# AYVANA: Deep Learning-Based Disease Diagnosis in Dermatology and Agriculture

Chakrala Keerthi Sai
Student, AI & ML, Holy Mary Institute of Technology and Science, Hyderabad, TS, India

Basani Dhanush
Student, AI & ML, Holy Mary Institute of Technology and Science, Hyderabad, TS, India

Gandla Ashritha
Student, AI & ML, Holy Mary Institute of Technology and Science, Hyderabad, TS, India

Dr. P. Sumithabhashini
Head of Department, AI & ML, Holy Mary Institute of Technology and Science, Hyderabad, TS, India

Battula Aarthi
Student, AI & ML, Holy Mary Institute of Technology and Science, Hyderabad, TS, India

Dr. Venkataramana. B
Associate Professor, CSE, Holy Mary Institute of Technology and Science, Hyderabad, TS, India

*Abstract*—**Visual disease recognition in healthcare and agriculture involves learning discriminative image patterns across biologically distinct domains. While convolutional neural networks (CNNs) perform well in domain-specific tasks, most prior work relies on customized pipelines that limit architectural reuse. This work presents AYVANA, a consistent CNN-based framework with shared methodology and domain-specific optimization that evaluates the independent reuse of convolutional neural network (CNN) paradigm across dermatological and agricultural image-based disease classification tasks. Using established transfer learning, domain-specific pipelines are trained independently using domain-appropriate optimization strategies, while sharing a unified architectural paradigm, ensemble inference mechanism, evaluation, and interpretability approach. An EfficientNetB5–DenseNet169 ensemble achieves 85.61% accuracy and a macro F1-score of 0.814 on a ten-class dermatology dataset, while an EfficientNetB3–InceptionV3 ensemble attains 97.03% accuracy and a macro F1-score of 0.968 on a seventy-nine-class plant disease dataset. Soft-voting ensembles improve prediction consistency and balanced class-wise performance, and Grad-CAM visualizations provide qualitative insight. The results demonstrate that established CNN architectures can be independently retrained and reused across heterogeneous visual diagnostic under a consistent experimental protocol with shared evaluation and inference procedures, without shared representation learning or cross-domain feature transfer.**

*Keywords—Convolutional Neural Networks, Visual Disease Diagnosis, Dermatology, Agriculture, Transfer Learning, Ensemble Learning, Interpretability.*

## I. INTRODUCTION

Visual disease diagnosis plays a critical role in both healthcare and agriculture, where observable symptoms guide expert decision-making for clinical treatment and crop management. In dermatology, accurate identification of skin diseases enables timely intervention and improved patient outcomes, while in agriculture, early detection of plant diseases is essential for minimizing yield loss and ensuring food security. Across both domains, diagnosis relies heavily on visual cues such as texture variation, color change, lesion morphology, and structural distortions. Recent advances in deep learning, particularly convolutional neural networks (CNNs), have substantially improved automated disease recognition through transfer learning and the availability of annotated datasets. CNN-based models have demonstrated strong performance in dermatology by learning discriminative features related to pigmentation and lesion boundaries, and in agriculture by capturing visual symptoms such as discoloration and necrosis. Motivated by the observation that CNNs fundamentally operate as general-purpose visual pattern learners—detecting textures, shapes, and spatial structures independent of biological context—this work investigates whether CNN-based architectures can be independently reused across multiple visual disease diagnosis tasks spanning distinct domains. Despite significant biological differences, dermatological and agricultural datasets share common visual learning challenges, including multi-class classification, class imbalance, and high intra-class variability. To address this, we present AYVANA, a methodological deep learning framework that standardizes architectural selection, preprocessing, training structure, ensemble inference, evaluation, and interpretability across visually driven diagnostic tasks, while allowing domain-specific optimization and independent model training. Derived from the Sanskrit roots *"Aya"* (body or human health) and *"Vana"* (forest or plants), AYVANA reflects a unified perspective on human and plant health. Importantly, AYVANA is a methodological experimental framework rather than a novel neural architecture or learning algorithm, and it does not involve shared biological modelling or cross-domain representation learning.

## II. RELATED WORKS

Deep learning–based image analysis has become a dominant approach for automated disease diagnosis across multiple domains. In visually driven diagnostic tasks, convolutional neural networks (CNNs) effectively learn discriminative features from image data, enabling accurate classification of complex disease patterns. This section reviews prior work in dermatological and agricultural disease diagnosis, as well as ensemble and interpretability-based approaches, and outlines the research gap addressed by the proposed framework.

## A. Deep Learning for Dermatological Disease Diagnosis

CNN-based models have achieved significant success in dermatological image analysis, particularly in the classification of skin lesions and dermatological conditions from clinical and dermoscopic images. A landmark study by Esteva et al. demonstrated that deep CNNs can achieve dermatologist-level performance in skin cancer classification, establishing the feasibility of AI-assisted dermatological diagnosis [3]. Subsequent studies have expanded this direction by leveraging transfer learning with pretrained CNN architectures to improve diagnostic accuracy and data efficiency. Several works have explored the application of modern CNN architectures for dermatological disease classification. Venkataiah et al. applied transfer learning–based CNN models for diagnosing inflammatory skin diseases such as eczema and psoriasis, reporting improved classification performance [1]. Shakya et al. investigated hybrid frameworks combining CNN feature extraction with traditional classifiers for skin cancer analysis [2], while Prottasha et al. employed Inception-based architectures for real-time skin disease detection [4]. Comparative studies by Tschandl et al. and Haenssle et al. evaluated deep learning systems against professional dermatologists, demonstrating that CNN-based models often achieve competitive or superior diagnostic accuracy for specific lesion categories [6,8]. Han et al. further demonstrated the effectiveness of CNNs for distinguishing benign and malignant cutaneous tumours from clinical images [7]. Popescu et al. provided a comprehensive review of emerging neural network architectures for melanoma detection, highlighting current trends and challenges in dermatological AI systems [5]. Despite recent advances, many dermatological diagnostic systems still suffer from class imbalance, dataset bias, and limited interpretability. Moreover, most studies focus on single-domain performance optimization, with little attention to architectural reuse or methodological consistency across visually driven diagnostic tasks.

## B. Deep Learning for Agricultural Plant Disease Diagnosis

In agriculture, CNN-based models have been widely adopted for plant disease detection due to the visually observable nature of disease symptoms on leaves and crops. Large-scale datasets such as PlantVillage and other publicly available repositories have enabled deep learning models to achieve high classification accuracy across multiple crop species and disease categories. Recent studies have explored ensemble learning and detection-based approaches to reduce prediction variance and improve classification reliability in agricultural disease diagnosis. Ali et al. demonstrated that ensembles of deep learning architectures outperform individual CNN models by reducing prediction variance and improving minority-class recognition [11]. Object detection–based frameworks have also been investigated to localize disease regions under complex visual conditions. Miao et al. proposed an enhanced YOLOv8-based architecture for accurate plant disease detection [10], while Sambana et al. applied transfer learning with YOLO-based models to achieve high classification accuracy across multiple plant disease categories [9]. Lightweight CNN architectures suitable for smart farming and edge deployment have been explored by Vo et al. and Kirar, highlighting the potential of EfficientNet- and MobileNet-based systems for real-world agricultural applications [13,12].

Demilie presented a comparative analysis of plant disease detection techniques, emphasizing the strong empirical performance of CNN-based approaches compared to traditional machine learning methods [14]. Bao et al. proposed AX-RetinaNet for detecting tea leaf diseases with 93% accuracy [15], while Aldakheel et al. built a YOLOv4-based model achieving nearly perfect recognition of leaf infections [16]. Although agricultural disease detection systems often report high accuracy on benchmark datasets, many rely on curated images with limited environmental variability, raising concerns about real-world generalization. Like dermatological systems, most agricultural models remain domain-specific, using customized architectures and pipelines without emphasizing architectural consistency or reuse across diagnostic domains.

## C. Ensemble Learning & Interpretability in Visual Diagnosis

Ensemble learning is widely used in visual classification to improve performance consistency and reduce prediction variance by combining complementary model architectures. Soft-voting ensembles of CNNs have been shown in multiple studies to outperform individual models in both medical and agricultural imaging tasks, particularly under visually ambiguous conditions. Interpretability has likewise become an important consideration in applied deep learning. Techniques such as Gradient-weighted Class Activation Mapping (Grad-CAM) provide qualitative insight into model attention by highlighting image regions associated with predictions. However, in many existing studies, interpretability is treated as a post hoc visualization rather than as a systematically integrated component of the evaluation framework.

## D. Research Gap & Positioning of the Proposed Framework

While prior research demonstrates the effectiveness of CNN-based models for dermatological and agricultural disease diagnosis individually, existing approaches remain fragmented across multiple visual diagnostic domains, relying on domain-specific architectures, training strategies, and evaluation protocols. Consequently, there is limited investigation into whether deep learning architectures from convolutional neural network (CNN) paradigm can be independently applied across biologically distinct but visually diagnosable domains under a consistent methodological framework. Motivated by this gap, the proposed AYVANA framework emphasizes architectural paradigm reuse rather than shared cross-domain learning, applying established CNNs independently to both domains using a unified architectural paradigm, ensemble inference strategy, evaluation protocol, and interpretability approach, while allowing domain-specific training and optimization configurations.

## III. METHODOLOGY AND TRAINING STRATEGY

### A. Problem Formulation

This study investigates whether convolutional neural networks (CNNs), as a general-purpose visual learning paradigm, can be independently reused across biologically distinct but visually diagnosable domains under domain-appropriate training configurations. The problem is formulated as a multi-class image classification task, where RGB images are mapped to disease categories by learning discriminative

visual features such as texture irregularities, color variations, lesion morphology, and structural distortions. Importantly, the proposed framework does not involve joint training, shared representation learning, or cross-domain feature transfer. Dermatological and agricultural diagnostic pipelines are trained independently using domain-specific datasets. The shared framework in this work refers strictly to the reuse of a common architectural paradigm, consistent preprocessing, training structure, ensemble inference, evaluation protocol, and interpretability mechanism across visual diagnostic domains. This design enables an empirical assessment of architectural family reuse without introducing cross-domain learning.

### B.  Dermatological Disease Diagnosis Pipeline

*Dataset Preparation and Preprocessing*

The dermatological diagnostic pipeline operates on a curated dataset constructed from publicly available sources, comprising ten skin disease categories, including Actinic Keratosis, Basal Cell Carcinoma, Benign Keratosis, Fungal Infection, Melanocytic Nevus, Melanoma, Psoriasis, Seborrheic Keratoses, Squamous Cell Carcinoma, and Viral Infection [17,18]. Images corresponding to these categories are aggregated from established Kaggle repositories, including ISIC-labelled datasets and curated skin disease image collections. The use of multiple data sources introduces variability in acquisition conditions, including differences in lesion scale, color distribution, skin tone, illumination, and background characteristics, resulting in substantial intra-class variability. All images are processed in RGB format and resized to $224 \times 224$ pixels to match the input requirements of ImageNet-pretrained CNN backbones. Input normalization is performed using ImageNet mean and standard deviation values to ensure compatibility with pretrained weights. The consolidated dataset is organized into training, validation, and test partitions following the predefined directory structure of the source datasets. While explicit separation between training, validation, and test folders is maintained, the aggregation of multiple public sources may introduce low-level redundancy or correlated samples across splits, as patient-level identifiers and duplicate detection are not consistently available. Data augmentation is applied only to the training set and includes random rotations along with horizontal and vertical flips. Validation and test sets remain augmentation-free to support unbiased experimental evaluation.

*Model Architecture*

Two CNN backbones, EfficientNetB5 and DenseNet169, are employed to capture complementary dermatological features. EfficientNetB5 is selected for its strong performance in modelling fine-grained texture patterns, while DenseNet169 facilitates feature reuse through dense connectivity, supporting stable gradient propagation in deep networks. Both models are initialized with ImageNet-pretrained weights, and their original classification layers are replaced with task-specific fully connected heads corresponding to the ten dermatological disease classes. Global Average Pooling, as implemented in the selected CNN architectures, is applied prior to classification, to reduce feature dimensionality and mitigate overfitting. Model training follows a transfer learning setup in which all network parameters are trained end-to-end from ImageNet-initialized

weights using the Adam optimizer in conjunction with the OneCycleLR learning rate scheduler, without staged freezing of layers. Training proceeds for a fixed number of epochs, and model checkpoint selection is performed based on minimum validation loss rather than patience-based early stopping.

*Algorithm*

Input: Dermatology image dataset $D_{dermatology}$,
Output: Predicted disease class

1. Load dataset $D_{dermatology}$ and split into train/val/test
2. Resize images to $224 \times 224$ and normalize using ImageNet statistics
3. Apply data augmentation to training set only
4. Initialize EfficientNetB5 and DenseNet169 with pretrained weights
5. Replace classifiers with task-specific heads (10 classes)
6. Train models using the Adam optimizer with OneCycleLR learning rate scheduling
7. Select best model checkpoints based on validation loss.
8. Perform soft-voting ensemble on test set predictions
9. Apply Grad-CAM for lesion localization
10. Evaluate using Accuracy, F1-score, and Confusion Matrix
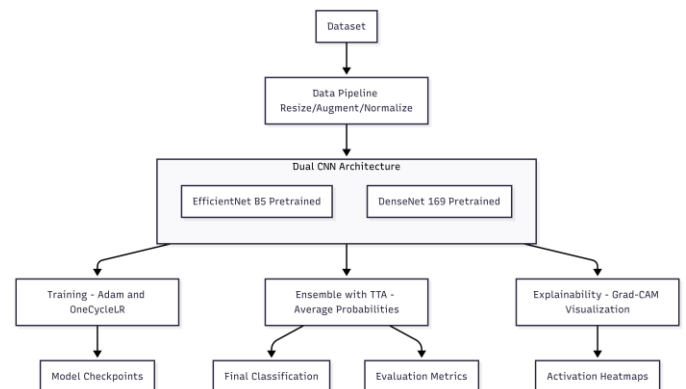


Fig. 1.  Block diagram of the dermatological image classification pipeline, showing preprocessing, EfficientNetB5 and DenseNet169 feature extraction, Adam training with OneCycleLR, soft-voting ensemble inference, evaluation, and Grad-CAM interpretability.

*Ensemble Inference and Interpretability*

During inference, predictions from EfficientNetB5 and DenseNet169 are combined using soft-voting ensemble averaging, where class probability outputs from both models are averaged to produce the final prediction. This ensemble strategy reduces prediction variance and improves prediction consistency for visually ambiguous skin conditions. To support interpretability, Gradient-weighted Class Activation Mapping (Grad-CAM) is applied to the final convolutional layers of both models. Grad-CAM heatmaps highlight lesion regions that contribute most strongly to model predictions, enabling qualitative verification that the models attend to visually salient regions associated with predicted classes rather than background artifacts.

**Published by :**
**https://www.ijert.org/**
**An International Peer-Reviewed Journal**

**International Journal of Engineering Research & Technology (IJERT)**
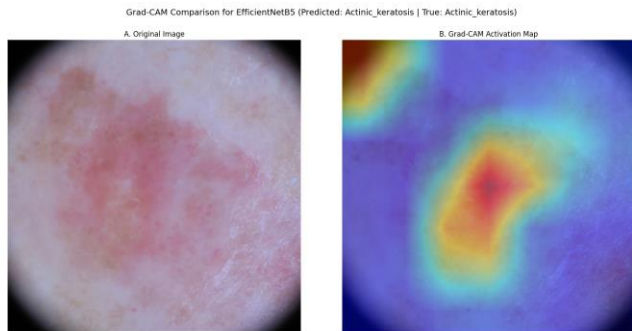**ISSN: 2278-0181**
**Vol. 15 Issue 01 , January - 2026**

Fig. 2. Grad-CAM visualization for EfficientNetB5 on an Actinic Keratosis image, showing the input image (left) and activation map (right).



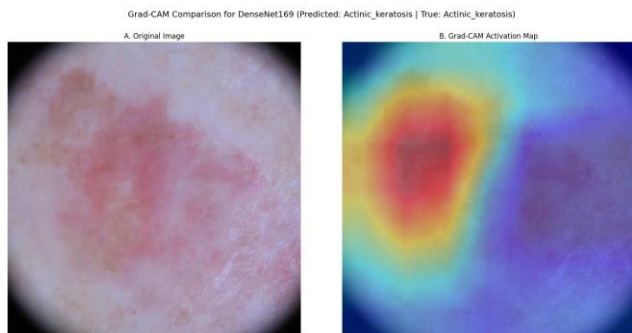Fig. 3. Grad-CAM visualization for DenseNet169 on an Actinic Keratosis image, with the input image (left) and activation map (right).

### C. Agricultural Plant Disease Diagnosis Pipeline

#### Dataset Preparation and Preprocessing

The agricultural diagnostic pipeline operates on a curated, multi-source dataset aggregated from publicly available Kaggle and Mendeley repositories, comprising seventy-nine disease and healthy classes across multiple crop species, including Apple[19,20], Cashew[21], Cassava[21], Corn[19,20], Cotton[32], Cucumber[22,24], Grape[19,20], Groundnut[25], Lentil[26], Onion[23], Potato[19,20], Rice[28,29,30], Soyabean[27], Tomato[19,20] and Wheat[31]. Images are collected under heterogeneous imaging conditions, with variations in illumination, background, leaf orientation, scale, and disease severity, resulting in substantial visual diversity. All images are processed in RGB format, resized to $299 \times 299$ pixels, and normalized using ImageNet statistics to ensure compatibility with pretrained CNN backbones. The dataset is organized into training, validation, and test partitions following predefined directory structures, without additional duplicate detection or cross-source overlap analysis. Data augmentation, including random rotations and horizontal and vertical flips, is applied only to the training set, while validation and test sets remain augmentation-free to support unbiased evaluation.

#### Model Architecture

Two CNN architectures, EfficientNetB3 and InceptionV3, are employed in the agricultural diagnostic pipeline to capture complementary visual features. EfficientNetB3 emphasizes efficient texture-sensitive representation, while InceptionV3 enhances multi-scale spatial feature extraction through parallel convolutional pathways. Both models are initialized with ImageNet-pretrained weights, and their classification heads are replaced with task-specific fully connected layers

corresponding to the 79 agricultural disease and healthy classes. Global Average Pooling, as implemented in the selected CNN architectures, is applied prior to classification. Model training follows a transfer learning paradigm involving a two-stage fine-tuning strategy. In the initial stage, ImageNet-pretrained EfficientNetB3 and InceptionV3 backbones are frozen while task-specific classification heads are optimized using the Adam optimizer. In the subsequent stage, the full network is unfrozen and fine-tuned with reduced learning rates using StepLR-based scheduling. Model checkpoint selection is based on minimum validation loss to mitigate overfitting. Backbone freezing is implemented through parameter freezing and optimizer parameter scoping rather than explicit gradient disabling.

#### Algorithm

Input: Agricultural image dataset $D_{agriculture}$,
Output: Predicted disease class

1. Load $D_{agriculture}$ dataset and split into train/val/test
2. Resize images to 299×299 and normalize using ImageNet statistics
3. Apply data augmentation to training data only
4. Initialize pretrained EfficientNetB3 and InceptionV3 with pretrained-weights
5. Replace final classifiers with 79-class task-specific heads
6. Train models using Adam optimizer with StepLR scheduling
7. Reduce learning rate in later epochs
8. Save best checkpoints based on validation loss
9. Perform soft-voting ensemble on test predictions
10. Generate Grad-CAM visual explanations
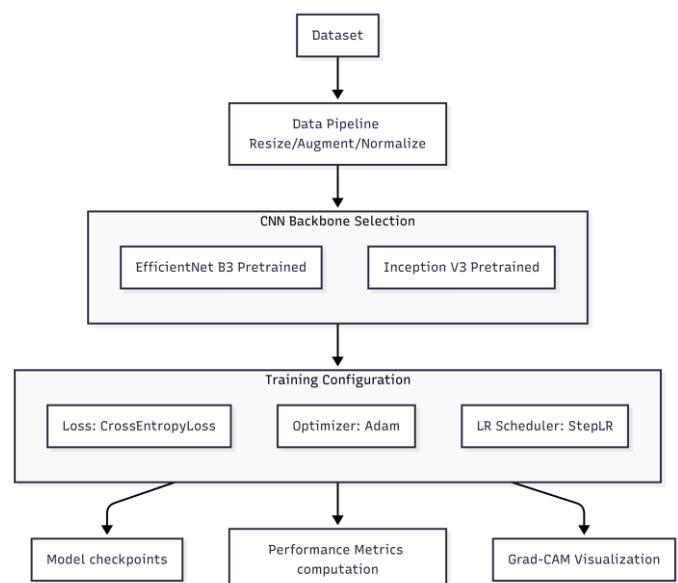11. Evaluate using Accuracy, F1-Score and Confusion matrix



Fig. 4. Block diagram of the agricultural plant disease classification pipeline, including preprocessing, EfficientNetB3 and InceptionV3 backbones, Adam with StepLR, evaluation, and Grad-CAM interpretability.

**Published by :**
**https://www.ijert.org/**
**An International Peer-Reviewed Journal**

**International Journal of Engineering Research & Technology (IJERT)**
**ISSN: 2278-0181**
**Vol. 15 Issue 01 , January - 2026**

*Ensemble Inference and Interpretability*

At inference time, probability outputs from EfficientNetB3 and InceptionV3 are combined using soft-voting ensemble averaging. This ensemble strategy improves classification stability and reduces misclassification among visually similar plant disease categories. Interpretability is achieved using Grad-CAM, which is applied to the final convolutional layers of both models. The resulting heatmaps qualitatively highlight infected regions such as necrotic tissue and discoloration on plant leaves, indicating that model predictions are driven by disease-relevant visual cues rather than background artifacts. This qualitative evidence supports the prediction consistency of the ensemble model's decision-making process in complex agricultural imaging scenarios.
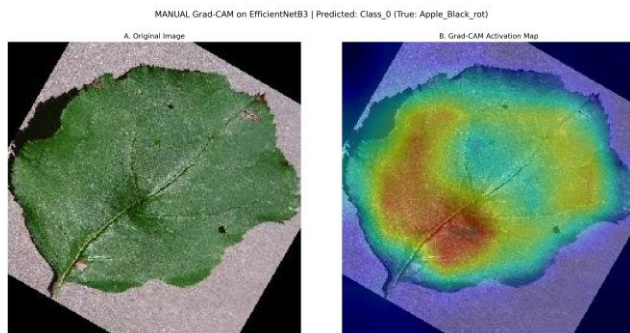


Fig. 5. Grad-CAM visualization for EfficientNetB3 on an Apple Black Rot leaf image, showing the input image (left) and activation map (right).
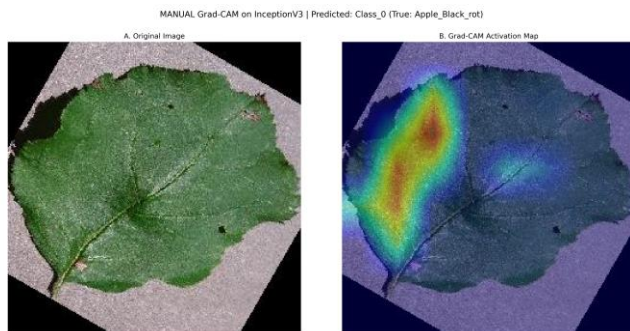


Fig. 6. Grad-CAM visualization for InceptionV3 on an Apple Black Rot leaf image, with the input image (left) and activation map (right).

### D. Evaluation Protocol and Reproducibility

Model performance across both domains is evaluated using classification accuracy, macro-averaged F1-score, and confusion matrix analysis. Macro-averaged metrics are emphasized to ensure balanced assessment across classes with varying sample frequencies. All experiments are implemented using the PyTorch framework with pretrained models sourced from torchvision and are conducted in a GPU-accelerated environment using NVIDIA Tesla T4 GPUs. Fixed random seeds are applied across Python, NumPy, and PyTorch environments, and consistent preprocessing pipelines and evaluation procedures are maintained to ensure procedural consistency across experiments. Minor numerical variations may still arise due to GPU-level nondeterminism. Dataset splits follow the predefined directory structure of the aggregated public datasets. While explicit separation between training, validation, and test folders is preserved, no additional duplicate

detection or cross-source overlap analysis is performed. Consequently, reported results should be interpreted as representative experimental outcomes rather than statistically bounded estimates. No multi-seed variance analysis or statistical significance testing is conducted. The proposed methodology does not involve shared parameter learning, multi-task optimization, or cross-domain feature transfer. Dermatological and agricultural pipelines are trained independently using domain-specific datasets. While both pipelines employ transfer learning with ImageNet-pretrained CNN architectures, learning rate scheduling strategies differ by domain: dermatological models utilize the OneCycleLR learning rate scheduler, whereas agricultural models employ an epoch-based StepLR scheduler, reflecting domain-specific optimization configurations. The methodological unification in this work refers to architectural family reuse, consistent preprocessing, ensemble inference, standardized evaluation metrics, and qualitative interpretability mechanisms, rather than identical optimization configurations.

## IV. RESULTS AND ANALYSIS

This section reports experimental results obtained using the proposed AYVANA framework for dermatological and agricultural image-based disease classification. Performance is evaluated using complementary metrics to account for class imbalance, dataset scale, and visual complexity, with comparisons between individual CNN backbones and ensemble configurations. Grad-CAM visualizations provide qualitative insight into model attention but do not constitute clinical or agronomic validation. All results are based on single experimental runs with fixed random seeds; no multi-seed variance analysis or statistical significance testing is performed, and conclusions are limited to observed performance under the evaluated experimental configurations.

### A. Dermatological Disease Diagnosis

The dermatological diagnostic pipeline is evaluated on a ten-class skin disease classification task using a held-out test set. Performance is primarily assessed using classification accuracy, macro F1-score, and confusion matrix analysis, which together provide a balanced evaluation under class imbalance and clinical visual ambiguity. Table I summarizes the performance of individual CNN models and the ensemble configuration.

TABLE I. PERFORMANCE COMPARISION FOR DERMATOLOGICAL DISEASE DIAGNOSIS

| Metrics | CNN Models Performance | | |
|---|---|---|---|
| | *DenseNet169* | *EfficientNetB5* | *Ensemble* |
| *Accuracy* | 81.15% | 83.24% | 85.61% |
| *F1-Score* | 0.7645 | 0.7820 | 0.8140 |

The ensemble model achieves the highest overall performance, improving accuracy by approximately 2–4% over individual backbones. The increase in macro F1-score indicates improved balanced classification across both common and minority disease classes. These results suggest that combining complementary CNN architectures reduces prediction variance and improves prediction consistency in visually ambiguous

dermatological conditions. Training and validation loss curves exhibit smooth convergence behavior under the applied learning rate scheduling. Both EfficientNetB5 and DenseNet169 exhibit smooth training and validation loss trends over epochs, with validation accuracy steadily improving. EfficientNetB5 shows a slightly larger train–validation gap, while DenseNet169 demonstrates smoother and more balanced learning behaviour.
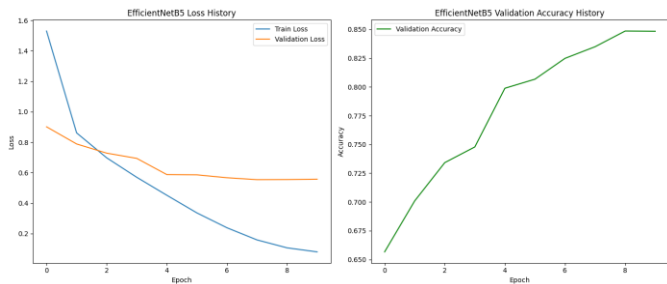


Fig. 7. Training loss and validation accuracy curves for EfficientNetB5 on the dermatological dataset, demonstrationg model convergence across training epochs.
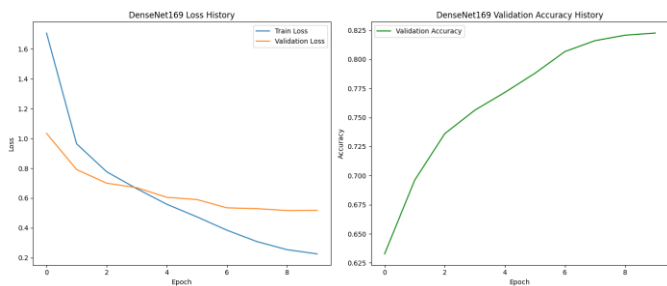


Fig. 8. Training loss and validation accuracy curves for DenseNet169 on the dermatological dataset, demonstrating learning progression across training epochs.

### B. Agricultural Disease Diagnosis

The agricultural diagnostic pipeline is evaluated on a large-scale dataset comprising seventy-nine disease and healthy classes across multiple crop species. Table II presents the comparative performance of individual CNN models and their ensemble.

TABLE II. PERFORMANCE COMPARISION FOR AGRICULTURAL DISEASE DIAGNOSIS

| Metrics | CNN Models Performance | | |
| --- | --- | --- | --- |
| | *InceptionV3* | *EfficientNetB3* | *Ensemble* |
| *Accuracy* | 95.12% | 96.45% | 97.03% |
| *F1-Score* | 0.942 | 0.958 | 0.968 |

Both individual models demonstrate strong performance, reflecting the effectiveness of transfer learning for large-scale plant disease classification. The ensemble configuration consistently outperforms individual models, yielding improvements in both accuracy and macro F1-score. Training and validation accuracy for both EfficientNetB3 and InceptionV3 increase steadily across epochs, with a clear performance boost observed as training progresses. Corresponding loss curves exhibit smooth convergence, with a

consistent reduction in loss throughout training. Validation metrics closely track training metrics, and both models exhibit smooth training and validation loss trends over epochs under the evaluated experimental configuration, with no evidence of severe overfitting within the observed training regime. Overall, these trends suggest that the chosen optimization strategy and learning rate scheduling are well-suited for the agricultural dataset, enabling efficient model convergence.
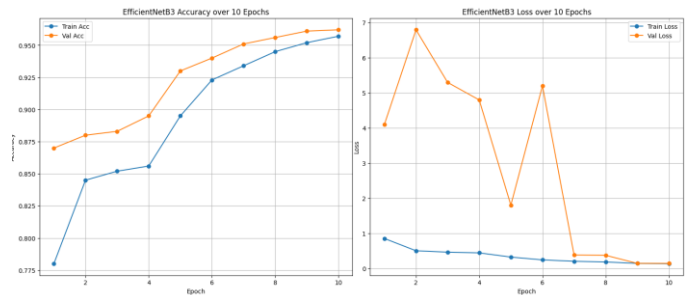


Fig. 9. Training and validation accuracy and loss curves for EfficientNetB3 over 10 epochs on the agricultural dataset, illustrating convergence behavior under Adam optimization with StepLR scheduling.
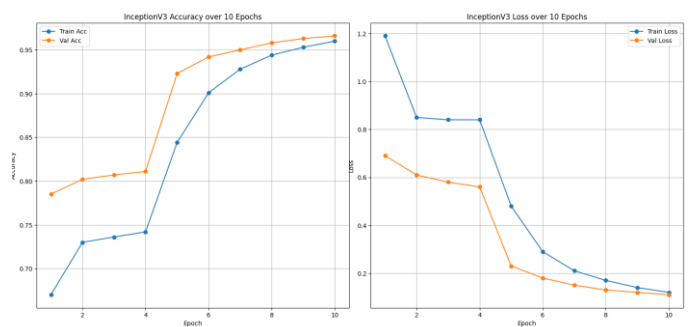


Fig. 10. Training and validation accuracy and loss curves for InceptionV3 over 10 epochs on the agricultural disease dataset, illustrating learning progression under Adam optimization with StepLR scheduling.

### C. Ensemble vs Individual Model Analysis

This analysis compares individual CNN backbones with their corresponding soft-voting ensemble configurations to assess the effectiveness of ensemble inference under a unified training and evaluation protocol. For dermatological disease classification, the ensemble combining EfficientNetB5 and DenseNet169 surpasses both standalone models, achieving an accuracy of 85.61% and a macro F1-score of 0.814. The gain in macro F1-score reflects improved class-wise balance, particularly for visually similar and underrepresented disease categories. In the agricultural domain, EfficientNetB3 and InceptionV3 already demonstrate strong individual performance; however, their ensemble further improves prediction consistency, attaining 97.03% accuracy and a macro F1-score of 0.968 across 79 disease and healthy classes. Across both domains, soft-voting ensemble inference consistently outperforms single-model configurations while maintaining identical inference pipelines and without introducing additional training complexity. These findings indicate that model complementarity can be effectively leveraged at the inference stage, enhancing prediction consistency within independently trained domain-specific classification tasks. Moreover, the consistent gains across two biologically distinct domains

reinforce the consistency of the proposed ensemble strategy across independently trained domain-specific classification tasks.

### D. Summary of Performance of Each Domain

The proposed framework demonstrates performance improvements across both dermatological and agricultural disease classification tasks despite substantial differences in dataset scale, class count, and visual complexity. In the dermatology domain, the ensemble model achieves an accuracy of 85.61% with a macro F1-score of 0.814, reflecting the challenge of visually ambiguous and clinically complex skin disease categories. In contrast, the agricultural domain attains higher absolute performance, with 97.03% accuracy and a macro F1-score of 0.968, attributable to clearer visual separability and larger dataset size.

TABLE III.          DOMAIN PERFORMANCE COMPARISION

| Metrics | Domain | |
|---|---|---|
| | *Dermatology* | *Agriculture* |
| *Classes* | 10 | 79 |
| *Ensemble Accuracy* | 85.61% | 97.03% |
| *Macro F1-Score* | 0.8140 | 0.968 |

These results indicate that CNN-based architectures, when trained independently under a consistent methodological framework, can be reused across biologically distinct visual diagnostic domains without domain-specific architectural redesign.

### E. Evaluation Metrics

Model performance is assessed using classification accuracy, macro F1-score and confusion matrix. Accuracy provides an overall performance measure, while macro F1-score evaluates balanced performance across classes by treating all classes equally. Confusion matrices are used to analyze class-wise prediction behavior and misclassification patterns.

#### Confusion Matrix

The confusion matrices illustrate the performance of the ensemble CNN framework across both dermatology and agriculture domains by comparing true class labels with predicted labels. In the dermatology task, the matrix exhibits strong diagonal dominance, indicating that most skin diseases are correctly classified, with a macro F1-score of 0.814 reflecting balanced precision and recall across all 10 classes. Misclassifications primarily occur between visually similar conditions such as melanoma and melanocytic nevus or actinic keratosis and basal cell carcinoma, which is expected due to overlapping visual patterns and inherent class imbalance in medical datasets. In contrast, the agricultural confusion matrix displays an exceptionally sharp diagonal with minimal off-diagonal errors, resulting in a high overall accuracy of 97.03%. This highlights the model's strong discriminative capability across a large number of plant disease and healthy classes. Collectively, these matrices demonstrate that although dermatological diagnosis presents greater visual ambiguity, the proposed ensemble framework achieves strong empirical performance across both domains under the evaluated experimental conditions. The results further emphasize the

benefit of ensemble inference in enhancing class separability and consistency in complex, real-world classification scenarios.
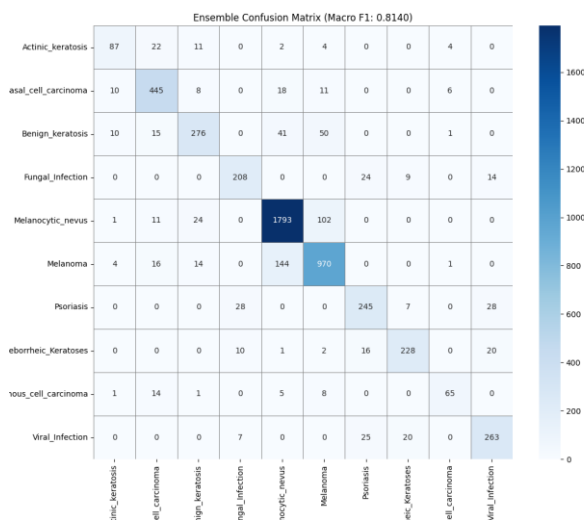


Fig. 11. Confusion matrix of the ensemble CNN model for dermatological disease classification across 10 classes, reporting a macro F1-score of 0.814 and highlighting class-wise prediction behavior.
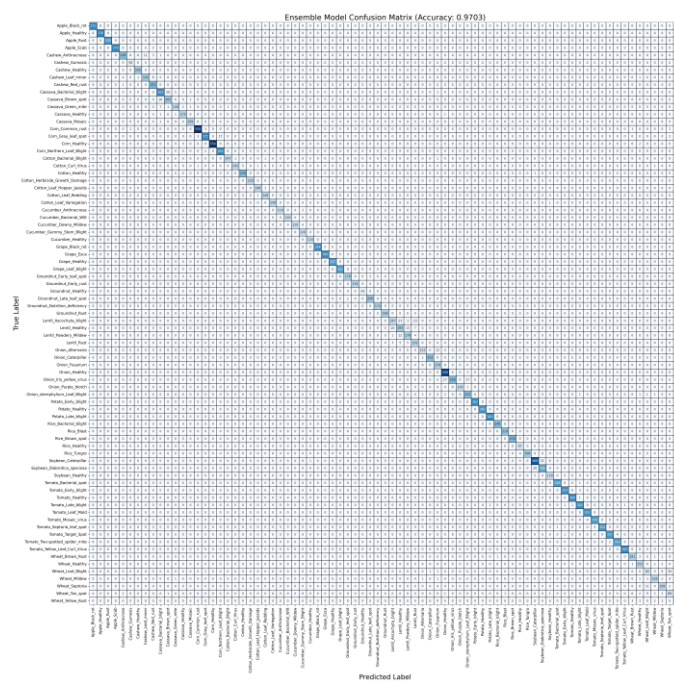


Fig. 12. Confusion matrix of the ensemble CNN model for agricultural disease classification across 79 disease and healthy classes, illustrating prediction distribution with an overall test accuracy of 97.03%.

### F. Ablation Analysis

The selection of EfficientNetB5, DenseNet169, EfficientNetB3, and InceptionV3 is guided by domain-specific visual characteristics and engineering trade-offs rather than an exhaustive comparison across all deep learning architectures. Transformer-based models such as Vision Transformers are excluded due to their higher data requirements and weaker inductive bias for local texture learning, which is critical for both dermatological lesion analysis and plant disease

recognition. Object detection frameworks such as YOLO are also not considered, as this work addresses closed-set image-level disease classification rather than region-level detection, and introducing detection-based models would alter the task formulation and evaluation protocol. Within the CNN family, the chosen architectures represent complementary design principles. EfficientNet is selected for its compound scaling strategy, enabling strong performance with balanced computational cost. DenseNet169 is paired with EfficientNetB5 in the dermatology domain due to its dense feature reuse, which preserves fine-grained texture and color information required for visually ambiguous skin diseases. In contrast, InceptionV3 is paired with EfficientNetB3 in the agricultural domain, as its multi-branch architecture effectively captures multi-scale spatial patterns commonly observed in plant disease imagery. The use of EfficientNetB5 for dermatology and EfficientNetB3 for agriculture reflects differences in dataset complexity and visual ambiguity. Dermatological images exhibit subtle intra-class variations that benefit from higher-capacity models, while agricultural images are larger in scale with clearer visual separability, allowing EfficientNetB3 to achieve strong performance with improved efficiency. Soft-voting ensemble inference is employed to exploit architectural complementarity without increasing training complexity. Across both domains, ensembles consistently improve macro-averaged F1-score, indicating reduced prediction variance and more balanced class-wise performance, particularly for visually similar or underrepresented disease categories. These gains demonstrates enchanced performance under the evaluated experimental conditions beyond raw accuracy, supporting the reuse of established CNNs under a consistent experimental framework.

## V. CONCLUSION AND FUTURE WORKS

### A. Conclusion

This work presented AYVANA, a unified deep learning framework that investigates the applicability of the convolutional neural network (CNN) paradigm across two biologically distinct yet visually diagnosable domains: dermatological disease identification and agricultural crop disease classification. By employing consistent preprocessing strategies, transfer learning–based CNN backbones, and a standardized experimental pipeline, the study demonstrates that modern CNN architectures can be independently retrained and reused across multiple diagnostic contexts without requiring domain-specific architectural redesign. Through extensive experiments on publicly available benchmark datasets, the framework achieved competitive classification performance in both domains, confirming the strong capacity of CNNs to learn discriminative visual features such as texture irregularities, colour variations, lesion morphology, and structural distortions. The results indicate that, despite fundamental biological differences between human skin and plant pathology, visually driven diagnostic tasks can be addressed using a shared methodological foundation when supported by appropriate dataset-specific training and evaluation. Importantly, AYVANA does not seek to establish cross-domain biological generalization or shared representation learning between

dermatology and agriculture. Instead, the primary contribution lies in demonstrating architectural reusability and methodological consistency, showing that a single deep learning framework can be systematically applied across heterogeneous visual diagnostic problems. This finding reinforces the architectural flexibility of CNN-based models for visual pattern recognition tasks when applied under domain-specific training settings. In addition to predictive performance, the integration of Grad-CAM–based interpretability provides qualitative insights into the regions influencing model decisions, enhancing transparency and supporting trust in automated diagnostic systems. While these explanations are not clinically definitive, they offer a useful mechanism for visual validation of learned features in both domains. Overall, this study contributes empirical evidence supporting the use of unified CNN architectures for multi-domain visual diagnostics and highlights the potential of such frameworks to streamline development across application areas. The proposed approach serves as a technical foundation for future research exploring multimodal learning, uncertainty estimation, real-world validation, and domain adaptation, thereby advancing the development of reliable and scalable AI-driven diagnostic systems.

### B. Limitations

While the proposed AYVANA framework demonstrates that convolutional neural network (CNN) architectures can be effectively applied to both dermatological and agricultural disease diagnosis, several limitations must be acknowledged. Although the framework unifies the experimental pipeline, the dermatological and agricultural models are trained and evaluated independently, with no shared feature learning, cross-domain representation transfer, or domain adaptation. Accordingly, AYVANA does not claim biological or semantic generalization between human skin and plant pathology, but rather emphasizes architectural reusability under a consistent methodology. The experimental evaluation relies exclusively on publicly available benchmark datasets, which, while standardized, may not fully reflect real-world variability. In dermatology, this includes differences in skin tone, lesion progression, and imaging conditions, while in agriculture, real field environments introduce challenges such as occlusion, variable lighting, and background clutter. As a result, the reported performance may not directly translate to unconstrained deployment settings. Furthermore, the framework operates under a closed-set classification assumption, limiting its ability to handle rare, emerging, or out-of-distribution disease cases, as no explicit uncertainty estimation or unknown-class rejection mechanisms are incorporated. Finally, the proposed system has not undergone clinical or field-level validation, and its predictions are not benchmarked against expert-confirmed diagnoses. In addition, AYVANA relies solely on single-modality RGB image analysis, without incorporating complementary contextual or multimodal information such as patient metadata, environmental factors, or temporal disease progression. These limitations indicate that the results should be interpreted strictly as evidence of methodological consistency and architectural reuse, rather than real-world diagnostic readiness or cross-domain biological generalization.

## C. Future Work

Several research directions can extend the AYVANA framework and enhance its applicability across visual diagnostic domains. A key avenue involves exploring cross-domain representation learning through shared feature extractors or domain-adaptive training to examine whether certain visual patterns generalize beyond independent domain-specific learning. Another important direction is the integration of multimodal inputs. In dermatology, incorporating patient metadata or dermoscopic imagery may improve diagnostic consistency, while in agriculture, combining visual data with environmental factors, crop growth stage information, or temporal progression could enhance robustness under real-world conditions. Future work may also focus on uncertainty estimation and out-of-distribution detection to improve prediction safety. Techniques such as predictive calibration or confidence-aware rejection mechanisms would enable the system to abstain from uncertain predictions, which is critical for real-world deployment. In addition, clinical and field-level validation involving dermatologist-confirmed diagnoses and agronomist-verified field trials is essential to bridge the gap between benchmark evaluation and practical usability. Finally, deployment-oriented research, including model compression, pruning, and quantization, as well as more advanced interpretability methods beyond Grad-CAM, could further improve efficiency, transparency, and user trust.

## REFERENCES

[1] Venkataiah, H., Prasad, T., Gopinath, G., Charitha, B., Teja, G., and Reddy, B., "CNN and transfer learning methods for enhanced dermatological disease detection," *International Journal of Electronics and Telecommunications*, pp. 337–344, 2025. https://doi.org/10.24425/ijet.2025.153607

[2] Shakya, M., Patel, R., and Joshi, S., "A comprehensive analysis of deep learning and transfer learning techniques for skin cancer classification," *Scientific Reports*, vol. 15, p. 4633, 2025. https://doi.org/10.1038/s41598-024-82241-w

[3] Esteva, A., Kuprel, B., Novoa, R. A., et al., "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, pp. 115–118, 2017. https://doi.org/10.1038/nature21056

[4] Prottasha, M. S. I., Farin, S., Ahmed, M., Rahman, M., Hossain, A. B. M., and Kaiser, M. S., "Deep learning-based skin disease detection using convolutional neural networks," in *Proc. Int. Conf. on Intelligent Systems*, 2023. https://doi.org/10.1007/978-981-19-80329_39

[5] Popescu, D., El-Khatib, M., El-Khatib, H., and Ichim, L., "New trends in melanoma detection using neural networks: A systematic review," *Sensors*, vol. 22, no. 2, p. 496, 2022. https://doi.org/10.3390/s22020496

[6] Tschandl, P., Codella, N., Akay, B. N., et al., "Comparison of the accuracy of human readers versus machine-learning algorithms for pigmented skin lesion classification," *The Lancet Oncology*, vol. 20, pp. 938–947, 2019. https://doi.org/10.1016/S1470-2045(19)30333-X

[7] Han, S. S., Kim, M. S., Lim, W., Park, G. H., Park, I., and Chang, S. E., "Classification of clinical images for benign and malignant cutaneous tumors using deep learning," *Journal of Investigative Dermatology*, vol. 138, no. 7, pp. 1529–1538, 2018. https://doi.org/10.1016/j.jid.2018.01.028

[8] Haenssle, H. A., Fink, C., Schneiderbauer, R., et al., "Man against machine: Diagnostic performance of a deep learning CNN for dermoscopic melanoma recognition," *Annals of Oncology*, vol. 29, no. 8, pp. 1836–1842, 2018. https://doi.org/10.1093/annonc/mdy166

[9] Sambana, B., Nnadi, H. S., and Wajid, M. A., "An efficient plant disease detection using transfer learning approach," *Scientific Reports*, vol. 15, p. 19082, 2025. https://doi.org/10.1038/s41598-025-02271-w

[10] Miao, Y., Meng, W., and Zhou, X., "SerpensGate-YOLOv8: An enhanced YOLOv8 model for accurate plant disease detection," *Frontiers in Plant Science*, vol. 15, p. 1514832, 2025. https://doi.org/10.3389/fpls.2024.1514832

[11] Ali, A., Yousef, A., Raja, M. A., and Abdelal, M., "An ensemble of deep learning architectures for accurate plant disease classification," *Ecological Informatics*, vol. 81, p. 102618, 2024. https://doi.org/10.1016/j.ecoinf.2024.102618

[12] Kirar, J. R., "A next-generation plant disease detection system using transfer learning and Edge Impulse," *International Journal of Scientific Research in Engineering and Technology*, vol. 10, no. 3, 2024. https://ijsret.com/wp-content/uploads/2024/05/IJSRET_V10_issue3_174.pdf

[13] Vo, H.-T., Quach, L.-D., and Ngoc, H. T., "Ensemble of deep learning models for multi-plant disease classification in smart farming," *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 5, 2023. https://doi.org/10.14569/IJACSA.2023.01405108

[14] Demilie, W. B., "Plant disease detection and classification techniques: A comparative study," *Journal of Big Data*, vol. 11, p. 5, 2024. https://doi.org/10.1186/s40537-023-00863-9

[15] Bao, W., Fan, T., and Hu, G., "Detection and identification of tea leaf diseases based on AX-RetinaNet," *Scientific Reports*, vol. 12, p. 2183, 2022. https://doi.org/10.1038/s41598-022-06181-z

[16] Aldakheel, E. A., Zakariah, M., and Alabdalall, A. H., "Detection and identification of plant leaf diseases using YOLOv4," *Frontiers in Plant Science*, vol. 15, p. 1355941, 2024. https://doi.org/10.3389/fpls.2024.1355941

[17] Shaju, R. E., *ISIC Skin Disease Image Dataset (Labelled)* [Dataset]. Kaggle, 2025. https://www.kaggle.com/datasets/riyaelizashaju/isic-skin-disease-image-dataset-labelled

[18] International Skin Imaging Collaboration, *ISIC 2019 Challenge Dataset* [Dataset], 2019. https://challenge2019.isic-archive.com/data.htm

[19] Mohanty, S. P., Hughes, D. P., and Salathé, M., "Using deep learning for image-based plant disease detection," *Frontiers in Plant Science*, vol. 7, p. 1419, 2016. https://doi.org/10.3389/fpls.2016.01419

[20] Sheneamer, A., *New Plant Diseases Dataset* [Dataset]. Kaggle, 2024. https://www.kaggle.com/datasets/vipooool/new-plant-diseases-dataset

[21] Mensah, P. K., Akoto-Adjepong, V., Adu, K., et al., *Dataset for Crop Pest and Disease Detection* [Dataset]. Mendeley Data, v1, 2023. https://doi.org/10.17632/bwh3zbpkpv.1

[22] Sultana, N., Shorif, S. B., Akter, M., and Uddin, M. S., *Cucumber Disease Recognition Dataset* [Dataset]. Mendeley Data, v1, 2022. https://doi.org/10.17632/y6d3z6f8z9.1

[23] Patil, T., *Onion Diseases Dataset* [Dataset]. Kaggle, 2025. https://www.kaggle.com/datasets/tejasbargujepatil/onion-diseases

[24] Kapadnis, S., *Cucumber Disease Recognition Dataset* [Dataset]. Kaggle, 2025. https://www.kaggle.com/datasets/sujaykapadnis/cucumber-disease-recognition-dataset

[25] Prem Kumar, E., *Multi-Crop Disease Dataset* [Dataset]. Mendeley Data, v1, 2025. https://doi.org/10.17632/6243z8r6t6.1

[26] Mahamud, E., and Tapos, M. A., *Lentil Plant Disease Image Dataset (4 Class)* [Dataset]. Mendeley Data, v2, 2024. https://doi.org/10.17632/7vb77bz2st.2

[27] Misignoni, M., *Soybean Leaf Dataset* [Dataset]. Kaggle, 2025. https://www.kaggle.com/datasets/maeloisamignoni/soybeanleafdataset

[28] Antony, L., and Prasanth, L., *Rice Leaf Diseases Dataset* [Dataset]. Mendeley Data, v1, 2023. https://doi.org/10.17632/dwtn3c6w6p.1

[29] Anshul, M., *Rice Disease Dataset* [Dataset]. Kaggle, 2025. https://www.kaggle.com/datasets/anshulm257/rice-disease-dataset

[30] Jawadali, A., *20K Multi-Class Crop Disease Images* [Dataset]. Kaggle, 2025. https://www.kaggle.com/datasets/jawadali1045/20k-multi-class-crop-disease-images

[31] Agarwal, K., Yadav, V., and Suthar, T., *Wheat Plant Diseases Dataset* [Dataset]. Kaggle, 2025. https://www.kaggle.com/datasets/kushagra3204/wheat-plant-diseases

[32] Bishshash, P., Nirob, M. A. S., Shikder, M. H., and Sarower, A., *SAR-CLD-2024: A Comprehensive Dataset for Cotton Leaf Disease Detection* [Dataset]. Mendeley Data, v2, 2024. https://doi.org/10.17632/b3jy2p6k8w.2