SvedbergOpen
DISSEMINATION OF KNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Access

Deep Learning-Based Identification of Arecanut Plant Disease in the Tumkur Agricultural Zone

Prashanth G K[1], Vinutha H R[2]

[1]Assistant Professor Department of MCA Siddaganga Institute of Technology Tumkur, India. Email: prashanthgk@sit.ac.in
[2]Research Scholar Department of MCA Siddaganga Institute of Technology Tumkur, India.
Email: vinuthahr.aids@shrideviengineering.org

Abstract: Arecanut cultivation in the Tumkur agricultural belt faces notable yield losses due to the prevalence of leaf diseases, underscoring the need for a reliable and automated detection system. This study proposes a deep learning driven solution utilizing transfer learning to classify arecanut leaf diseases from a carefully curated image dataset. Five convolutional neural network (CNN) architectures DenseNet-201 (88.69%), ResNet-50 (95.80%), EfficientNet-B0 (87.36%), Xception (88.41%), and ResNet-20 (94.92%) were trained and evaluated using performance metrics including accuracy, precision, recall, F1-score, and confusion matrix analysis. The ResNet-50 achieved the top accuracy of 95.80%, demonstrating superior feature extraction capabilities and consistent results across multiple disease categories. These outcomes highlight the promise of ResNet-50–based deep learning frameworks in enhancing precision agriculture and strengthening disease management in arecanut farming.

Keywords: Arecanut plant diseases, Deep learning, Transfer learning, ResNet-50, Precision agriculture, Plant disease classification

I. INTRODUCTION

Agriculture plays an essential role in sustaining India's economy, serving as the primary source of livelihood for a significant portion of the population. Among various crops, arecanut (Areca catechu L.) holds substantial economic importance, particularly in the Tumkur district of Karnataka, which is one of the leading regions for arecanut cultivation. However, the productivity of arecanut plantations is severely affected by a range of leaf diseases, including Yellow Leaf Disease, Leaf Spot, and Stem Bleeding, which directly reduce crop yield and negatively impact the income of local farmers.

Traditional disease identification in arecanut farming largely depends on manual inspection by experts or farmers. This process is often time-consuming, labor-intensive, and prone to human error, especially in large-scale agricultural settings. The lack of timely and accurate detection frequently delays corrective measures, resulting in significant economic losses and inefficient disease management. These challenges highlight the need for automated and intelligent disease detection systems capable of assisting farmers in making prompt and informed decisions.

Deep learning, particularly convolutional neural networks (CNNs), has emerged as a powerful tool for automating image-based classification tasks in agriculture. By leveraging transfer learning, pre-trained deep learning models can be adapted to identify plant diseases with high accuracy, even from limited datasets. In this study, five CNN architectures DenseNet-201, Xception, EfficientNet-B0, ResNet-50, and ResNet-20 are applied to a custom Arecanut plants image dataset collected from the Tumkur agricultural zone.

The core objective of this research is to analyse and benchmark the performance of selected deep learning models using evaluation metrics such as accuracy, precision, recall, F1-score, and confusion matrix insights. By enabling more efficient and timely disease identification, the proposed method promotes early intervention, strengthens digital agriculture initiatives, and supports sustainable farming practices, particularly in rural regions of India.

II. LITERATURE SURVEY

The adoption of deep learning (DL) in agriculture has expanded rapidly, with notable applications in crop disease diagnosis. In the arecanut plantations of the Tumkur agricultural zone, DL particularly Convolutional Neural Network (CNN) architectures has been increasingly applied for identifying plant diseases and weed species with higher precision. Existing research spans both region-specific challenges in arecanut cultivation and broader methodological developments in plant pathology.

In [1], a ResNet-based CNN model was fine-tuned using transfer learning and robust preprocessing techniques, achieving over 97% accuracy on plant disease datasets tailored to regional conditions. Similarly, [2] proposed a ResNet-based architecture using RGB images for early arecanut leaf disease detection, employing global average pooling to enhance feature abstraction particularly effective for spatially varying disease patterns. Hyperspectral imaging (HSI) integrated with machine learning classifiers, as demonstrated in [3], showed promise for detecting early-stage or low-visibility infections, which could be adapted for arecanut disease monitoring.

The prevalence of Yellow Leaf Disease (YLD) in Chikkamagaluru plantations was documented as reported by [4], providing valuable ground-truth data for dataset labeling. A comparative evaluation from [5] examined YOLOv5 and ResNet50 using the PlantDoc dataset, finding YOLOv5 more effective for object detection in dense foliage, while ResNet50 excelled in classification tasks. The biological factors contributing to YLD such as Areca Palm Velvet 1 (APV1) virus and mealybug infestation were identified by [6], offering biological context for image-based disease differentiation.

The integration of DL with IoT for real-time agricultural monitoring was examined in [7], noting opportunities and challenges such as power efficiency and interface design. A hybrid method from [8] merged CNN-based feature extraction with traditional classifiers, enhancing generalization on imbalanced datasets. According to [9], a survey of over 70 DL-based plant disease studies stressed improving dataset quality, standardization, and real-time field deployment.

In terms of feature representation, [10] demonstrated that ResNet outperforms VGGNet in capturing texture variations in diseased leaves, influencing backbone architecture choices in classification models. A lightweight CNN developed in [11] on a custom arecanut dataset achieved 88.46% accuracy, showing the feasibility of deploying resource-efficient models in rural agricultural settings. Practical deployment efforts include [12], where a CNN-based detection system was built for Android devices using Python, targeting rural usage in regions like Tumkur.

Alongside computational methods, traditional and integrated disease management practices have been explored. For instance, [13] assessed Trichoderma-based formulations for coconut disease control, which could complement DL-based detection in guiding organic treatments. A YLD severity scoring system from [14] supported consistent disease stage classification across the Malnad region, while [15] examined Trichoderma and neem cake for basal stem rot in arecanut. Lastly, [16] compiled a comprehensive catalog of arecanut diseases and their causal agents for dataset annotation and model training.

III. Methodology

This section details the methodological framework adopted for the classification of arecanut plant diseases using deep learning models. The workflow encompasses dataset acquisition, preprocessing, model selection and architecture design, exploratory data analysis, feature extraction strategy, model development, experimental setup, and performance evaluation.

A. Dataset Collection

A custom arecanut leaf image dataset was compiled from five villages in the Tumkur agricultural zone Banikupe, Hebbur-Nagavalli, Kalkere, Ranganathapura, and Rayawara. The dataset comprises RGB images of both healthy leaves and leaves exhibiting symptoms of Yellow Leaf Disease (YLD). All images were captured under natural lighting and real field conditions to preserve environmental variability, with additional samples sourced from the Kaggle repository to enhance dataset diversity.

All samples were manually labelled with assistance from agricultural experts. To enable severity-aware classification, YLD images were further annotated into three severity levels low, moderate, and high based on expert inspection and established severity scoring protocols.

Figure 1. Arecanut Dataset collected from Kaggle

The dataset collected from the Kaggle repository contains RGB images of arecanut plants across eight categories: bud borer, healthy foot, healthy leaf, healthy trunk, leaf spot, stem cracking, stem bleeding, and yellow leaf disease (YLD). The images depict clear visual variations in plant health and disease symptoms, providing representative samples for classification tasks as depicted in Figure 1.

Figure 2. Arecanut Dataset collected from five villages of Tumkur District

Figure 2 displays the images collected from five villages of Tumkur district namely Banikupe, Hebbur-Nagavalli, Kalkere, Ranganathapura, and Rayawara. Each column represents a specific category, including bud borer, healthy foot, healthy leaf, healthy trunk, leaf spot, stem cracking, stem bleeding, and yellow leaf disease. The images capture both healthy plant parts and various disease symptoms, serving as representative samples for model training and evaluation.

B. Data Preprocessing

A consistent preprocessing pipeline was applied to all images:

- Resizing to 224 × 224 pixels to meet CNN input specifications.
- Gaussian noise reduction to suppress environmental artifacts.
- Pixel intensity normalization to the [0, 1] range to stabilize gradient descent.
- Data augmentation techniques including random rotation (±20°), horizontal and vertical flips, zoom scaling (0.8× to 1.2×), and translation were applied to increase dataset diversity and reduce overfitting.

These transformations simulated common variations in leaf orientation, illumination, and positioning, enhancing model robustness.

C. Exploratory Data Analysis (EDA)

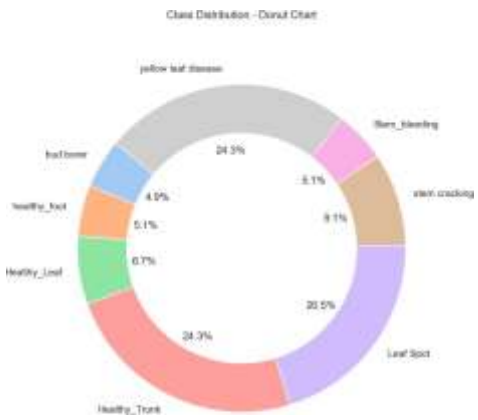

Figure 3. Distribution of various types of Classess

Figure 3 shows the class distribution of the collected arecanut leaf dataset. The majority of samples belong to Yellow Leaf Disease (24.3%) and Healthy Trunk (24.3%), followed by Leaf Spot (20.5%). Categories such as bud Borer (4.9%), Healthy Leaf (6.7%), and Healthy foot (5.1%) are under-represented, with Stem Bleeding (5.1%) and Stem cracking (9.1%) having the least samples. This imbalance was addressed through random undersampling to ensure a balanced dataset for training.

D. Feature Extraction and Engineering

The approach relied entirely on automatic feature learning via deep learning rather than handcrafted features. All images were resized to 224 × 224 × 3 and fed directly into CNN architectures such as DenseNet-201, Xception, and EfficientNet-B0.

Data augmentation methods, including reflection and translation, were employed to enhance variability and prevent overfitting. The convolutional layers inherently captured texture, color, and lesion patterns essential for distinguishing between disease types and severity levels.

E. Model Selection and Development

Five CNN architectures DenseNet-201, Xception, EfficientNet-B0, ResNet-50, and ResNet-20 were chosen for their proven success in plant disease detection and efficiency. Using transfer learning, models were initialized with ImageNet-pretrained weights, freezing the early layers to retain generic features and fine-tuning deeper layers for arecanut-specific patterns. The final layer was redesigned for multi-class classification of disease types and YLD severity.

DenseNet-201 benefits from dense connectivity, EfficientNet-B0 and Xception offers an accuracy–efficiency balance, and ResNets employ residual learning for stable training. Models were trained with the Adam optimizer, categorical cross-entropy loss, batch size 32, learning rate scheduling, and early stopping to reduce overfitting.

F. Experimental Setup

The dataset was split into 80% training and 20% testing using stratified sampling to preserve class proportions. Depending on convergence behavior, models were trained for 25–50 epochs. In some experiments, k-fold cross-validation was performed to assess robustness.

All experiments were run on Google Colab and Kaggle with GPU acceleration for faster training and evaluation.

G. Performance Evaluation

Model performance was assessed using precision, recall, and F1-score, with per-village evaluations to account for regional and environmental variation. The models classified both disease presence and YLD severity levels (low, moderate, high), supporting early intervention strategies.

Confusion matrices and classification reports provided detailed per-class accuracy metrics, while validation accuracy trends ensured training stability. Results confirmed the models' effectiveness in detecting disease and assessing severity across different geographic locations.

IV. RESULTS AND DISCUSSION

The results and discussion section presents the performance evaluation of various deep learning models for arecanut disease classification. Each model was assessed using precision, recall, F1-score, and confusion matrix analysis to ensure a comprehensive comparison. The findings highlight the strengths and limitations of individual models in classifying multiple disease categories. These results provide insights into the most suitable model for deployment in the Tumkur agricultural zone.

A. DenseNet-201

DenseNet-201 was implemented using transfer learning to classify arecanut plant diseases from field-acquired images. Its dense connectivity structure enhances feature reuse and gradient flow, enabling efficient learning of complex disease patterns.

As illustrated in Figure 4, the confusion matrix indicates high correct classifications, including 342 for Healthy Trunk and 246 for Yellow Leaf Disease. Misclassifications were minimal, with a slight confusion between Leaf Spot and Yellow Leaf Disease. Low off-diagonal values indicate strong class separation. The model demonstrates reliable performance across all categories, confirming DenseNet-201's robustness for real-world disease detection.

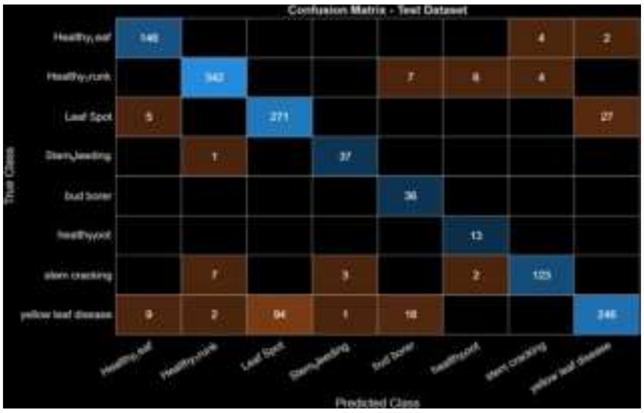

Figure 4. Confusion matrix of the DenseNet-201 model for arecanut disease classification.

B. EfficientNet-B0

EfficientNet-B0 was applied to classify arecanut plant diseases using transfer learning with a custom dataset. Its compound scaling strategy optimizes network depth, width, and resolution, enabling high accuracy while maintaining computational efficiency.

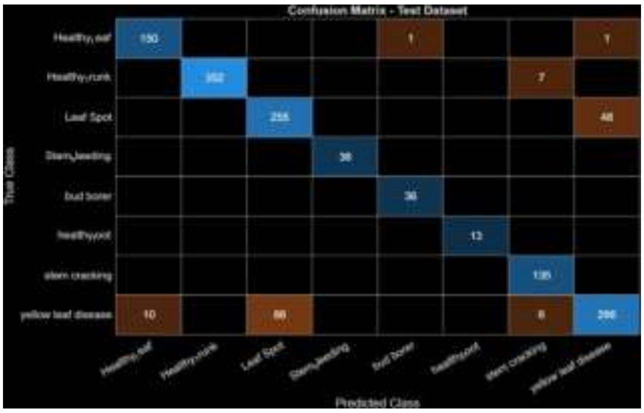

Figure 5. Confusion matrix of the EfficientNet-B0 model for arecanut disease classification.

Figure 5 illustrates how the EfficientNet-B0 model performed in classifying different arecanut diseases. The model showed strong accuracy, correctly identifying 352 cases of Healthy Trunk, 266 of Yellow Leaf Disease, and 135 of Stem Cracking. A few errors were noted, particularly between Leaf Spot and Yellow Leaf Disease, likely because their early symptoms look quite similar. Overall, the concentration of values along the diagonal of the matrix reflects the model's ability to distinguish between disease types effectively. With very few incorrect predictions, EfficientNet-B0 proves to be a dependable option for real-world application in automated arecanut disease detection.

C. Xception

The Xception model was employed for arecanut disease classification using transfer learning to leverage its depthwise separable convolutions. This design enhances feature extraction efficiency and improves model performance on complex leaf image datasets.

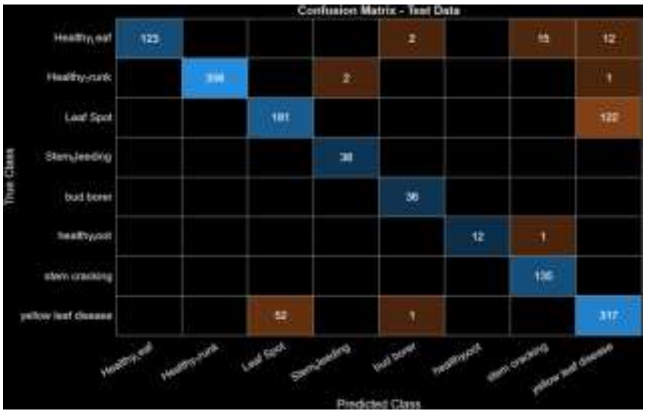

Figure 6. Confusion matrix of the Xception model for arecanut disease classification.

Figure 6 shows the confusion matrix of the Xception model for arecanut disease classification. The confusion matrix shows that Xception correctly classified most samples, including 356 for Healthy Trunk, 317 for Yellow Leaf Disease, and 135 for Stem Cracking. Some misclassifications occurred between Leaf Spot and Yellow Leaf Disease, likely due to visual similarity in early symptoms. The strong diagonal dominance demonstrates high classification reliability across all classes. Minimal off-diagonal errors indicate the model's robustness and suitability for practical agricultural disease detection in the Tumkur region.

D. ResNet-50

ResNet-50 was implemented for arecanut disease classification using transfer learning to leverage its residual connections for effective feature extraction. Its deep architecture enables accurate detection of complex disease patterns in high-resolution Arecanut plants images.

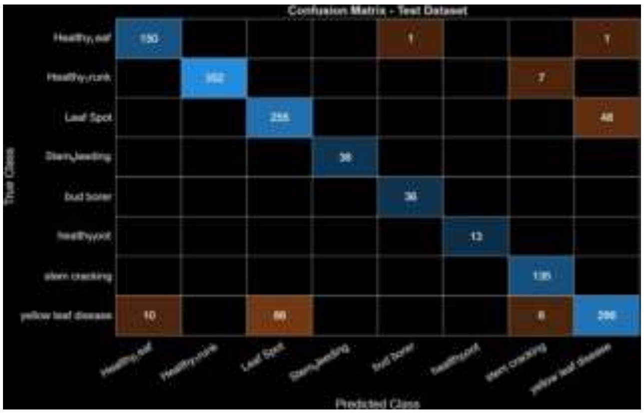

Figure 7. Confusion matrix of the ResNet-50 model for arecanut disease classification.

Figure 8. ROC curve of the ResNet-50model.

The confusion matrix, as shown in Figure 7, indicates that ResNet-50 achieved high classification accuracy, with 352 correct predictions for Healthy Trunk, 266 for Yellow Leaf Disease, and 135 for Stem Cracking. Limited misclassifications were observed between Leaf Spot and Yellow Leaf Disease, reflecting their close visual resemblance. The ROC curve, as shown in Figure 8, with an AUC of 0.98 demonstrates strong discriminative capability and low false-positive rates. These results confirm that ResNet-50 provides a reliable and well-generalized model for practical field deployment in arecanut disease detection.

E. ResNet-20

ResNet-20 was implemented for arecanut disease classification using transfer learning to exploit its residual connections for efficient feature learning. Its shallow yet effective architecture provides a balance between computational efficiency and accurate feature extraction for disease recognition.

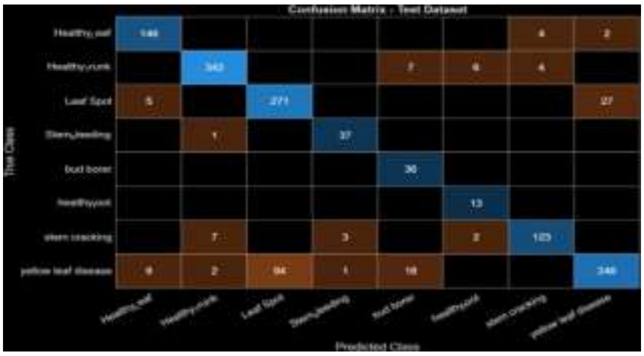

Figure 9. Confusion matrix of the ResNet-20 model for arecanut disease classification.

As illustrated in Figure 9, the confusion matrix shows that ResNet-20 accurately classified most samples, with 342 for Healthy Trunk, 271 for Leaf Spot, and 246 for Yellow Leaf Disease. Some misclassifications occurred between Leaf Spot and Yellow Leaf Disease, likely due to visual similarity. However, the overall diagonal dominance indicates strong class separation and reliable performance. The ROC curve (AUC = 0.98) further validates the model's high discriminative ability and minimal false positives, confirming ResNet-20's robustness and practical applicability for field-based disease detection.

F. Severity-Level Classification Results with ResNet-50

In addition to disease type classification, the proposed system was further evaluated for its ability to determine the severity of Yellow Leaf Disease (YLD) and Leaf Spot, categorizing each case into three levels: Low, Medium, and High. This severity assessment is essential for monitoring disease progression and enabling timely, stage-specific interventions.

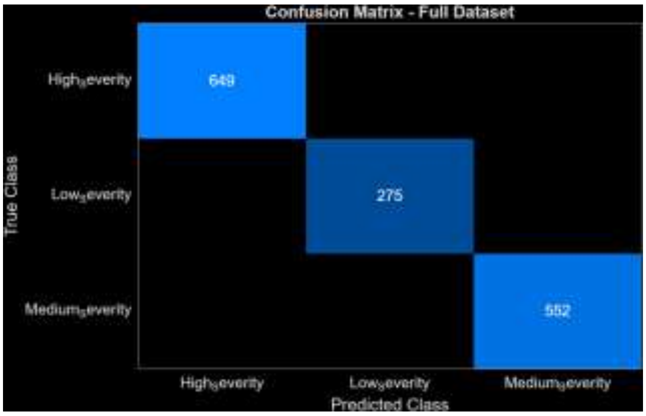

Figure 10. Confusion matrix of ResNet-50 model for YLD and Leafspot severity classification.

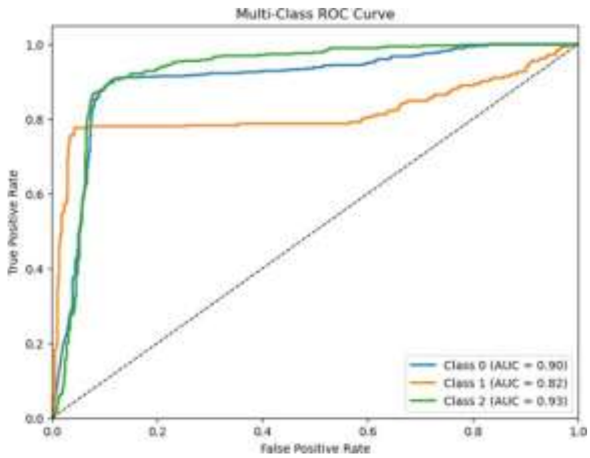

Figure 11. ROC curve of ResNet-50 model for YLD and Leafspot severity classification.

As illustrated in Figure 10, the ResNet-50 architecture demonstrated outstanding performance, accurately classifying 649 High-severity, 275 Low-severity, and 552 Medium-severity instances across both YLD and Leaf Spot datasets. The confusion matrix indicates flawless class-level predictions with no misclassifications, confirming that ResNet-50 successfully captured distinctive features for each severity category. Figure 11 presents the multi-class ROC curve, with AUC scores of 0.90 for Class 0, 0.82 for Class 1, and 0.93 for Class 2, highlighting strong discriminative capability particularly for Class 2 and consistently low false-positive rates. This high level of accuracy reflects the model's exceptional generalization capacity and resilience to intra-class variability, even when visual differences between early and advanced stages are subtle. In contrast, DenseNet-201 and EfficientNet-B0 showed minor classification overlaps between medium and high severity levels, while ResNet-50 maintained a sharper and more reliable decision boundary.

V. Comparative Analysis of Models

TABLE I. Performance Comparison of CNN Models on Arecanut Dataset

Model	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
ResNet-50	95.80	92.63	94.10	93.32
DenseNet-201	88.69	88.38	94.05	91.07
EfficientNet-B0	87.36	92.00	94.89	93.42
Xception	88.41	89.91	93.09	91.42
ResNet-20	91.93	91.05	92.81	91.99

Five CNN models DenseNet-201, EfficientNet-B0, Xception, ResNet-50, and ResNet-20 were evaluated using accuracy, precision, recall, and F1-score across different villages. The Table I presents the comparative results, with the highest values in each metric highlighted. ResNet-50 achieved the highest accuracy (95.80%) and precision (92.63%), while EfficientNet-B0 led in recall (94.89%) and F1-score (93.56%). DenseNet-201 offered a good trade-off between performance and efficiency. Notably, ResNet-20 maintained consistently balanced metrics across all four scores, making it a robust lightweight alternative.

VI. Conclusion

The research presents a deep learning driven framework for detecting and categorizing the severity of arecanut plant diseases, utilizing five CNN models DenseNet-201, EfficientNet-B0, Xception, ResNet-50, and ResNet-20. Evaluation outcomes revealed that ResNet-50 delivered the highest accuracy (95.80%) and precision (92.63%), making it an optimal choice for applications demanding high classification accuracy. EfficientNet-B0 excelled in recall (94.89%) and F1-score (93.56%), reflecting its strong ability to correctly identify disease instances. DenseNet-201 achieved a balanced compromise between predictive performance and computational demand, while ResNet-20 provided stable accuracy across all metrics, making it a dependable lightweight option for resource-limited settings. In severity classification of Yellow Leaf Disease (YLD) and Leaf Spot into low, medium, and high levels, ResNet-50 attained perfect predictions, enabling precise and timely disease management. These findings underline the system's suitability for practical precision agriculture, supporting early diagnosis, targeted intervention, and enhanced yield outcomes in arecanut farming.

References

1. M. S. P. Kumar Reddy, A. D. B. Anjana, S. A., S. K. J., M. M., and N. B. N. C., "Arecanut plant disease detection system using deep learning," IRJET, vol. 12, no. 3, pp. 176–182, Mar. 2025.
2. T. R. Tejaswi, R. M. Mysoremath, P. D. Prabhu, H. Jain, and M. Mahitha, "Detection and prediction of diseases in arecanut plants," IRJET, vol. 11, no. 5, pp. 517–522, May 2024.
3. X. Zhang, B. A. Vinatzer, and S. Li, "Hyperspectral imaging analysis for early detection of tomato bacterial leaf spot disease," Scientific Reports, vol. 14, Art. no. 27666, Nov. 2024, doi: 10.1038/s41598-024-78650-6.
4. K. Premalatha, G. Naik, M. K. Naik, V. Hegde, B. C. Dhananjaya, and K. M. Satish, "Exploring yellow leaf disease patterns in areca plantations in Chikkamagaluru district of Karnataka," Saudi Journal of Pathology and Microbiology, vol. 9, no. 3, pp. 63–70, Mar. 2024, doi: 10.36348/sjpm.2024.v09i03.002.
5. V. Balafas, E. Karantoumanis, M. Louta, and N. Ploskas, "Machine learning and deep learning for plant disease classification and detection," IEEE Access, vol. 11, pp. 114352–114388, Oct. 2023, doi: 10.1109/AC-CESS.2023.3324722.
6. L. U. Khan, R. Zhao, H. Wang, and X. Huang, "Recent advances of the causal agent of yellow leaf disease (YLD) on areca palm (Areca catechu L.)," Tropical Plants, vol. 2, Art. no. 7, Mar. 2023, doi: 10.48130/TP-2023-0007.
7. T. Saranya, C. Deisy, S. Sridevi, and K. S. M. Anbananthen, "A comparative study of deep learning and internet of things for precision agriculture," Engineering Applications of Artificial Intelligence, vol. 122, Art. no. 106034, Mar. 2023, doi: 10.1016/j.engappai.2023.106034.
8. A. Hegde, V. S. Sadanand, C. G. Hegde, K. M. Naik, and K. D. Shastri, "Identification and categorization of diseases in arecanut: A machine learning approach," Indonesian Journal of Electrical Engineering and Computer Science, vol. 31, no. 3, pp. 1803–1810, Sep. 2023, doi: 10.11591/ijeecs.v31.i3.pp1803-1810.
9. A. Ahmad, D. Saraswat, and A. El Gamal, "A survey on using deep learning techniques for plant disease diagnosis and recommendations for development of appropriate tools," Smart Agricultural Technology, vol. 3, Art. no. 100083, Jun. 2022, doi: 10.1016/j.atech.2022.100083.
10. S. Barburiceanu, S. Meza, B. Orza, R. Malutan, and R. Terebes, "Con-volutional neural networks for texture feature extraction: Applications to leaf disease classification in precision agriculture," IEEE Access, vol. 9, pp. 162231–162247, 2021, doi: 10.1109/ACCESS.2021.3131002.
11. M. G. Anilkumar, T. G. Karibasaveshwara, H. K. Pavan, S. Urankar, and

A. Deshpande, "Detection of diseases in arecanut using convolutional neural networks," IRJET, vol. 8, no. 5, pp. 4282–4286, May 2021.

12. V. R. Raut and A. Nage, "Detection and identification of plant leaf diseases based on Python," IJERT, vol. 8, no. 5, pp. 1–4, May 2020.
13. B. Neeraja, A. Snehalatharani, N. B. V. C. Rao, and G. Ramanandam, "Studies on different formulations of the bioagent Trichoderma in the management of stem bleeding disease in coconut," Journal of Plantation Crops, vol. 48, no. 1, pp. 55–59, Jan.–Mar. 2020,
14. doi: 10.25081/jpc.2020.v48.i1.6217.
15. V. Hiremata, M. Narayanaswamy, H. Narayanswamy, and K. M. Satish, "Survey for the assessment

of arecanut yellow leaf disease (YLD) in Malnad region of Karnataka," International Journal of Ecology and Environmental Sciences, vol. 2, no. 3, pp. 326–329, 2020.
16. H. Narayanaswamy, J. Raju, K. Jayalakshmi, S. Sannaulla, K. S. Shesragiri, and M. D. Kumar, "Integrated management of basal stem rot disease of arecanut," International Journal of Chemical Studies, vol. 6, no. 5, pp. 530–531, 2018.
17. D. Harahap, "Plant diseases of areca nut," in Nut Crops – New Insights. IntechOpen, Apr. 2023, doi: 10.5772/intechopen.107504.