An Interpretable, UAV Imagery-Based Deep Learning Approach for Plant Disease Detection and Severity Estimation in Indian Agriculture

Gunde Veeraswami, Kadurla Nagaraju , Kamatam Hruthikesh, and Mrs.M .Vijaya lakshmi(*Project Guide*)

Abstract—In the vast and diverse agricultural landscape of India, plant diseases remain a persistent and escalating threat to both national food security and the economic stability of millions of farming households. Traditional methods of disease monitoring, which rely heavily on manual field surveys, are increasingly becoming a bottleneck due to chronic labor shortages and the sheer geographic scale of farming operations. To address these systemic inefficiencies, we propose a scalable, deep learning-based solution that leverages Unmanned Aerial Vehicle (UAV) imagery for automated disease detection. Utilizing the MH-SoyaHealthVision dataset, specifically tailored to Indian soybean crops, we have completed the critical first phase of data curation and balancing to address severe class imbalances. Our proposed system involves a robust five-phase architecture designed to identify critical diseases such as Mosaic and Rust with high precision. Crucially, our approach addresses the "trust gap" in AI adoption by moving beyond simple predictions; we integrate Explainable AI (XAI) mechanisms to generate visual heat maps. This report synthesizes findings from recent pivotal studies, details the completed data analysis phase, and outlines the architectural roadmap designed to provide farmers with actionable, interpretable insights for proactive crop management.

Index Terms—Deep Learning, UAV, Precision Agriculture, Plant Disease Detection, Explainable AI (XAI), Data Balancing, Indian Agriculture.

I. Introduction

For the Indian economy, agriculture is far more than just a commercial sector; it is a lifeline that supports a vast majority of the population and serves as the fundamental backbone of the nation's Gross Domestic Product (GDP). However, this critical industry faces a constant, invisible battle against biological threats. Plant diseases are a leading cause of yield reduction and quality degradation, resulting in significant financial losses every harvest season [6]. The complexity of managing these diseases is compounded by the immense diversity of crops grown across the subcontinent, particularly vital cash crops like soybean.

Historically, the primary line of defense has been "fieldwalking"—the age-old practice where farmers or agricultural extension officers physically patrol rows of crops to spot signs of infection on leaves and stems. While this method has served humanity for centuries, it is struggling to keep pace with the demands of modern, high-yield agriculture.

A. The Limitations of Manual Monitoring

The most significant drawback of manual monitoring is that it is inherently reactive rather than proactive. By the time a human observer notices a disease outbreak with the naked eye, the infection may have already spread significantly across the field, making containment difficult and costly. Furthermore, physically surveying acres of land under the hot sun is an exhausting, time-consuming, and expensive endeavor, particularly given the rising cost of agricultural labor.

A more subtle, yet critical issue is the reliance on subjective human judgment. Distinguishing between a nutrient deficiency and a fungal infection often requires expert knowledge. Without immediate access to agronomists, diagnosis can be inconsistent, leading to incorrect treatments or the "preventative" overuse of chemicals. As the agricultural workforce shrinks and landholdings remain fragmented, the need for an automated, consistent, and scalable monitoring solution has never been more urgent.

B. The Technological Shift: UAVs and AI

Unmanned Aerial Vehicles (UAVs), commonly known as drones, provide a compelling solution to these ground-level limitations. Capable of covering large areas rapidly, UAVs equipped with high-resolution cameras can capture field data from angles impossible to achieve on foot. The science behind this is grounded in plant physiology: stress caused by disease alters how a plant reflects light, often before physical symptoms become visible to the human eye.

By pairing this aerial perspective with Deep Learning, specifically Convolutional Neural Networks (CNNs), we can automate the diagnosis process. These AI models mimic the human visual cortex, learning to recognize the unique visual "fingerprints" of specific diseases with high precision. Our project aims to operationalize this technology, providing farmers with a tool that enables timely, data-driven decisions—ultimately reducing the overuse of pesticides and securing crop yields.

II. Motivation

A. Contextualizing for the Indian Farmer

Most global research in precision agriculture focuses on vast, monoculture farms typical of the West. India's reality is markedly different; it is dominated by small, diverse landholdings. For a small-scale Indian farmer, crop loss is not just a statistical variation—it is a direct threat to their livelihood and family's food security. Early detection can mean the difference between a profitable season and significant debt. We are motivated by the need for a precise, localized solution that fits this unique context.

B. Addressing the "Data Gap"

A major hurdle in applying global AI advancements to Indian agriculture is the "data gap." While sophisticated models exist, they are typically trained on datasets from Europe, North America, or China. These models often fail when applied to Indian fields due to subtle differences in crop varieties, soil colors, and lighting conditions. We are driven by the recent release of the MH-SoyaHealthVision dataset, which provides high-quality, localized data. This allows us to build a tool that is not just theoretically sound, but practically relevant to the Indian environment.

III. Literature Survey

To ground our work in the current state-of-the-art, we analyzed ten key papers published between 2021 and 2025. This survey specifically focuses on the model architectures employed and their suitability for agricultural tasks.

A. Advanced Hybrid Architectures

A significant advancement was demonstrated by Samha et al. (2025) in their work "Advanced Crop Health Monitoring," where they utilized a fusion model combining Transfer Learning with Attentive Convolutional Recurrent Neural Networks (ACRNN) [1]. Their approach leveraged CNNs for spatial feature extraction and RNNs to capture sequential dependencies, significantly improving interpretability through attention mechanisms. While highly accurate, the computational complexity of ACRNN poses challenges for real-time deployment on standard drone hardware.

B. Lightweight Models for Edge Computing

Addressing the need for efficiency, Kumar and Murugan (2023) focused on edge deployment for cashew farming [2]. They adapted the MobileNetV2 architecture, a model specifically designed for mobile and embedded vision applications. By prioritizing depth-wise separable convolutions, they achieved real-time processing speeds directly on the UAV. However, this efficiency came with a trade-off in accuracy compared to deeper networks, and their application was limited to detecting a single disease type.

C. Semantic Segmentation Models

Precise localization of disease was the focus for Pan et al. (2021) in their study on Wheat Yellow Rust [4]. They employed the Pyramid Scene Parsing Network (PSPNet), a model known for its ability to aggregate global context information. This allowed for pixel-level classification (segmentation) rather than just image-level classification. While achieving high accuracy, segmentation models like PSPNet require pixellevel annotated datasets, which are labor-intensive to create.

Similarly, Ye et al. (2025) improved upon the U-Net architecture, developing the MGA-UNet for citrus disease detection [9]. U-Net is renowned for biomedical image segmentation, and its adaptation here proved effective for isolating small disease lesions. However, their reliance on multispectral data makes the solution expensive for small-scale farmers.

D. Transformer-Based Approaches

The shift towards Transformer architectures was highlighted by Gunder et al. (2024) in "SugarViT" [7]. They utilized Vision Transformers (ViTs), which process images as sequences of patches. This allows the model to capture longrange dependencies across the image, which is crucial for estimating overall disease severity. While ViTs represent the cutting edge, they are notoriously data-hungry and require massive computational resources for training.

E. Generative Models for Data Augmentation

To tackle the perennial issue of data scarcity, Faria et al. (2025) introduced "PotatoGANs" [10]. They employed Generative Adversarial Networks (GANs) to creating synthetic images of diseased plants. This innovative approach allows for the expansion of training datasets without waiting for natural disease outbreaks, directly addressing the class imbalance problem we face in our project.

F. Dataset Contributions

Finally, Shinde and Attar (2023) provided the foundational contribution for our work by releasing the "MHSoyaHealthVision" dataset [3]. While this was a data paper and did not propose a new model architecture, it provided the essential localized data needed to train any of the aforementioned sophisticated models for the Indian context.

IV. Problem Definition & Objectives

A. Problem Definition

The core problem is to design and evaluate a deep learning system that accurately detects and maps multiple diseases in Indian fields using UAV imagery. The system must move beyond simple detection to provide actionable insights, such as severity estimation and interpretable heat maps, addressing the lack of localized AI solutions.

B. Objectives

- To develop a custom Convolutional Neural Network (CNN) model tailored for detecting key diseases in Indian crops.
- To produce a practical output: a simple, color-coded "heat map" for farmers to easily visualize problem areas.
- To create a working template that can be adapted for developing AI solutions for other crops across India.

V. Proposed System Architecture

Our proposed system is structured into five distinct phases, creating a pipeline that moves from raw data acquisition to actionable, interpretable insights.

TABLE I: Comparative Analysis of Model Architectures in Literature

Author (Year)	Crop	Model Architecture	Key Contribution	Limitations/Gaps
Samha et al. (2025) [1]	Generic	CNN + ACRNN Fusion	Attention mechanism improves feature focus.	High computational cost for edge devices.
Kumar & Murugan (2023) [2]	Cashew	MobileNetV2	Optimized for speed on low-power hardware.	Lower accuracy; limited to single disease.
Pan et al. (2021) [4]	Wheat	PSPNet (Segmentation)	Global context aggregation for precise mapping.	Requires expensive pixel-level labeling.
Shi et al. (2021) [5]	Potato	CropDocNet	Custom architecture for hyperspectral data.	High sensor cost makes it impractical.
Gunder et al. (2024) [7]	Sugar Beet	Vision Transformer (ViT)	Captures long-range dependencies for severity.	Extremely high training resource needs.
Faria et al. (2025) [10]	Potato	GANs (Generative)	Synthesizes new data to fix imbalances.	Synthetic data may lack realworld variance.

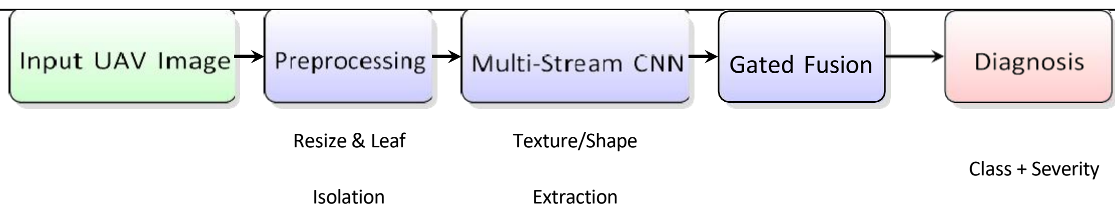


Fig. 1: Operational Flow: A high-level visualization of how a single input image travels through the system pipeline.

A. Phase 1: Data Acquisition and Curation (Completed)

The foundation of any AI model is data. We utilized the MH-SoyaHealthVision dataset [3]. In this phase, we performed a rigorous audit of the 2,845 images. We identified a critical class imbalance—Healthy leaves were represented by only 281 images, while Rust disease had 1,001 images.

As shown in Table II, the raw data was heavily biased towards diseased samples. To prevent model bias, we completed a data balancing strategy involving augmentation (rotation, flipping) to equalize the classes.

TABLE II: Original Dataset Distribution (Pre-Balancing)

Class Name	Count	Status
Healthy	281	Severely Underrepresented
Mosaic Disease	773	Underrepresented
Rust Disease	1001	Overrepresented
Pest Attack	790	Moderate
Total	2845	Imbalanced

This imbalance presented a significant risk: a model trained on Table II would likely ignore healthy plants. Therefore, we applied synthetic augmentation to bring all classes to parity, resulting in the balanced distribution shown in Table III.

TABLE III: Final Balanced Dataset Distribution (PostAugmentation)

Class Name	Final Count	Status
Healthy	1,000	Augmented to Parity
Mosaic Disease	1,000	Augmented to Parity
Rust Disease	1,000	Original Volume
Pest Attack	1,000	Original Volume
Total	4,000	Fully Equalized

The completion of this phase ensures that the model provides a fair and reliable foundation for the subsequent learning phases.

B. Phase 2: Smart Preprocessing Gatekeeper

Raw drone imagery contains noise, varying lighting, and background artifacts like soil and weeds. In this phase, we implement a ”Preprocessing Gatekeeper” to standardize the input.

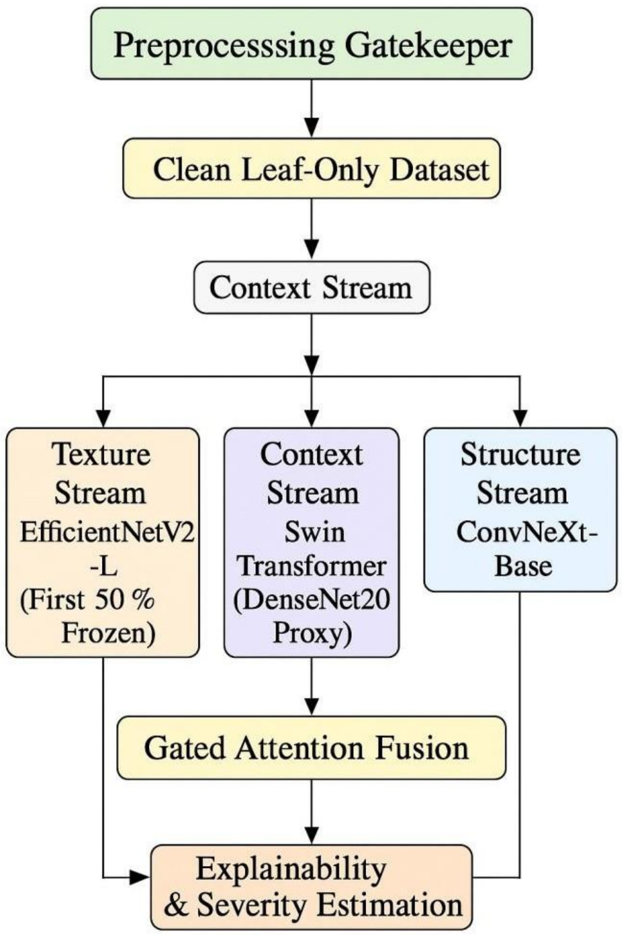
- Leaf Isolation: We utilize region-based analysis to automatically segment the image. This process creates a binary mask that separates the plant leaves from the soil background. By feeding only the ”Leaf-Only” data into the model, we force the AI to focus strictly on biological health indicators rather than environmental artifacts.
- Normalization: Scaling pixel intensity values ensures the model converges quickly during training.

C. Phase 3: Multi-Stream Model Architecture

This phase involves the design of a sophisticated multistream Convolutional Neural Network (CNN). Unlike standard models that process an image in a single pass, our architecture uses three distinct streams simultaneously:

- Texture Stream: This stream utilizes efficient convolutional layers to focus on high-frequency surface details. It is specifically designed to identify fungal infections like Rust, which manifest as distinct textural changes on the leaf surface.

- Structure Stream: This stream analyzes the geometric shape and edges of the leaf. It is crucial for detecting Pest Attacks, where the primary symptom is physical damage (bites or holes) rather than discoloration.
- Context Stream: This stream employs attention-based mechanisms (like Transformers) to consider the broader leaf area. It is optimized to identify patterns like Mosaic virus mottling, which requires understanding the relationship between different parts of the leaf.



Final model architecture for detecting soybean leaf health classes.

Fig. 2: Multi-Stream Architecture: The model splits the input into Texture, Context, and Structure streams, fuses the features, and provides explainable outputs.

D. Phase 4: Gated Fusion and Optimization

The insights from the three streams in Phase 3 must be combined intelligently. We implement a ”Gated Attention Fusion” mechanism. This layer acts as a smart arbiter; it learns to weigh the importance of each stream dynamically based on the specific image. For example, if an image shows clear physical damage, the gate will prioritize the Structure Stream’s output. The model is then trained using adaptive optimization algorithms to minimize error rates and ensure high precision.

E. Phase 5: Explainability and Severity (XAI)

The final phase addresses the ”black box” problem. We integrate Gradient-weighted Class Activation Mapping (GradCAM). This technique generates a heat map overlay, highlighting exactly which pixels in the image led to the disease prediction. Additionally, this phase calculates a ”Severity Score” based on the density of the infection, converting the technical output into a practical alert level (Low, Medium, High) for the farmer.

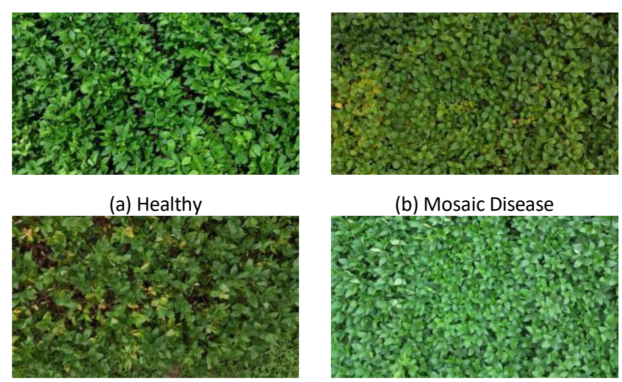
VI. Dataset Analysis (Work Done)

A. The MH-SoyaHealthVision Dataset

Our work relies on the MH-SoyaHealthVision dataset, a localized collection containing 2,845 images of Indian soybean crops. This dataset is crucial because it captures the specific lighting conditions and crop varieties found in Maharashtra, filling the global data gap.

B. Visual Class Representation

To visualize the distinct features our model must learn, sample images are provided in Fig. 3.



(a) Healthy (b) Mosaic Disease

(c) Rust Disease (d) Pest Attack

Fig. 3: Visual samples of the four target classes from the MHSoyaHealthVision dataset. The XAI module will be trained to highlight the specific lesion areas seen in (b), (c), and (d).

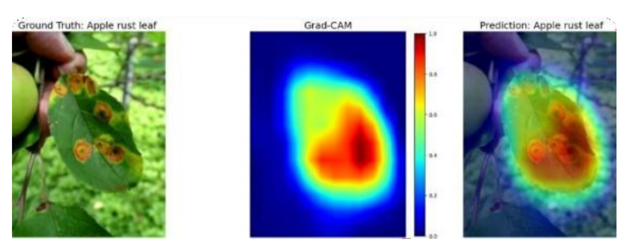


Fig. 4: Expected Explainable AI Output: The original image (left) compared with the Grad-CAM heatmap (right), which highlights the specific infected regions used for diagnosis.

VII. Future Scope & Conclusion

A. Scalability and Future Work

The architecture developed in this project acts as a robust template for Indian agriculture. While our current focus is on Soybean, the underlying methodology is crop-agnostic. Future iterations of this project will focus on adapting the model to other regionally significant crops that face similar biological threats. By creating a pipeline that integrates new datasets, we aim to ensure the system remains adaptable to changing agricultural conditions across the country.

B. Conclusion

Modernizing Indian agriculture requires tools that are not only technologically advanced but also practically accessible and trustworthy. This project addresses the critical gap in localized AI solutions by leveraging the MH-SoyaHealthVision dataset. We have successfully completed the first phase of our roadmap: identifying the research gap and curating a balanced, localized dataset.

By designing a system that goes beyond simple detection to offer interpretability (via XAI) and severity estimation, we aim to provide farmers with a transparent, useful tool. The integration of interpretable heatmaps solves the "black box" problem, fostering trust among users. Moving forward, our focus will shift to the rigorous training of the CNN backbone and fine-tuning the severity estimation modules to ensure high accuracy in real-world field conditions.

References

[1] A. K. Samha et al., "Advanced Crop Health Monitoring for Smart Agriculture Using Fusion of Transfer Learning With Attentive Convolutional Recurrent Neural Network on Remote Sensing Images," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, 2025.

[2] M. Kumar R and B. Murugan MS., "Artificial Intelligence based drone for early disease detection and precision pesticide management in cashew farming," *arXiv preprint arXiv:2303.08556*, 2023.

[3] S. Shinde and V. N. Attar, "An Indian UAV and leaf image dataset for integrated crop health assessment of soybean crop," *Data in Brief*, vol. 50, Art. no. 109517, 2023.

[4] Q. Pan, M. Gao, P. Wu, J. Yan, and S. Li, "A Deep-Learning-Based Approach for Wheat Yellow Rust Disease Recognition from Unmanned Aerial Vehicle Images," *Sensors*, vol. 21, no. 19, 2021.

[5] Y. Shi, L. Han, A. Kleerekoper, S. Chang, and T. Hu, "A Novel CropDocNet for Automated Potato Late Blight Disease Detection from the Unmanned Aerial Vehicle-based Hyperspectral Imagery," *arXiv preprint arXiv:2107.13277*, 2021.

[6] S. N. A. M. Robi, N. Ahmad, M. A. M. Izhar, H. M. Kaidi, and N. M. Noor, "Utilizing UAV Data for Neural Network-based Classification of Melon Leaf Diseases in Smart Agriculture," *Int. J. Adv. Comput. Sci. Appl.*, vol. 15, no. 1, 2024.

[7] M. Gunder, et al., "SugarViT-Multi-Objective Regression of UAV Images with Vision Transformers and Deep Label Distribution Learning Demonstrated on Disease Severity Prediction in Sugar Beet," *arXiv preprint arXiv:2311.03076*, 2024.

[8] Y. Liu, Y. Song, P. Cui, Y. Fang, and B. Su, "Diagnosis of grapevine leafroll disease severity infection via UAV remote sensing and deep learning," *Smart Agriculture*, vol. 5, no. 3, pp. 49-61, 2023.

[9] N. Ye, et al., "Early detection of Citrus Huanglongbing by UAV remote sensing based on MGA-UNET," *Frontiers in Plant Science*, vol. 14, Art. no. 1303645, 2023.

[10] F. T. J. Faria, et al., "PotatoGANs: Utilizing Generative Adversarial Networks, Instance Segmentation, and Explainable AI for Enhanced Potato Disease Identification and Classification," *arXiv preprint arXiv:2405.07332*, 2024.

[11] S. P. Mohanty, D. P. Hughes, and M. Salathe, "Using deep learning for´ image-based plant disease detection," *Frontiers in Plant Science*, vol. 7, p. 1419, 2016.

[12] E. C. Tetila et al., "Detection and classification of soybean pests using deep learning with UAV images," *Computers and Electronics in Agriculture*, vol. 179, p. 105836, 2020.

[13] J. G. A. Barbedo, "Plant disease severity estimation using optical imaging," *Remote Sensing*, vol. 11, no. 11, p. 1376, 2019.

[14] A. Kamilaris and F. X. Prenafeta-Boldu, "Deep learning in agriculture: A´ survey," *Computers and Electronics in Agriculture*, vol. 147, pp. 70-90, 2018.

[15] X. Zhang et al., "Swin Transformer for plant disease classification," *Agriculture*, vol. 12, no. 8, p. 1123, 2022.