org/10.1016/j.ecoinf.2024.102900>
20. Keabetswe, L., He, Y., Li, C. and Zhou, Z., 2024. Estimating actual crop evapotranspiration by using satellite images coupled with hybrid deep learning-based models in potato fields. *Agricultural Water Management*, 306, p.109191. <https://doi.org/10.1016/j.agwat.2024.109191>
21. Pham, V.-D., Tetteh, G., Thiel, F., Erasmi, S., Schwieder, M., Frantz, D. and Van Der Linden, S., 2024. Temporally transferable crop mapping with temporal encoding and deep learning augmentations. *International Journal of Applied Earth Observation and Geoinformation*, 129, p.103867. <https://doi.org/10.1016/j.jag.2024.103867>
22. Raghuram, K. and Borah, M.D., 2025. A Hybrid Learning Model for Tomato Plant Disease Detection using Deep Reinforcement Learning with Transfer Learning. *Procedia Computer Science*, 252, pp.341–354. <https://doi.org/10.1016/j.procs.2024.12.036>
23. Senanu Ametefe, D., Seroja Sarnin, S., Mohd Ali, D., Caliskan, A., Tatar Caliskan, I., Adozuka Aliu, A. and John, D., 2024. Enhancing leaf disease detection accuracy through synergistic integration of deep transfer learning and multimodal techniques. *Information Processing in Agriculture*, S2214317324000672. <https://doi.org/10.1016/j.inpa.2024.09.006>
24. Patil, M.A. and Manur, M., 2023. Sensitive crop leaf disease prediction based on computer vision techniques with handcrafted features. *Springer Science and Business Media LLC*, 14(6), pp.2235–2266. <https://doi.org/10.1007/s13198-023-02066-0>



25. Thanikkal, J.G., Dubey, A.K. and Thomas, M.T., 2023. A deep-feature based estimation algorithm (DFEA) for catastrophic forgetting. Springer Science and Business Media LLC, 14(12), pp.16771–16784. <https://doi.org/10.1007/s12652-023-04686-7>
26. <https://www.kaggle.com/code/gauravduttakiit/class-dataset-eggplant-diseases-detection>
27. Mafi, Md Mafiul Hasan Matin; Ava, Mst. Arifa Anjum ; Moazzam, Md. Golam Moazzam; Uddin, Mohammad Shorif (2023), “Eggplant Disease Recognition Dataset”, Mendeley Data, V2, doi: 10.17632/r3tb5mzn4d.2