

Development of a UAV-based crop disease detection system using deep learning algorithms to enhance precision agriculture

Rupanjali Debbarma^{1,2,*},
Aditya Sankar Sengupta¹

¹ Department of ECE, ICFAI
University, Tripura, India

² Department of EEE, Tripura
Institute of Technology, Narsingarh,
Agartala, India

*Email: rupanjaldebbarma23@
outlook.com; rupanjali.db22@gmail.
com

Received for publication
December 23, 2025.

Abstract

The shift toward precision agriculture has made it clear that there is a need for smart and automated systems that can detect crop diseases promptly and with high precision. A research project is being discussed here, which presents a comprehensive drone-based solution for detecting cotton leaf disease, combining the latest advances in deep learning with the gradual training of the machine for rapid decision-making in agriculture. In the beginning, the system implements adaptive median filtering (AMF), followed by the contrast-limited adaptive histogram equalization (CLAHE) process, which removes noise and enhances the image's contrast. The Swin Transformer is used to identify spatial and spectral features at different levels, while the combination of ConvNeXt-V2 and the squeeze-and-excitation blocks with bidirectional long short-term memory (SE-BiLSTM) classifier refines spatial cues and learns sequential dependencies for precise disease classification. Moreover, the framework includes a Deep Q-Learning (DQL) recommendation module that supplies contextually aware crop management activities, like the optimized use of pesticides and scheduling of irrigation. The experimental results, derived from the SAR-CLD-2024 dataset, indicate a remarkable performance with an accuracy of 99.30%, thereby surpassing existing convolutional neural network (CNN)-based and hybrid models in terms of precision, recall, and F1-score. The combination of deep learning with reinforcement learning in Unmanned Aerial Vehicle (UAV)-based systems indicates a future of sustainable, adaptable, and data-driven precision agriculture.

Keywords

UAV-based monitoring, precision agriculture, cotton leaf disease, Swin Transformer, ConvNeXt-V2, SE-BiLSTM, Deep Q-Learning

Abbreviations

3D-CNN: Three-Dimensional Convolutional Neural Network; AI: Artificial Intelligence; ConvNeXt-V2: Convolutional Next Version 2; DenseNet: Densely Connected Convolutional Network; DRL: Deep Reinforcement Learning; FNR: False Negative; Rate FPR: False Positive Rate; GELU: Gaussian Error Linear Unit; LSTM-H: Long Short-Term Memory with Hidden layers; ResNet: Residual Network; ROC-AUC: Receiver Operating Characteristic – Area Under the Curve; SAR-CLD-2024: Synthetic Aperture Radar–Cotton Leaf

Disease 2024 dataset; SE-BiLSTM: Squeeze-and-Excitation Bidirectional Long Short-Term Memory; t-SNE: t-distributed Stochastic Neighbor Embedding; UAV: Unmanned Aerial Vehicle. VGGNet: Visual Geometry Group Network

I. Introduction

Precision agriculture operates as a modern farming method that employs contemporary technologies for data processing and sensing, automation to boost productivity, and improve crop production [1], [2], [3]. Drone technology is one of the most advanced monitoring methods, enabling coverage of large agricultural areas and capturing high-quality, multi-temporal images. Aerial photographs and satellite images reveal the health of the plants, moisture, and nutrients in the soil, which aids decision making based on data [4], [5]. However, the precise and quick recognition of cotton diseases in yields still remains a challenge in terms of yield and quality control to date. The combination of UAV-based imaging and deep learning has been acknowledged as a great improvement in disease monitoring and a remarkable decrease in the need for human support during field inspections [6].

The main reason for crop diseases is attributed to the interaction of various biological, environmental, and management factors, which are quite complex [7], [8]. Pathogenic microorganisms are one of the primary causes of crop infections and are composed of fungi, bacteria, and viruses. Climate is also a significant factor, influencing its components, such as temperature, humidity, and precipitation [9], [10]. Additionally, cotton plants undergo a lot of stress due to improper watering, lack of nutrients, and insect problems, which in turn make them more susceptible to diseases. These factors vary not only between different locations but also change rapidly in response to the climate. Consequently, traditional disease monitoring methods in cotton plants become unreliable, slow, and often ineffective in large or mixed-crop areas [11].

In recent years, a number of methods have been investigated to use machine learning and deep learning techniques to automatically detect cotton crop illnesses [12], [13]. The first methods were based on the use of conventional imaging processing techniques that included color thresholding, texture analysis, and specialized feature extraction. After that,

the convolutional neural networks (CNN), ResNet, VGGNet, and DenseNet architectures proved that they were much better at automated feature learning compared to earlier models [14], [15]. Additionally, UAV-based imaging, combined with UNet and DeepLabv3+ models, enabled the achievement of very precise disease segmentation in aerial photographs [16]. Nevertheless, these solutions still have drawbacks, including being highly dependent on lighting conditions, a limited ability to generalize across diverse crops or seasons, and, lastly, being costly in terms of computing power, which altogether limits their application in the field [17].

To address these difficulties, the researchers propose a UAV-based crop disease detection system utilizing deep learning techniques, which are the most advanced for achieving high precision in agriculture. The suggested approach consists of a sequence of steps that first applies noise filtering and contrast enhancement to the image in the pre-processing phase to increase the clarity of the image, then utilizes UNet++ to perform highly accurate disease segmentation. The Swin Transformer captures features at different levels, while the ConvNeXt-V2 classifier ensures strong and precise disease classification. After that, the development of a Deep Q-Learning (DQL)-based recommendation module is carried out to supply adaptive and optimized farm management decisions, such as pesticide application and irrigation scheduling. The integrated framework aims to facilitate the precise and instantaneous disease monitoring, while also supporting informed decision making. It thereby directly links AI innovations to agricultural activities in the field. The contributions of the paper are summarized in the following points:

- Uses adaptive median filtering (AMF) and contrast-limited adaptive histogram equalization (CLAHE) for enhancing UAV-acquired images, leading to higher visibility and more accurate disease diagnosis regardless of the outdoor field conditions.

- Employs a hierarchical attention mechanism grounded on Swin Transformer to proficiently capture both localized and global disease traits from aerial images.
- The proposed model is a combination of ConvNeXt-V2 and squeeze-and-excitation (SE) blocks with bidirectional long short-term memory (SE-BiLSTM), which is capable not only of filtering out disease-relevant channels but also of learning different patterns of sequential texture to achieve maximum accuracy in classification.
- A DQL-powered recommendation module is developed, which communicates the suggestions of adaptive and data-derived crop management practices. Hence, the system becomes more than just a detector and is a practical decision support tool.

The organization of the paper is as follows: A literature review is given in Section 2, which also mentions the most recent developments. Section 3 explains the suggested method with the steps of UAV image pre-processing, feature extraction using Swin Transformer, classifying through ConvNeXt-V2 with SE-BiLSTM, and the recommendation module of DQL. Section 4 shows the experimental results and provides a discussion on them. Last but not least, Section 5 wraps up the research and proposes some future research paths.

II. Literature Review

Drones in agriculture and agricultural monitoring have revolutionized precision agriculture, enabling the collection of superior-quality data in a matter of seconds. Drones carrying multispectral, hyperspectral, and RGB sensors have brought about an improved level of precision in disease detection, trustworthy yield determination, and in-depth research of plant characterizations. As per the recent article by Casas et al. [18], the use of UAVs to provide the multispectral images has made machine learning algorithms so powerful that they can accurately identify the diseases in *Phoenix canariensis* palm groves. Aerial imaging has shown its remarkable benefits for extensive crop monitoring in this manner. Similarly, Fei et al. [19] proposed the application of multi-sensor data fusion along with machine learning techniques as a more accurate way of predicting wheat yields, thus pointing out the necessity of mixing sources for better agricultural decision making. Zhu et al. [20] likewise reported such results for oilseed rape, where

the integration of spatial and spectral data not only enhanced the accuracy of predictions but also the robustness of the model.

Deep learning frameworks have been developed in a manner that has indeed changed the process of detecting and diagnosing diseases using UAVs, making it significantly faster and more accurate. Das and Raghuvanshi [21] introduced a deep Radial Basis Function Network (RBFN) with multidisciplinary mixed attention employed to the problem of leaf disease classification from UAV images. The performance of their model was excellent in detecting very small disease symptoms, even under varying lighting and background conditions. Linero-Using UAV-based multispectral photos and deep learning methods, Ramos et al. [22] studied the scalability dataset for black Sigatoka disease classification in banana plantations. The outcomes of their research showed that the factors affecting the performance of neural networks are primarily the diversity and quality of the dataset; thus, the necessity of well-annotated and scalable UAV datasets for trustworthy agricultural AI systems has been highlighted.

The application of UAVs in agriculture has expanded beyond disease detection to include crop maturity analysis and pest management. A UAV-mounted deep learning framework for strawberry maturity detection was created by Singh et al. [23], which provided very high real-time classification accuracy in natural field settings. Gokeda and Yalavarthi [24] presented a deep hybrid IoT-UAV model for pest detection that combined real-time image analytics with communication networks, consequently offering a scalable and effective method for precise pest surveillance. The research not only underscores the importance of disease diagnosis but also highlights the ever-growing role of UAVs in automating critical agricultural tasks.

Recent research has focused on intelligent UAV navigation and decision-making frameworks aimed at enhancing operational efficiency. A one-tiered UAV path-planning system introduced by Fu et al. [25] combined the use of an augmented Double Deep Q-Network (D3QN) algorithm with remote sensing data, resulting in a significant increase in path coverage and a considerable reduction in energy consumption during field operations. Conversely, livestock movements in UAV monitoring networks were forecasted using the LSTM-H model, a hybrid deep learning system introduced by Bokani et al. [26], which highlighted the significance of temporal sequence modelling in agricultural monitoring. In addition, Raptis et al. [27] demonstrated the development of a

comprehensive precision agriculture system that leverages UAV-based functionalities for resource application and mapping activities tailored to the specific characteristics of individual fields.

Alavilli et al. [28] proposed an integrated AI framework combining transformer-guided graph learning, 3D-CNN, BiLSTM, and neuro-symbolic fuzzy reasoning for accurate and interpretable cancer prediction. This multi-modal fusion strategy inspired the proposed architecture's hybrid feature learning and decision support design.

The studies investigated together demonstrate the increasing use of UAVs and deep learning algorithms in precision agriculture, which in turn has paved the way for applications in disease detection, yield prediction, pest management, and field automation. Nevertheless, most of the approaches that are currently in place are prone to noise interference in aerial imagery, vary in segmentation accuracy, and are not very adaptable to different crop species. To address these challenges, a UAV-based crop disease detection system is proposed in this paper, equipped with both deep learning and reinforcement learning components. The system utilizes noise filtering and contrast enhancement for pre-processing, UNet++ for segmentation, Swin Transformer for feature extraction, and ConvNeXt-V2 for disease classification, with a DQL-based recommendation module added last. The integrated design is expected to provide real-time, accurate, and intelligent decision support for sustainable precision agriculture.

a. Research gap

The majority of the literature on UAV-automated monitoring in agriculture has built upon existing research [29]; however, the literature still lacks an in-depth examination of some important issues. The primary focus of earlier research has been on detecting diseases or estimating yield. Furthermore, no single method exists that offers pre-processing of images, segmentation, classification, and support for decision-making all at once [30]. Powerful conventional convolutional neural networks (CNNs) that are the most frequently used in models, in spite of their power, often fail to give a high accuracy due to changing lighting conditions, highly complex backgrounds, and varying drone altitudes [31]. In addition to this, the prevailing UAV-based deep learning systems often rely on limited or crop-specific datasets, which results in poor generalization to the different field conditions. One of the main disadvantages of the present situation is that it is void of intelligent optimization

and reinforcement learning methods [32], which are very good at adaptive decision-making. Static methods prevail, and they are unable to adjust not only to different environmental conditions but also to varying health conditions of the crop. Moreover, the earlier models fell short of the necessary segmentation accuracy, primarily due to the use of simple architectures that could not effectively capture the multi-scale contextual information from UAV images [33]. The specific research puts forward an innovative cohesive and intelligent system for UAV-assisted crop disease identification whose components are denoising and contrast improving for sophisticated pre-processing, UNet++ for reliable segmenting, Swin Transformer for feature extracting, and ConvNeXt-V2 for classifying, and then the entire setup being rounded off with a DQL-based recommendation module for farm management that is dynamic thus permitting to make exact and data-driven choices for the efficient and environmental friendly precision agriculture.

III. Proposed Methodology

The proposed UAV-based cotton leaf disease detection framework integrates advanced deep learning and attention-driven mechanisms to enhance precision agriculture decision-making. Initially, UAV-acquired images of the field undergo a comprehensive pre-processing process, where the AMF removes high-frequency noise while the CLAHE enhances local contrast, resulting in a noise-free and visually improved dataset. The pre-processed images are then passed to the Swin Transformer for feature extraction, where hierarchical shifted-window self-attention captures both local and global dependencies, enabling effective representation of spatial textures and disease-specific patterns. Subsequently, the multi-scale features extracted are adjusted and categorized using a hybrid ConvNeXt-V2 with SE-BiLSTM framework. The block of ConvNeXt-V2 amplifies the ability to differentiate between various spatial cues. The SE mechanism highlights the channels most pertinent to the disease in a dynamic manner, and the BiLSTM monitors the temporal correlations that exist in the feature maps of the different patches throughout its process. The disease classifications are linked to practical measures through a precision agriculture recommendation layer, which simultaneously provides field-level instructions, such as pesticide spraying, irrigation control, and nutrient adjustment. This entire process ensures high diagnostic accuracy, robustness that is continually improved,

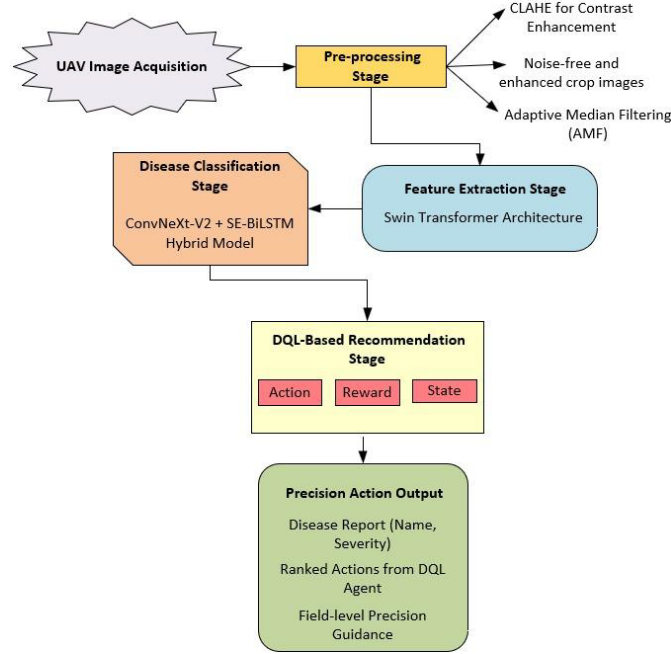


Figure 1: Proposed cotton leaf disease detection framework. AMF, adaptive median filtering; CLAHE, contrast-limited adaptive histogram equalization; DQL, Deep Q-Learning.

and the capacity to handle various UAV imaging conditions, thereby facilitating direct integration with future smart farming systems. The workflow diagram for the cotton leaf disease detection model is illustrated in Figure 1.

a. Pre-processing stage

UAVs frequently capture images of crops that reveal differences due to lighting, atmospheric disturbances, and high-frequency noise generated by the camera. These differences are mainly caused by the flight altitude, the movement of the UAV, and the prevailing environmental conditions. Thus, image preprocessing becomes a necessary step to ensure the stability and precision of the next classification model. The proposed system incorporates two main processes in the preprocessing step: (1) noise removal through AMF and (2) contrast boost by CLAHE. The output of this phase is images that have had their noise reduced and contrast increased, while retaining the texture and color information important for crop disease recognition.

a.i AMF

The AMF regulates the window size according to the local intensity features of the image, effectively

suppressing impulse noise. If we denote the gray-level intensity of the image at point (x, y) as $I(x, y)$ and the square window of size $s \times s$ as W_s , then the minimum, maximum, and median intensities in this window are determined according to the equations provided in Eq. (1).

$$I_{min} = \min(W_s), I_{max} = \max(W_s), I_{med} = \text{median}(W_s) \quad (1)$$

To check if the median pixel is a non-noisy value or not, the following calculations are made according to Eq. (2):

$$A_1 = I_{med} - I_{min}, A_2 = I_{med} - I_{max} \quad (2)$$

In case $A_1 > 0$ and $A_2 < 0$, then I_{med} is taken as a valid intensity value; otherwise, the size of the window s is increased till the maximum allowed s_{max} is reached. In the second phase, the pixel $I(x, y)$ is evaluated with the help of Eq. (3):

$$B_1 = I(x, y) - I_{min}, B_2 = I(x, y) - I_{max} \quad (3)$$

In the case where $B_1 > 0$ and $B_2 < 0$, the pixel can be considered as noise-free; in any other case, it is substituted with the median value I_{med} . The filtering step to get the final pixel intensity is done according to Eq. (4):

$$I_f(x, y) = \{I(x, y), \text{ if } B_1 > 0 \text{ and } B_2 < 0 \text{ } I_{med}, \text{ otherwise}\} \quad (4)$$

The two-phase operation of AMF removes salt-and-pepper noise completely, while simultaneously preserving important textural boundaries and lesion details necessary for disease region analysis.

a.ii Contrast enhancement using CLAHE

After reducing noise, CLAHE is used to enhance local contrast and highlight disease patterns. The image that has been filtered $I_f(x, y)$ is sliced into context areas or tiles $T_{i,j}$, and the histogram of every tile is calculated as explained in Eq. (5):

$$h_{i,j}(k) = \sum_{(x,y) \in T_{i,j}} \delta(I_f(x, y) - k) \quad (5)$$

Where $\delta(\cdot)$ illustrates the Kronecker delta function, and k stands for gray levels. The cumulative distribution function (CDF) for each tile is acquired by using Eq. (6):

$$CDF_{i,j}(k) = \sum_{t=0}^k \frac{h_{i,j}(t)}{N_{T_{i,j}}} \quad (6)$$

In this case, $N_{T_{i,j}}$ stands for the sum of pixels in the tile $T_{i,j}$. To curb noise amplification, the histogram is cut off at a certain level C_L which is given in Eq. (7):

$$h'_{i,j}(k) = \min(h_{i,j}(k), C_L) \quad (7)$$

Subsequently, the pixel intensity of the enhanced image for each location is estimated by enlarging the clipped CDF as per Eq. (8):

$$I_c(x, y) = \text{round}(CDF_{i,j}(I_f(x, y)) \times (L - 1)) \quad (8)$$

Where L denotes the number of gray levels (usually $L = 256$). Bilinear interpolation in neighboring tiles guarantees seamless transitions and avoids the emergence of block artifacts.

a.iii Final pre-processed image

Mathematically, the total transformation that was executed on the UAV image is represented as a series of the AMF and CLAHE operators, together with Eq. (9):

$$I_p(x, y) = \text{CLAHE}(\text{AMF}(I(x, y))) \quad (9)$$

The image that was finally pre-processed $I_p(x, y)$ is characterized by very high local contrast and

very little noise, which in turn allows the details of the affected areas to be seen clearly. The improved image quality enhances the reliability and discrimination power of the following deep learning modules, which are aimed at disease feature extraction and classification. Sample pre-processed images are shown in Figure 2.

b. Feature extraction using Swin Transformer

The improved UAV images $I_p(x, y)$ are sent to a Swin Transformer backbone after pre-processing for extracting the spatial-spectral features that are most discriminative, hence showing the condition of the crops. The Swin Transformer, a hierarchical vision transformer architecture, successfully captures both local and global dependencies through the use of shifted window multi-head self-attention (SW-MSA), thereby eliminating the drawbacks of traditional CNN models in terms of computational inefficiency and limited receptive field. The network can, therefore, distinguish between the different scales of disease symptoms-leaf spots, blight patches, and color distortions-throughout a single attention framework. The hierarchical feature extraction yields multi-level semantic representations, which are crucial for robust disease classification, as shown in Figure 3.

The Swin Transformer in Figure 3 consists of four hierarchical stages that operate according to their fundamental design by using shifted window self-attention and feed-forward networks (FFNs) and normalization. The stages operate through the same transformer blocks which they combine with patch merging layers to create their functionality. The four-layer design enables progressive feature abstraction which allows the model to capture low-level textures in early stages and high-level disease patterns in deeper stages. The hierarchical structure improves multi-scale feature learning which results in better classification accuracy.

b.i Patch partition and linear embedding

Initially, the image was pre-processed $I_p(x, y) \in R^{H \times W \times 3}$ is divided into non-overlapping segments of size $P \times P$. Then, each segment is flattened and embedded in a D -dimensional space according to the equation in Eq. (10).

$$z_0 = [x_p^1 E; x_p^2 E; \dots; x_p^N E], E \in R^{(P^2 \cdot 3) \times D} \quad (10)$$

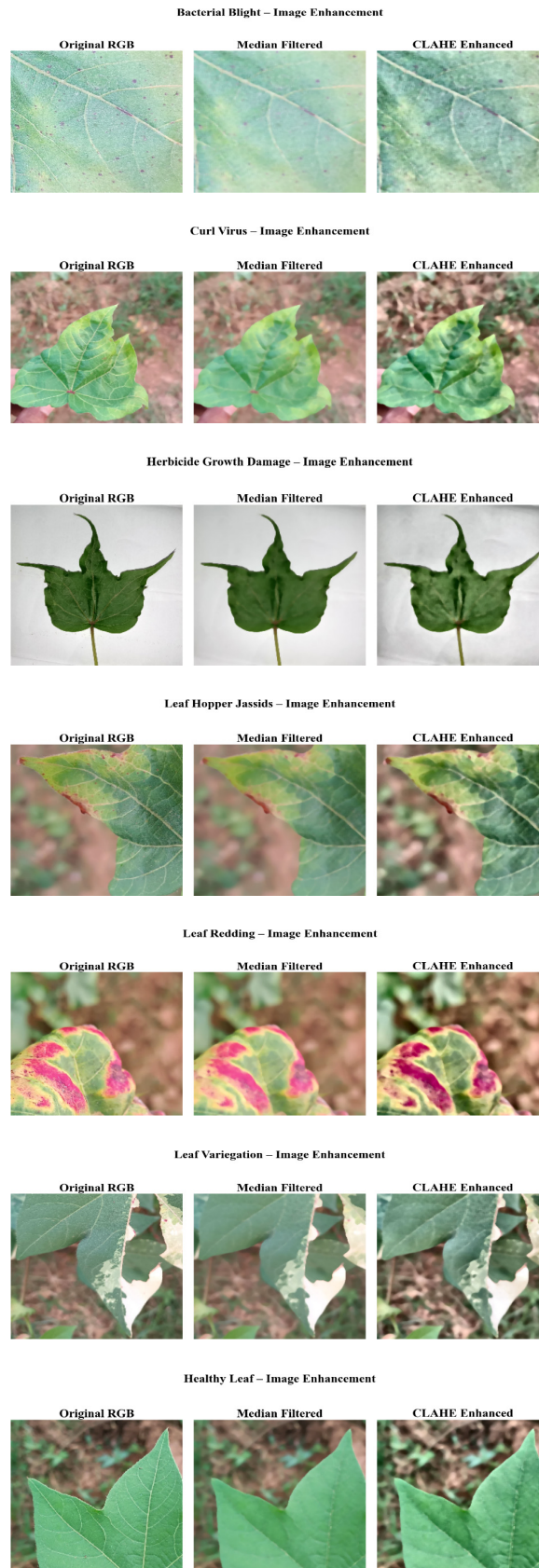


Figure 2: Sample pre-processed images for different disease types. CLAHE, contrast-limited adaptive histogram equalization.

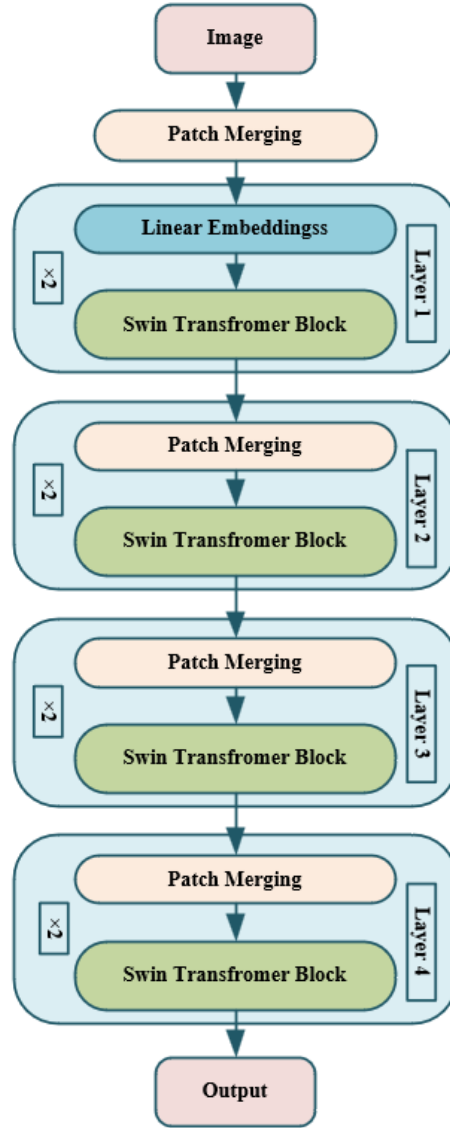


Figure 3: Architecture of the Swin Transformer model.

Here, $N = \frac{H \times W}{P^2}$ indicates the sum of patches,

and E is the learnable embedding matrix. The sequence z_0 . The input that is created is then put through the first Swin Transformer block.

b.ii SW-MSA

The Swin Transformer differs from conventional Vision Transformers in that it does not compute attention globally, but rather limits self-attention computation to non-overlapping local windows. For the specific window w , the self-attention computed for every head is as shown in Eq. (11):

$$Attention(Q, K, V) = \text{Softmax} \left(\frac{QK^T}{\sqrt{d_k}} + B \right) V \quad (11)$$

Where $Q = XW_Q$, $K = XW_K$, and $V = XW_V$ represent the projections for query, key, and value, respectively; d_k is the size of the key vectors, and B denotes the relative position bias. To enlarge the receptive field, the window configuration is alternately shifted by half the window size in each layer. This shift enables information flow across windows, thereby enhancing global context understanding without incurring any additional computational cost. The final result of the shifted window attention is determined as shown in Eq. (12):

$$z' = SW - \text{MSA}(LN(z)) + z \quad (12)$$

Where $LN(\cdot)$ denotes Layer Normalization, and this residual connection aids gradient stability during training.

b.iii FFN and patch merging

The result z' is fed to a two-layer FFN for the purpose of extracting features by adding a non-linearity. The FFN transformation is specified as in Eq. (13):

$$z'' = \text{FFN}(LN(z')) + z' \quad (13)$$

The FFN consists of two linear layers and a GELU activation function. To obtain the hierarchical information, the Swin blocks are linked by a patch merging layer that combines the features of adjacent patches, reducing the spatial resolution by applying a linear projection while increasing the channel dimension. This merging process is represented mathematically in Eq. (14):

$$z_{l+1} = \text{Concat}(z_{2i,2j}, z_{2i+1,2j}, z_{2i,2j+1}, z_{2i+1,2j+1})W_m \quad (14)$$

In this equation, W_m stands for the learnable linear projection matrix, and l is the layer index. The hierarchical representation enables a gradual transition from low-level texture signals to high-level semantic patterns relevant to crop diseases.

b.iv Final feature map generation

Following through numerous hierarchical stages, the globally averaged output of the last Swin Transformer

layer is deemed to be the final feature descriptor F , which is represented as in Eq. (15):

$$F = \frac{1}{N} \sum_{i=1}^N z_i^{(L)} \quad (15)$$

Where $z_i^{(L)}$ is used to refer to the token representation of the i^{th} position in the last layer L . The resulting feature vector $F \in R^{D_L}$ is able to encode, in an effective manner, the entire range of multi-scale spatial dependencies, leaf color gradients, and shape deformations that are characteristic of healthy and diseased crop conditions.

b.v Significance in UAV-based disease detection

The hierarchical Swin Transformer (as defined via Eqs. (10)–(15)) is characterized by its ability to fuse local and global interactions throughout the UAV-obtained images, resulting in better discrimination power than CNN-based models. Its window-shifting technique guarantees that even the tiniest lesion patterns or uneven textures from fungal or bacterial infections are completely and accurately captured. As a result, the feature vectors F thus extracted are fed into the classification module that follows for distinguishing particular crop diseases. Figure 4 shows the visualization of features.

c. Disease classification stage

The Swin Transformer encoder provided the extracted feature vector F that represented the

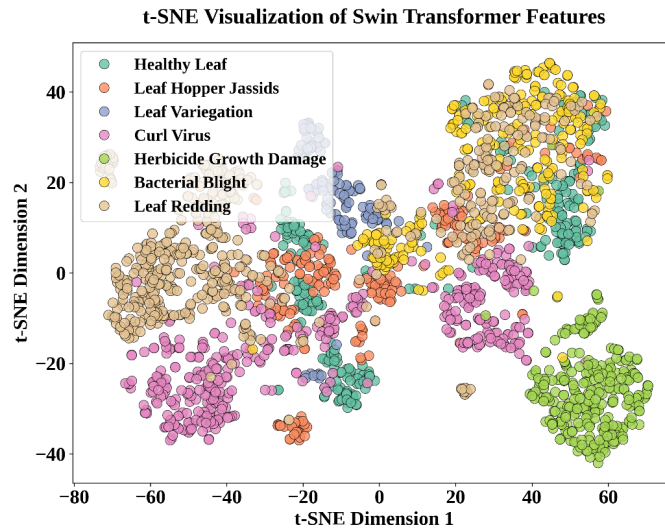


Figure 4: t-SNE visualization of Swin Transformer features.

spatial and texture information of crop leaves. Yet, the proposed framework accepts a hybrid ConvNeXt-V2 and Squeeze-and-Excite Bi-LSTM (SE-BiLSTM) model to capture not only the sequential dependencies among patch-level features but also to highlight the discriminative regions. The hybrid design has successfully integrated convolutional refinement and temporal correlation learning, making it proficient and accurate in identifying crop diseases, even under varying illumination and imaging conditions.

c.i ConvNeXt-V2 feature refinement

The ConvNeXt-V2 block functions as an extremely lightweight spatial refiner, retaining both hierarchical spatial information and enhancing deep semantic representations. ConvNeXtV2 operates depth-wise convolution over the Swin Transformer features $F \in R^{N \times D_L}$ as input, followed by normalization and activation afterward. The resultant output features F_c are then formed as in Eq. (16).

$$F_c = \sigma(LN(W_d * F + b_d)) \quad (16)$$

Where W_d denotes the kernel for the depth-wise convolution; $*$ is the symbol for convolution; b_d is used for the bias term; $LN(\cdot)$ represents Layer Normalization; and $\sigma(\cdot)$ denotes the GELU activation function. The process makes the local spatial smoothness better, but at the same time preserves the discriminative boundaries of lesions that are crucial for the recognition of the patterns of infection spread.

c.ii Squeeze-and-Excite attention mechanism

In order to enhance the network's representational capability, a Squeeze-and-Excite (SE) module is utilized at the output of ConvNeXt-V2. This technique, as explained in Eqs (17) and (18), dynamically reweights feature channels depending on the global context:

$$s = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W F_c(i, j) \quad (17)$$

$$w = \sigma(W_2 \delta(W_1 s)) \quad (18)$$

Where s is the global descriptor for each channel computed through global average pooling, W_1 and W_2 are the matrices of weights that are subject to learning, $\delta(\cdot)$ is the ReLU activation function, and $\sigma(\cdot)$ is

the Sigmoid function. The feature map F_s which has been recalibrated, is obtained through scaling the channels as per Eq. (19):

$$F_s = w \odot F_c \quad (19)$$

Where $\odot$ denotes the operation of element-wise multiplication. This type of adaptive weighting assigns greater importance to disease-specific features (for instance, regions that are discolored or have spots) while simultaneously reducing the background information that is not relevant.

c.iii Bi-LSTM for temporal dependency modeling

Due to sequential scanning or progressive leaf patterns, the features derived from UAV imagery show inter-patch dependencies; thus, the SE block is followed by a BiLSTM network. The BiLSTM captures both forward and backward contextual correlations over the spatial sequences. As an illustration, the sequence $\{F_s^1, F_s^2, \dots, F_s^T\}$ is used for computing the forward and backward hidden states in the following manner:

$$\overrightarrow{h}_t = LSTM_f(F_s^t, \overrightarrow{h}_{t-1}), h_t \leftarrow LSTM_b(F_s^t, h_{t+1} \leftarrow) \quad (20)$$

The concatenated hidden representation h_t is obtained using Eq. (21):

$$h_t = [\overrightarrow{h}_t; h_t \leftarrow] \quad (21)$$

The bidirectional flow of information between the two directions helps the model identify the relations between changes in lesion and color, and thus classify with higher reliability.

c.iv Classification layer

The prediction of probabilities over K disease classes is identified by passing the BiLSTM output through a fully connected (FC) classification layer and then applying the Softmax function. The output probability vector P is calculated according to Eq. (22)

$$P_k = \frac{e^{(W_k h + b_k)}}{\sum_{j=1}^K e^{(W_j h + b_j)}}, k = 1, 2, \dots, K \quad (22)$$

Where W_k and b_k represent the weights and biases of the classification layer, respectively. The class label $\hat{y}$ predicted is found as in Eq. (23):

$$\hat{y} = \arg \max_k (P_k) \quad (23)$$

c.v Loss function and optimization

The categorical cross-entropy loss function, which gauges the discrepancy between the expected and actual disease labels, is used to optimize the model. The loss L is defined as in Eq. (24).

$$L = -\sum_{k=1}^K y_k \log(P_k) \quad (24)$$

Where y_k signifies the correct one-hot encoded label for class k . The AdamW optimizer has been applied in the optimization process to benefit from its properties of stable convergence and overfitting alleviation. The combination of ConvNeXt-V2 and SE-BiLSTM (Eqs. (16)–(24)) offers a supporting mechanism for both local feature enhancement and global temporal dependency modeling. The ConvNeXt-V2 enhances the spatial hints received from the Swin Transformer, while the SE module dynamically amplifies the significant channels. The BiLSTM operates on this enriched representation to identify disease progression patterns across different regions of the image. Consequently, the hybrid classifier outperforms the separate CNN and Transformer architectures in terms of performance metrics, making it very effective for UAV-based crop disease diagnosis. The pseudocode for the classification stage is shown in Algorithm 1.

Algorithm 1: Classification Stage of the Proposed Framework

Inputs:

$F \leftarrow$ Feature sequence extracted from Swin Transformer (dimension: $T \times D$)
 $C \leftarrow$ Number of disease classes

Output:

$p \leftarrow$ Predicted class probability distribution
 1: function CLASSIFICATION(F)
 2: # Step 1: Spatial Feature Refinement (ConvNeXt-V2 Block)
 3: $Z_1 \leftarrow$ DepthwiseConv(F , kernel size = 7, padding = same)
 4: $Z_1 \leftarrow$ LayerNormalization(Z_1)
 5: $Z_1 \leftarrow$ GELU(Z_1)
 6: # Step 2: Channel Attention (Squeeze-and-Excitation)
 7: $s \leftarrow$ GlobalAveragePooling(Z_1)

8: $z \leftarrow$ FullyConnected(s , $D \leftarrow D/r$); $z \leftarrow$ ReLU(z)
 9: $a \leftarrow$ FullyConnected(z , $D/r \leftarrow D$); $a \rightarrow$ Sigmoid(a)
 10: $Z_2 \leftarrow$ ElementWiseMultiplication(Z_1 , a) # Channel reweighting
 11: # Step 3: Sequential Dependency Learning (BiLSTM)
 12: (H_fwd , H_bwd) $\leftarrow$ BiLSTM(Z_2)
 13: $h_cls \leftarrow$ Concatenate($H_fwd[T]$, $H_bwd[1]$) # Final feature representation
 14: # Step 4: Final Classification
 15: logits $\leftarrow$ FullyConnected(h_cls , $2H \rightarrow C$)
 16: $p \leftarrow$ Softmax(logits)
 17: return p
 18: end function

d. Recommendation using the DRL model

The last component of the suggested disease detection framework utilizing UAV technology is to produce practical suggestions for precision farming, considering the results of disease classification and contextual environmental data. As soon as the disease D_i is recognized with a severity grade S_i and confidence percentage C_i , the system takes into account environmental conditions, for example, temperature (T) and humidity (H), in order to calculate an urgency factor U_i , which ranks the need for corrective actions. The urgency is formulated with mathematics as given in Eq. (25) as follows,

$$U_i = \alpha \cdot S_i + \beta \cdot C_i + \gamma \cdot \left(\frac{H}{100} \right) + \delta \cdot \left(\frac{T - T_{opt}}{T_{opt}} \right) \quad (25)$$

Where $\alpha, \beta, \gamma, \delta$ are partially determined through empirical methods and T_{opt} refers to the specific temperature preferred for that crop. A larger U_i signifies a priority for immediate intervention. The recommendation module for each diagnosed disease applies a rule-based decision matrix (RDM) that ties together disease type, degree, and environmental stressors to the best management measures. The Recommendation Score (R_j) associated with each possible action A_j is calculated using Eq. (26) as,

$$R_j = w_1 \cdot E_j + w_2 \cdot (1 - Cost_j) + w_3 \cdot C_i \quad (26)$$

Where E_j stands for the action's expected efficiency (like reducing fertilizer, increasing irrigation, or spraying the leaves), $Cost_j$ is the cost of implementation normalized, and w_1, w_2, w_3 are importance weights. The actions that have the highest R_j and U_i which is the urgency factor that surpasses a certain

limit ($U_i > U_{thr}$) are marked as “Immediate Action Required,” The pseudocode for the recommendation model is given in Algorithm 2.

Algorithm 2: Recommendation Stage Model

Inputs:

$\hat{y} \leftarrow$ Predicted disease class
 $s \leftarrow$ Disease severity score (0–1)
 $c \leftarrow$ Model confidence score (0–1)
 $E \leftarrow$ Environmental parameters (e.g., temperature, humidity)
 $A \leftarrow$ Set of possible treatment actions $\{a_1, a_2, \dots, a_M\}$
 $Eff(a) \leftarrow$ Estimated treatment effectiveness for action a
 $Cost(a) \leftarrow$ Normalized cost of action a
 $\theta_R \leftarrow$ Minimum recommendation score threshold
 $\tau_U \leftarrow$ Urgency threshold value

Output:

$\Pi \leftarrow$ Ranked list of recommended actions

```

1: function URGENCY( $\hat{y}, s, c, E$ )
2:  $U \leftarrow \alpha_1 \cdot s + \alpha_2 \cdot c + \alpha_3 \cdot EnvironmentImpact(E, \hat{y})$ 
3: return  $U$ 
4: end function

5: function SCORE( $a, \hat{y}, s, E$ )
6:  $R \leftarrow w_1 \cdot Eff(a) - w_2 \cdot Cost(a) +$ 
    $+ w_3 \cdot ContextSuitability(a, \hat{y}, E)$ 
7: return  $R$ 
8: end function

9: function RECOMMEND( $\hat{y}, s, c, E, A$ )
10:  $U \leftarrow URGENCY(\hat{y}, s, c, E)$ 
11:  $\Pi \leftarrow$  empty list
12: for each action  $a$  in  $A$  do
13:  $R_a \leftarrow SCORE(a, \hat{y}, s, E)$ 
14: if  $R_a \geq \theta_R$  then
15: if  $U \geq \tau_U$  then
16: Label( $a$ )  $\leftarrow$  “Immediate”
17: else
18: Label( $a$ )  $\leftarrow$  “Recommended”

```

```

20: end if
21: Add ( $a, R_a, Label(a)$ ) to  $\Pi$ 
22: end if
23: end for
24: if  $\Pi$  is empty then
25: return (“Monitor and Reassess”)
26: end if
27: Sort  $\Pi$  by  $R_a$  in descending order
28: return  $\Pi$ 
29: end function

```

The recommendation system dynamically updates its suggestions by combining knowledge of disease physiology and environmental analytics. For example, if bacterial blight is found in a humid environment, the solution is to remove the infected parts of the plant and apply foliar micronutrients to delay the growth of the fungus. Moreover, the same operation is conducted for leaf discoloration caused by a nutrient imbalance; the system recommends soil amendments, including a reduction in slow-release nitrogen fertilizer, plus adjustments to the irrigation schedule to aid the plant's recovery. This versatile recommendation system enables the model to not only recognize and categorize agricultural diseases with utmost accuracy but also to break down the findings into data-informed, economical, and context-sensitive agronomic measures, thereby making precision agriculture more effective in real-life applications.

d.i Validation of the DQL recommendation module

The research used a simulated decision-validation strategy to assess how well the DQL recommendation module performs its tasks. The researchers tracked the agent's reward patterns and action selection patterns during training sessions, which showed that the system learned its intended behavior across different disease levels and environmental factors. The recommendation outputs showed consistent alignment with agronomic rules encoded in the decision matrix, particularly in high-severity disease scenarios where urgent intervention actions were prioritized. The DQL module demonstrated its ability to create effective management recommendations, which maintain contextual awareness even though the study did not include actual field testing. The research team will conduct future studies that use field testing to obtain quantitative data about yield responses and cost-benefit analysis.

IV. Result and Discussion

The present research employed a model implemented in Python, utilizing two of the most popular machine learning libraries: TensorFlow and Scikit-learn. The model was trained, evaluated, and optimized using this approach, resulting in strong performance across multiple evaluation metrics. The dataset was divided into two parts: a training set and a testing set. The former consisted of 80% of the data for training, and the latter 20% was reserved for testing. During training, the model's accuracy, loss, and other metrics were improved iteratively with different epochs. After training, the model was tested on the test set to evaluate its generalization ability, thus guaranteeing reliable performance on unseen data. The Python toolset enabled the computation of key performance indicators, such as precision, recall, F1-score, and ROC-AUC, providing a different perspective on the model's performance and efficiency.

The designed framework requires multiple deep learning components, which creates challenges for both computing and system implementation. The Swin Transformer achieves decreased processing requirements through its shifted window attention system, which performs better than standard global self-attention systems. ConvNeXt-V2 functions as a basic lightweight enhancement system, while SE-BiLSTM processes short feature streams instead of working with complete image data, which enables better management of inference expenditure. The system operates as an edge-based solution, which enables UAVs to send their captured images to nearby ground stations or edge GPUs for immediate processing. The framework achieves practical agricultural use because its implementation enables almost real-time processing capabilities.

a. Dataset description

Researchers and practitioners can greatly benefit from the cotton leaf disease identification dataset, as well as the entire agricultural industry, providing them with the necessary tools and insights to tackle issues linked to cotton leaf diseases effectively. One of the key benefits of the dataset is early detection, which is ensured by the precise detection and classification of diseases in the cotton leaves, thereby interlinking farmers and crop management, allowing for timely actions and more efficient application of crop management measures. Furthermore, it supports innovation in agricultural research through the machine learning algorithms and disease detection

techniques that it has been developing. The collection consists of pictures taken at the National Cotton Research Institute field in Gazipur during various stages of the cotton leaf disease, following a careful selection process. The Redmi Note 11s smartphone was used to take the pictures, which show the disease manifestations in different sizes. The field surveys conducted from October 2023 to January 2024, with the assistance of local specialists, yielded excellent images even in challenging situations, such as varying lighting conditions. A total of 2,137 images were taken, classified into seven categories, and these showed different states of cotton leaves, including those infected with the curl virus, bacterial blight, and those with healthy leaves. The different classes correspond to various issues that threaten the growth of cotton plants, including insects, diseases, and drought. The dataset undergoes a lengthy sequence of processing and enhancement operations, which include tagging, data purification, and the application of enhancement methods such as lighting and rotation. Thus, the dataset comprises 7,000 enhanced images in addition to 2,137 original ones, demonstrating the reliability of deep learning methods in accurately diagnosing and classifying cotton leaf diseases [34].

The SAR-CLD-2024 dataset focuses on cotton leaf diseases, while the framework learns visual representations that include common texture irregularities, lesion patterns, and color distortions, which appear in multiple crop diseases. The model enables transfer learning to other crops through its crop-specific dataset fine-tuning process. The hierarchical Swin Transformer and hybrid classification architecture enable the framework to detect cross-crop diseases through its ability to capture multi-scale spatial features which extend beyond cotton leaf detection.

b. Performance comparison

The graphical representation of the model's performance metrics shows the four main evaluation metrics: Accuracy, Precision, Recall, and F1-Score. The score of each metric is indicated by a vertical bar, which can be read from Figure 5.

The optimal score for Accuracy is 0.9930, whereas the precision score is 0.9929, which is almost equal. Recall and F1 scores are a bit lower, with scores of 0.9928 each. All metrics show almost the same performance, which means that the model is very precise and works very well with all the measures. The colors of the bars are distinctly different, making it very easy to distinguish the metrics apart.

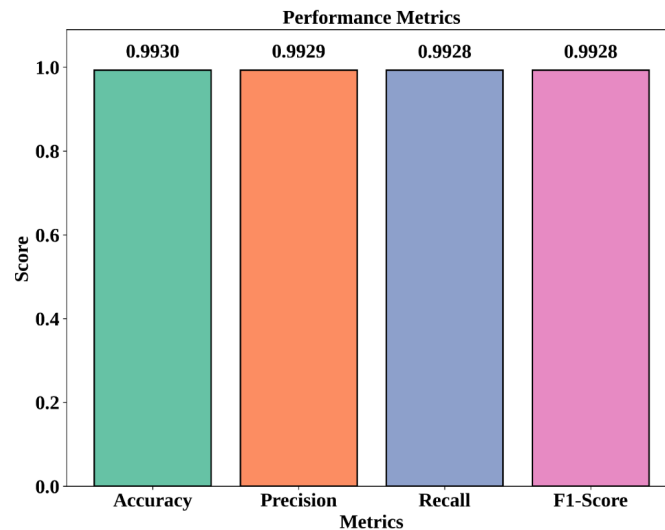


Figure 5: Performance metrics comparison.

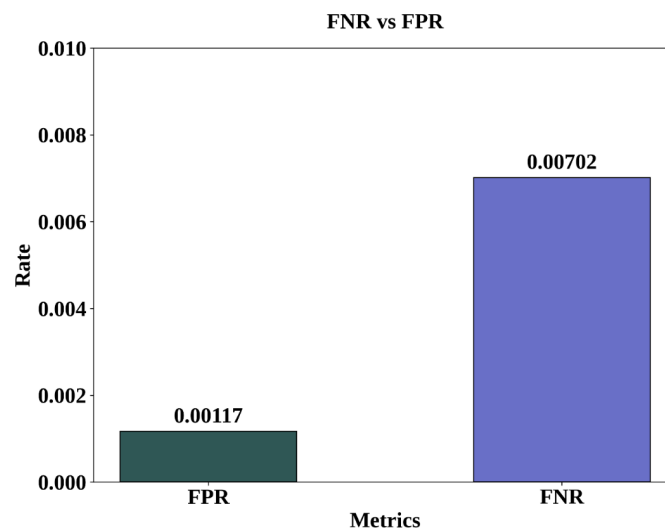


Figure 6: Comparison of FNR and FPR.

The comparison of FPR and FNR in error rates is shown in Figure 6. The FPR is represented by a tiny green bar, indicating a value of 0.00117, which represents a relatively low occurrence of false positives. The FNR, on the other hand, is represented by a large purple bar, which is valued at 0.00702, hence a much higher rate of false negatives. Therefore, it can be concluded that the model has a very low false positive rate; however, the problem with the higher rate of false negatives remains, which may negatively affect the model's performance in certain cases.

In Figure 7, a multi-class precision-recall (PR) curve is presented, with different plant conditions

represented by individual curves. The horizontal axis signifies recall, while the vertical axis shows precision. The curves for the various conditions being analyzed are shown in different colors. The conditions that exhibit a perfect performance with an average precision (AP) of 1.00000, are "Curl Virus," "Healthy Leaf," "Herbicide Growth Damage," and "Leaf Variegation," which are represented by the straight lines in the upper right corner. On the other hand, "Bacterial Blight" (AP = 0.99743) and "Leaf Hopper Jassids" (AP = 0.99486) have slightly lower AP values but still yield high precision and recall. This implies that the model's performance is not only excellent across

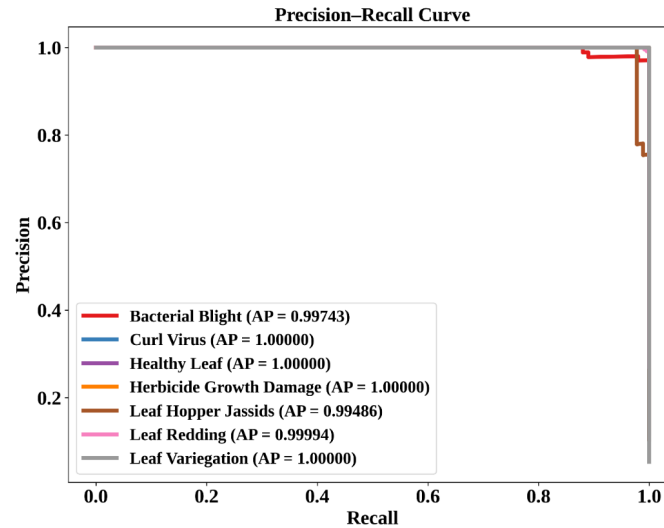


Figure 7: PR curve for plant conditions. PR, precision-recall; AP, average precision.

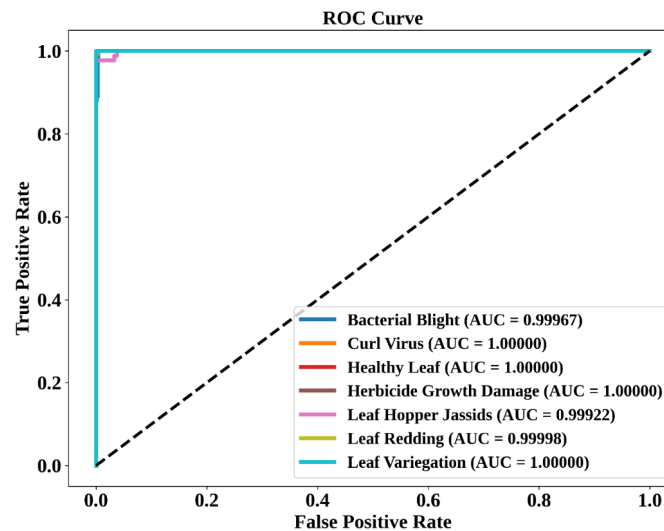


Figure 8: ROC curve for plant conditions. AUC, area under the curve.

most conditions but also with some that are slightly deviant.

A receiver operating characteristic (ROC) curve for several plant situations is shown in Figure 8; each curve represents the model's capacity to make precise assessments for a given condition. Plotting the true positive rate (TPR) on the y-axis and the FPR on the x-axis. The diagonal dashed line represents a random classifier, where the performance would match that of random guessing. All the curves that are above this line are assumed to have a better performance, and those that are close to the top-left

corner have good precision and recall. The conditions "Curl Virus," "Healthy Leaf," "Herbicide Growth Damage," and "Leaf Variegation" show a perfect performance with the area under the curve (AUC) being 1.00000, whereas in the case of "Bacterial Blight" (AUC = 0.99967) and "Leaf Hopper Jassids" (AUC = 0.99922), the AUC is still very high, but the performance is slightly lower.

Figure 9 shows the model's accuracy during training and validation for 55 epochs. The accuracy values are on the y-axis, while the epoch count is on the x-axis. The red line represents validation

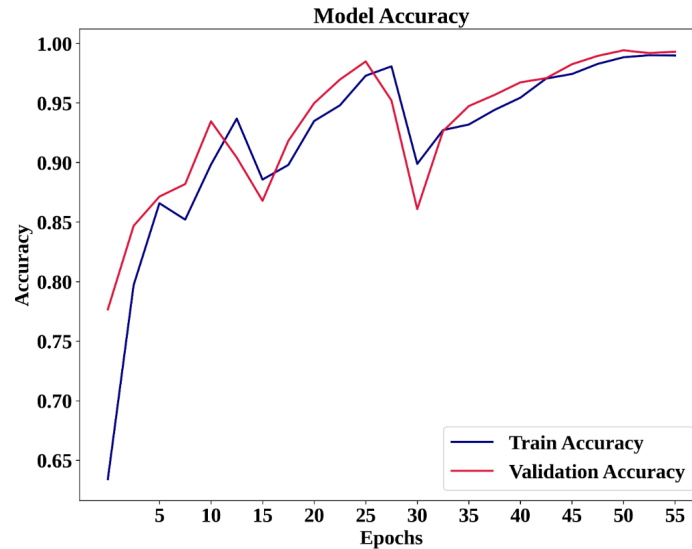


Figure 9: Model accuracy over epochs.

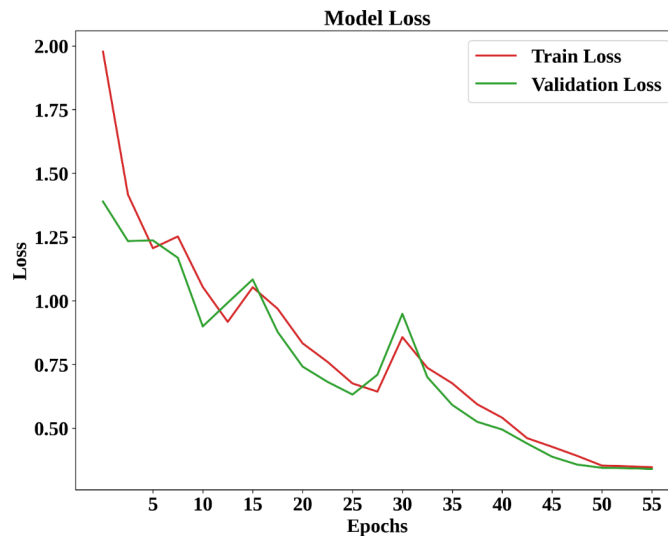


Figure 10: Model loss over epochs.

accuracy, whereas the blue line represents training accuracy. The training accuracy is typically higher than the validation accuracy, and both graphs demonstrate a steady increase in accuracy over the course of the epochs. The model's successful learning process during training is indicated by this; nevertheless, overfitting is suggested by the slight discrepancy between the two lines, as the training accuracy continues to outperform the validation accuracy in the subsequent epochs. However, as the overall performance seems to be robust, it can

still be said that the performance is satisfactory for both training and validation sets.

The model's loss throughout 55 epochs of training and validation is shown in Figure 10. The x-axis represents the number of epochs, and the y-axis again displays the loss values. The training loss is shown by a red line, while the validation loss is shown by a green line. The two transitions in loss show a very similar pattern and time-frame, suggesting that the model is not only learning but doing it pretty fast. The training loss graph starts with a high value and after that, it

Table 1: Comparison of cotton leaf disease detection models

Techniques	Accuracy	Precision	Recall	F1-Score
Proposed	0.9930	0.9929	0.9928	0.9928
[35]	0.87	0.88	0.85	0.87
[36]	0.96	0.9745	0.9689	0.9845
[37]	-	0.948	0.931	-

goes down smoothly during the training. The same is true for the validation loss which is going down gradually but not as fast as the training loss. This might lead to the conclusion that some overfitting is occurring since the training loss is dropping faster. In this case, the losses of both training and validation are getting closer to the lowest values which imply that the model performs well.

Table 1 summarizes the performance of the proposed model in comparison to three previous studies using the primary assessment metrics: Accuracy, Precision, Recall, and F1 Score. With an Accuracy of 0.9930, a Precision of 0.9929, a Recall of 0.9928, and an F1 Score of 0.9928, the suggested model was outstanding across the board. However, when compared to our model, the first quoted study [35] performs significantly worse, with an accuracy of 0.87, a precision of 0.88, a recall of 0.85, and an F1 score of 0.87. With an Accuracy of 0.96, a Precision of 0.9745, a Recall of 0.9689, and an F1 Score of 0.9845, the second study [33] shows a notable improvement, but the suggested model remains superior. The third study [34] focuses on detection and reports a Precision of 0.948 and a Recall of 0.931; however, it makes no mention of F1 Score or Accuracy numbers.

V. Conclusion

This research aimed to develop a drone-based detection framework for cotton leaf diseases, addressing the drawbacks of traditional plant monitoring techniques in agricultural lands. The framework integrates progressive image preprocessing, multi-level feature extraction using the Swin Transformer, and a hybrid ConvNeXt-V2 with SE-BiLSTM architecture for classification, resulting in highly accurate detection of disease patterns in aerial images. A DQL recommendation module was also added, designed to convert disease detection results into practical and context-sensitive crop management choices, thereby making the process even easier. By unifying these components into a single process, the system is

empowered to deliver more accurate disease diagnoses, consequently lessening the dependency on field inspections while allowing the development of intelligent decision support in precision agriculture.

The results of the trials showed that the novel method excelled in all four measures: Accuracy, Precision, Recall, and F1-Score; moreover, it was evident that it outperformed the existing methods for cotton leaf disease detection. The model maintained its ability to learn throughout the process, and at the same time, it was not affected by the changes in lighting in the field and the complexity of the background. To give an idea of the system's outstanding performance in both classification and recommendation, it can be suggested that future research could be focused on collecting a bigger dataset with multi-season UAV imagery, real-time deployment on edge devices, and the recommendation module covering the areas of adaptive fertilizer and irrigation analytics. Furthermore, the coupling of weather forecasting and drone path optimization algorithms might bring about a great improvement in both the reliability and the efficiency of operations in large-scale crop monitoring scenarios.

VI. Declaration

Data availability

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Funding statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Author contribution

Rupanjali Debbarma conceptualized the study, performed experiments, and drafted the manuscript, while Aditya Sankar Sengupta supervised the research, contributed to methodology, and reviewed the manuscript.

Ethical approval

This study does not involve human participants or animals and therefore did not require ethical approval.

Consent to participate

Not applicable

Consent to publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

- [1] Mgendi, G. (2024). *Unlocking the potential of precision agriculture for sustainable farming*. Discover Agriculture, 2(1), 87. DOI: 10.1007/s44279-024-00078-3
- [2] Khullar, V., et al. (2025). *Multiple model visual feature embedding and selection method for an efficient pest classification supporting precision agriculture*. Scientific Reports, 15(1), 31791. DOI: 10.1038/s41598-025-16942-1
- [3] Chen, X., Zhang, H., & Wong, C. U. I. (2025). *Dynamic monitoring and precision fertilization decision system for agricultural soil nutrients using UAV remote sensing and GIS*. Agriculture, 15(15), 1627. DOI: 10.3390/agriculture15151627
- [4] Vahidi, M., Shafian, S., & Frame, W. H. (2025). *Precision soil moisture monitoring through drone-based hyperspectral imaging and PCA-driven machine learning*. Sensors, 25(3), 782. DOI: 10.3390/s25030782
- [5] Sarkar, S., Osorio Leyton, J. M., Noa-Yarasca, E., Adhikari, K., Hajda, C. B., & Smith, D. R. (2025). *Integrating remote sensing and soil features for enhanced machine learning-based corn yield*

prediction in the southern US. Sensors, 25(2), 543. DOI: 10.3390/s25020543

[6] Al-Mahbashi, M., et al. (2025). *An effective approach to improving photovoltaic defect detection using the new DCD-YOLOv8s model*. Scientific Reports, 15(1), 38308. DOI: 10.1038/s41598-025-22307-5

[7] Tan, C., et al. (2024). *Detection of the infection stage of pine wilt disease and spread distance using monthly UAV-based imagery and a deep learning approach*. Remote Sensing, 16(2), 364. DOI: 10.3390/rs16020364

[8] Mo, Y., et al. (2024). *Agricultural practices influence soil microbiome assembly and interactions at different depths identified by machine learning*. Communications Biology, 7(1), 1349. DOI: 10.1038/s42003-024-07059-8

[9] Cao, Y., et al. (2024). *Phosphorus availability influences disease-suppressive soil microbiome through plant-microbe interactions*. Microbiome, 12(1), 185. DOI: 10.1186/s40168-024-01906-w

[10] Awad, D. A., Masoud, H. A., & Hamad, A. (2024). *Climate changes and food-borne pathogens: The impact on human health and mitigation strategy*. Climatic Change, 177(6), 92. DOI: 10.1007/s10584-024-03748-9

[11] Tekelchal, L., Hdgot, A., Gebregiorgis, G., & Zeweld, W. (2025). *Socioeconomic and management implications of earthen-dam irrigation practices in Ethiopia*. Sustainable Water Resources Management, 11(5), 89. DOI: 10.1007/s40899-025-01260-1

[12] Hyder, U., & Talpur, M. R. H. (2024). *Detection of cotton leaf disease with machine learning model*. Turkish Journal of Engineering, 8(2), 380–393. DOI: 10.31127/tuje.1406755

[13] Rehman, A., Akhtar, N., & Alhazmi, O. H. (2025). *Monitoring and predicting cotton leaf diseases using deep learning approaches and mathematical models*. Scientific Reports, 15(1), 22570. DOI: 10.1038/s41598-025-06985-9

[14] Faisal, H. M., et al. (2025). *Detection of cotton crops diseases using customized deep learning model*. Scientific Reports, 15(1), 10766. DOI: 10.1038/s41598-025-94636-4

[15] Aslam, A., Usman, S. M., Zubair, M., Yasin, A., Owais, M., & Hussain, I. (2025). *Multi-convolutional neural networks for cotton disease detection using synergistic deep learning paradigm*. PLOS ONE, 20(5), e0324293. DOI: 10.1371/journal.pone.0324293

[16] Vellingiri, A., Mohanasundaram, K., Tamilselvan, K. S., Maheswar, R., & Ganesh, N. (2023). *Multiple sensor based human detection robots: A review*. International Journal on Smart Sensing and Intelligent Systems, 16(1). DOI: 10.2478/ijssis-2023-0009

- [17] Singh, J., Singh, A. K., & Chauhan, S. S. (2025). *Enhanced edge-based steganography using image segmentation and random LSB substitution for secure data hiding*. International Journal on Smart Sensing and Intelligent Systems, 18(1). DOI: 10.2478/ijssis-2025-0049
- [18] Casas, E., Arbelo, M., Moreno-Ruiz, J. A., Hernández-Leal, P. A., & Reyes-Carlos, J. A. (2023). *UAV-based disease detection in palm groves of Phoenix canariensis using machine learning and multispectral imagery*. Remote Sensing, 15(14), 3584. DOI: 10.3390/rs15143584
- [19] Fei, S., et al. (2023). *UAV-based multi-sensor data fusion and machine learning algorithm for yield prediction in wheat*. Precision Agriculture, 24(1), 187–212. DOI: 10.1007/s11119-022-09938-8
- [20] Zhu, H., Liang, S., Lin, C., He, Y., & Xu, J.-L. (2024). *Using multi-sensor data fusion techniques and machine learning algorithms for improving UAV-based yield prediction of oilseed rape*. Drones, 8(11), 642. DOI: 10.3390/drones8110642
- [21] Das, B., & Raghuvanshi, C. S. (2025). *Advanced UAV-based leaf disease detection: Deep radial basis function networks with multidimensional mixed attention*. Multimedia Tools and Applications, 84(27), 32533–32561. DOI: 10.1007/s11042-024-20462-x
- [22] Linero-Ramos, R., Parra-Rodríguez, C., Espinosa-Valdez, A., Gómez-Rojas, J., & Gongora, M. (2024). *Assessment of dataset scalability for classification of black sigatoka in banana crops using UAV-based multispectral images and deep learning techniques*. Drones, 8(9), 503. DOI: 10.3390/drones8090503
- [23] Singh, R., Gadade, A. M., & Hussain, I. (2025). *Robust real-time strawberry maturity detection using UAV-mounted deep learning for precision agriculture*. BMC Plant Biology, 25(1), 1367. DOI: 10.1186/s12870-025-07246-7
- [24] Gokeda, V., & Yalavarthi, R. (2024). *Deep hybrid model for pest detection: IoT-UAV-based smart agriculture system*. Journal of Phytopathology, 172(5), e13381. DOI: 10.1111/jph.13381
- [25] Fu, H., et al. (2025). *A hierarchical path planning framework of plant protection UAV based on the improved D3QN algorithm and remote sensing image*. Remote Sensing, 17(15), 2704. DOI: 10.3390/rs17152704
- [26] Bokani, A., Yadegaridehkordi, E., & Kanhere, S. S. (2025). *LSTM-H: A hybrid deep learning model for accurate livestock movement prediction in UAV-based monitoring systems*. Drones, 9(5), 346. DOI: 10.3390/drones9050346
- [27] Raptis, E. K., et al. (2023). *End-to-end precision agriculture UAV-based functionalities tailored to field characteristics*. Journal of Intelligent & Robotic Systems, 107(2), 23. DOI: 10.1007/s10846-022-01761-7
- [28] Alavilli, S. K., Nippatla, R. P., Kadiyala, B., Boyapati, S., & Vasamsetty, C. (2025). *Unified Robotic Automation and AI-Driven Transformer-Guided Graph Neural Network with Hybrid 3D-CNN, BiLSTM, and Adaptive Neuro-Symbolic Fuzzy Decision Framework for Histological Subtype and Lymph Node-Aware Breast Cancer Prediction*. International Journal of Automation and Smart Technology, 15(1). doi:10.5875/j8f74e88
- [29] Khujamatov, H., Muksimova, S., Abdullaev, M., Cho, J., Lee, C., & Jeon, H.-S. (2025). *Empowering smallholder farmers with UAV-based early cotton disease detection using AI*. Drones, 9(5), 385. DOI: 10.3390/drones9050385
- [30] Saad, M. H., & Salman, A. E. (2024). *A plant disease classification using one-shot learning technique with field images*. Multimedia Tools and Applications, 83(20), 58935–58960. DOI: 10.1007/s11042-023-17830-4
- [31] Jayanthi, V., & Kanchana, M. (2025). *MSNet-LNet architecture with improved deep joint segmentation for tomato plant disease classification with multi-texton and statistical features*. Journal of Phytopathology, 173(3), e70088. DOI: 10.1111/jph.70088
- [32] Patil, B. V., & Patil, P. S. (2025). *IoT-enhanced meta-heuristic hybrid deep learning model for predicting cotton leaf diseases*. Journal of Phytopathology, 173(2), e70058. DOI: 10.1111/jph.70058
- [33] Yang, Z.-Y., Xia, W.-K., Chu, H.-Q., Su, W.-H., Wang, R.-F., & Wang, H. (2025). *A comprehensive review of deep learning applications in cotton industry: From field monitoring to smart processing*. Plants, 14(10), 1481. DOI: 10.3390/plants14101481
- [34] Bishshash, P., Nirob, M. A. S., Shikder, M. H., & Sarower, A. (2024). *SAR-CLD-2024: A comprehensive dataset for cotton leaf disease detection*, 2(1). DOI: 10.17632/b3jy2p6k8w.2
- [35] Meena, P. O. O. J. A., Gupta, P. R. I. Y. A. N. K. A., & Agarwal, D. S. (2024). *Integrated GLCM Texture Features and CNN for Automated Cotton Disease Identification*. Journal of Theoretical and Applied Information Technology, 102(21), 7589–7600.
- [36] Lakshmi, M. D. (2025). *Advancements in early detection of cotton leaf diseases in plants using multi-modal deep learning techniques*. Journal of Information Systems Engineering and Management, 10(2s), 167–175. DOI: 10.52783/jisem.v10i2s.211
- [37] Gao, L., Ran, T., Zou, H., & Wu, H. (2025). *Cotton leaf disease detection using LLM-synthetic data and DEMM-YOLO model*. Agriculture, 15(15), 1712. DOI: 10.3390/agriculture15151712