

TriAttnNet Based Deep Learning Model for Automated Cotton Pest Detection and Disease Classification

A. Mohan

amohanphd2020@gmail.com

University College of Engineering, Jawaharlal Nehru Technological University Kakinada (JNTUK)

P Chiranjeevi

Amrita Sai Institute of Science and Technology

Krishna Mohan

University College of Engineering, Jawaharlal Nehru Technological University Kakinada (JNTUK)

Article

Keywords: Spatial Generative Adversarial Networks, Hybrid Mongoose Ray Chaotic Optimization, Cotton Plant Pest identification and categorization, TriAttnNet

Posted Date: April 1st, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9110147/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Scientific Reports on June 26th, 2026. See the published version at <https://doi.org/10.1038/s41598-026-58795-2>.

TriAttnNet Based Deep Learning Model for Automated Cotton Pest Detection and Disease Classification

^{*1}A. Mohan, ²Dr P Chiranjeevi, ³Dr A Krishna Mohan

^{*1}*Research scholar, Department of Computer Science and Engineering, Jawaharlal Nehru Technological University Kakinada (JNTUK), Andhra Pradesh, India*

²*Professor, CSE Department, Amrita Sai Institute of Science and Technology, Bathinapadu, Andhra Pradesh, India*

³*Professor, CSE Department, University College of Engineering, Jawaharlal Nehru Technological University Kakinada (JNTUK), Andhra Pradesh, India*

^{*1}*Corresponding Author Email: amohanphd2020@gmail.com*

Abstract—This paper presents a deep learning model to detect cotton plant pests and classify diseases, which must overcome limited datasets, class imbalance, and feature redundancy. At the preprocessing phase, the Gaussian blur filtering and Contrast Limited Adaptive Histogram Equalization (CLAHE) are used to sharpen images by improving their clarity and contrast. To increase and diversify the data, SpGAN-based data augmentation is used to produce realistic synthetic samples. To obtain an accurate Region of Interest (RoI), an Attention-Guided Multi-Scale Residual U-Net (AGMS-UNet) is considered to segment local and global structural information. The proposed framework makes three major contributions: (i) a new attention-based feature extractor, TriAttnNet, that incorporates spatial, channel, and contextual attention to represent diseases on a fine-grained level; (ii) a new optimization strategy, Hybrid Mongoose Ray Chaotic Optimization (HMRCO), which includes chaotic strategies to better tune the parameters and explore the feature space; and (iii) classification layer with focal loss for final decision. Experimental analyses prove that the suggested method is much more effective than the current state-of-the-art models, providing a powerful and understandable solution to precision agriculture and sustainable cotton crop health monitoring. Experimental results show that TriAttnNet achieves 98.66% accuracy, 98.71% recall, and 98.81% F1-score, which is better than the state-of-the-art algorithms, such as EfficientNetB1-CBAM (96.38%) and BERT-ResNet-PSO (95.69%). The proposed system is computationally feasible and interpretable, and it is interpretable to provide a practical solution to precision agriculture and sustainable monitoring of the health of cotton crops.

Keywords - Spatial Generative Adversarial Networks, Hybrid Mongoose Ray Chaotic Optimization, Cotton Plant Pest identification and categorization, TriAttnNet.

1. INTRODUCTION

Cotton is one of the most commercially significant crops in the field of natural fibre production. There is a close relationship between the insect assaults and cotton harvesting [1,2]. Given the great degree of resemblance among the various species, the identification and categorization of crops is one of the biggest problems in agricultural areas. Keeping an eye on the pest situation at all times, particularly in the early stages, is essential for guaranteeing cotton growth and quality [3-5].

This is highly beneficial for farmers in choosing biological control techniques or efficient chemicals to stop the spread of pests. Farmers find it challenging and time-consuming to classify appropriately by hand. Crop disease is thought to be identified by the spots on the leaf, which are regarded as the primary indicator of disease presence [6-8]. Nectar is held in place by a tiny, cup-shaped structure found under each leaf on the cotton plant. Many bugs are drawn to the plant because of these deposits and its wet stems.

In recent years, many scholars have created a variety of computational models based on vision in order to address these issues in the field of agriculture [9-11]. These models cover many different technologies, including Deep learning (DL), Transfer learning (TL), Convolutional neural networks (CNN), and machine learning (ML). Typically, specialists use their naked eyes to assess the symptoms. However, in the case of large farms, this required constant professional monitoring, which is costly [12,13]. However, because farmers rely on their own experience to diagnose symptoms, a faulty diagnosis might result in erroneous management measures, including applying pesticides excessively or at the wrong time. The automatic identification of plant illnesses is a crucial area of study because it might be useful for tracking large agricultural fields and identifying disease signs as soon as they manifest on plant leaves [14,15]. It is essential to create a quick, automated, low-cost, and accurate method for identifying plant diseases.

The paper discusses the use of neural networks and image processing techniques to develop autonomous systems for early disease detection and monitoring in large crop fields. The main contribution of the paper is as follows: Proposed TriAttnNet, an attention-driven feature extraction model combining spatial, channel, and contextual attention for fine-grained symptom detection. Introduced HMRCO, a hybrid optimization algorithm that enhances parameter tuning, convergence stability, and robust feature selection.

Developed TriAttnNet, a tri-branch classification network integrating DenseNet, Improved MobileNet, and Classification with focal loss to boost accuracy and handle data imbalance. This study proposes and implements a few-shot cotton plant pest identification system. This study aims to address the inadequacies of current agricultural pest classification models by addressing hardware realization in terms of accuracy performance.

2. Related work

In 2021, Zekiwos and Bruck [16] introduced a CNN-DL-based image processing for cotton leaf disease and pest diagnosis. The goal of that work was to create a model that would improve the DL technique's ability to identify pests and cotton leaf disease. Researchers have investigated common cotton leaf diseases and pests, including leaf miners, spider mites, and bacterial blight.

In 2021, Caldeira, et. al., [17] have suggested a DL technique for the identification of cotton leaf lesions. It may be challenging for the producer to correctly identify a lesion, and the symptoms of the most common pests and illnesses cannot be distinguished in their early stages.

In 2023, Zhang et. al., [18] measured cotton verticillium wilt disease severity built using GWO-SVM and hyperspectral. As an experimental subject, this study examined the crown spectral data of cotton plants with canopy wilt disease. This study classified the cotton canopy wilt disease into five distinct classes. The continuous wavelet analysis approach was used to optimize four models: GA-SVM, GS-SVM, PSO-SVM, and GWO-SVM (mexh and db3).

In 2022, Jajja, et. al., [19] recommended an approach for detecting whitefly attacks on cotton crops using Compact Convolutional Transformers (CCTs). To facilitate future studies, agricultural areas do not yet have adequate datasets or online picture samples. To prepare for the future use of AI in agriculture in the future, fresh datasets must be created.

In 2021, Liang [20] suggested a metric learning-based few-shot categorization of cotton leaf spot diseases. This was used promptly to prevent and manage cotton illnesses. Several approaches were used to segment the disease spots on cotton leaf disease photos. These approaches were then evaluated using threshold segmentation and the Support Vector Machine (SVM) method to identify the most appropriate one.

In 2022, Memon, et. al., [21] suggested a cotton crop leaf disease identification model using Meta DL. Several cotton leaf diseases can be accurately identified using suggested meta-DL model. For this investigation, we collected field-collected photographs of cotton leaves. There were 2385 photos of both healthy and sick leaves included in the collection. The data augmentation method assisted in the number of datasets.

In 2025, Wang, et. al., [22] focused on lightweight models for crop disease detection but very few of these proved efficient for edge-device deployment. The researchers selected MobileViTv2 as a backbone because of its ability to balance local feature extraction with global context understanding. In this research, RF-Cott-Net applied an early exit mechanism to save inference time and quantization-aware training (QAT) to ensure that model accuracy was preserved after quantization.

In 2024, Bishshash et al. [23] proposed facilitating precision farming and disease management practices. An extensive dataset for cotton leaf diseases has been created. Their dataset was augmented to 7000 images, which increases the robustness of the deep-learning training. Researchers obtained an accuracy of 96.03% in Inception V3, which proved the effectiveness of the dataset in an automated detection task with

less manual inspection hindrance and thus could allow for timely and site-specific interventions in the field.

In 2022, Kumar et. al., [24] introduced the Tensor Flow ML technique for cotton disease detection. TensorFlow's Keras API was used to develop a prediction model using a CNN algorithm. The aforementioned approach finds additional applications in the creation of mobile applications that aid farmers in detecting cotton disease and suggesting pesticides for its treatment.

2.1. Problem statement

The major issues of DL methods of cotton crop disease detection include a narrow attention to specific diseases, the absence of various datasets, and the inability to detect diseases at an early stage. The models also have difficulties with generalization in different geographical and environmental conditions. Moreover, the connection with precision agriculture and its real-world implementation by farmers is not studied extensively. The challenges have to be overcome to come up with strong, accessible, and generalizable solutions capable of facilitating correct and timely detection of cotton diseases to improve sustainable crop management.

3. Proposed Methodology

The proposed TriAttnNet presents an integrated approach for cotton plant pest detection and disease classification in agriculture. By combining DL techniques with traditional image processing methods, this methodology addresses challenges such as limited data availability, model robustness, and interpretability. It begins with the pre-processing techniques such as Gaussian blur filtering, Contrast Limited Adaptive Histogram Equalization (CLAHE), and generation of synthetic training samples through Spatial Generative Adversarial Networks (Spa-GAN) and employs. Proposed TriAttnNet, an attention-driven feature extraction model combining spatial, channel, and contextual attention for fine-grained symptom detection. Introduced HMRCO, a hybrid optimization algorithm that enhances parameter tuning, convergence stability, and robust feature selection. Developed TriAttnNet, a tri-branch classification network integrating DenseNet, Improved MobileNet, and Classification with focal loss to boost accuracy and handle data imbalance. The detailed workflow of the proposed cotton plant pest detection and disease classification model is shown in Figure 1.

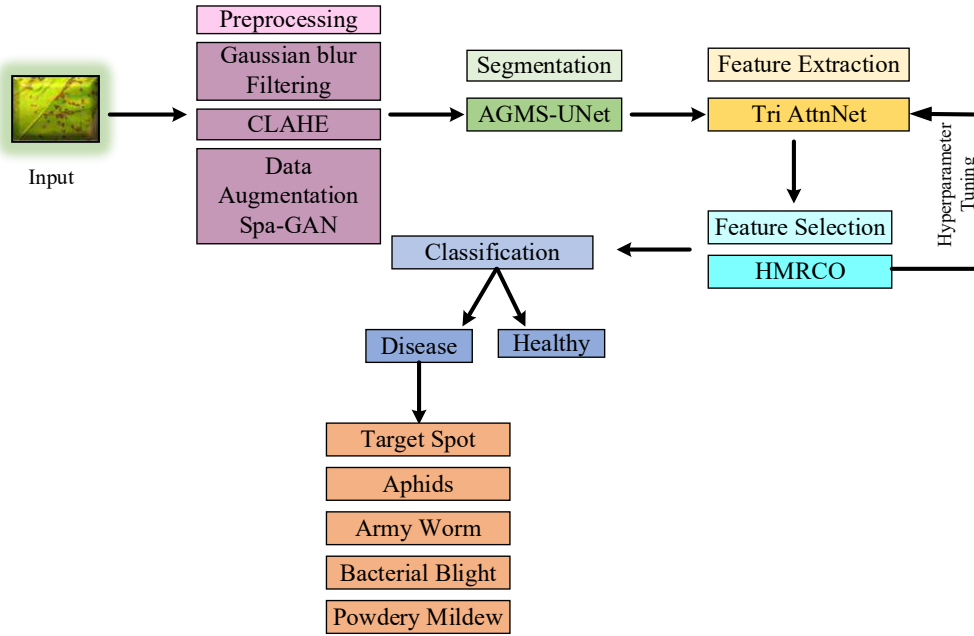


Fig. 1: Overview of Proposed Model

3.1. Data Collection

The dataset of this study is taken from the publicly available Kaggle repository, Cotton Plant Disease Dataset [25]. The data is made up of high-resolution images of the leaves of cotton plants that have been classified into various categories of healthy and diseased samples. In particular, it comprises five significant diseases of cotton plants, such as Aphids, Armyworm, Bacterial Blight, Powdery Mildew, and Target Spot, and a healthy leaf category. These classifications encompass a broad spectrum of the biotic stress conditions that are very pertinent to cotton farming and pest control.

The dataset is a wide range of samples that are required to train and validate DL models in pest detection and disease classification. Nevertheless, because of the nature of the limitations, including the unequal distribution of classes, the variability of the illumination, and the small size of the sample, additional preprocessing and augmentation measures are utilized to enhance the quality and diversity of the data to be used to train a robust model. Figure 2 shows the sample images in the dataset.

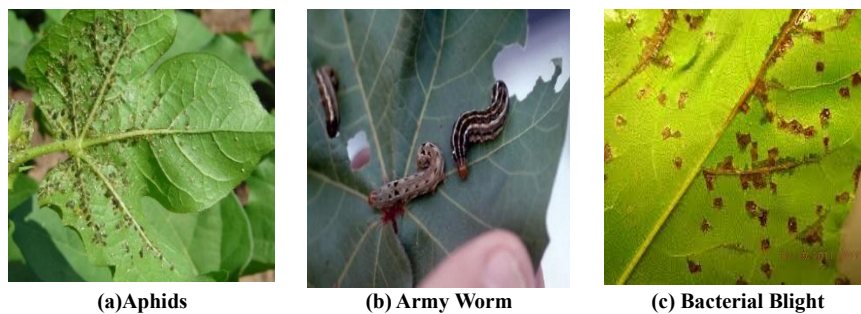




Fig. 2: Samples of cotton diseases and pests in the dataset

3.2. Preprocessing

Preprocessing is used to remove noise and enhance the quality of the image by applying a Gaussian blur filter and CLAHE to remove noise and enhance the contrast, respectively. The preprocessing pipeline followed a denoise-then-enhance strategy which eliminates the undesirable background noise, asymmetrical lighting, and brings out the disease-related areas on the cotton leaf, including pest spots, blight spots, and mildew spots.

1) Gaussian blur filtering

Gaussian blur filtering [26] is an image smoothing method that is very popular for eliminating the high-frequency noise without altering the necessary structures of the image. In contrast to a mean filter, which averages the pixels around the pixel, Gaussian blur uses a Gaussian distribution (bell-shaped curve) kernel. The kernel is defined in Eq. (1).

$$G(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \quad (1)$$

where σ represents the standard deviation of the distribution. Smaller values of σ preserve fine details, while larger values result in heavier blurring and potential information loss. Therefore, kernel size selection plays a critical role in balancing noise reduction and feature preservation.

2) CLAHE

In order to address the challenges of low visibility and unequal lighting in the images of cotton leaves, CLAHE [27] is used following the Gaussian blur filtering. In contrast to the global histogram equalization, which can over-enhance noise, CLAHE operates on small contextual blocks (tiles) and uses a contrast limiting factor to avoid enhancement of irrelevant details. The transformation operation is expressed in the equation. (2).

$$s = T(r) = \int_0^r P_r(w)dw \quad (2)$$

where $P_r(w)$ is the probability density function of pixel intensities within a localized window. By adaptively redistributing pixel intensities, CLAHE enhances disease-specific patterns like pest damage, ensuring robust feature extraction in the subsequent steps.

3) Data Augmentation using Spatial Generative Adversarial Networks (Spa-GAN)

Despite preprocessing, the dataset still suffers from a limited sample size and class imbalance, which leads to model overfitting and reduced generalization. To address this, Spa-GAN [28] is applied for data augmentation. Unlike conventional GANs, Spa-GAN integrates a spatial attention mechanism that ensures the generator focuses on disease-relevant regions of cotton leaves (blight lesions, pest infection spots, mildew patches) while generating synthetic images. This leads to higher-quality augmented data with enhanced variability and structural consistency.

$$L_{Spa-GAN} = \min_G \max_D \left[E_{x \sim p_{data(x)}} [\log D(x)] + E_{z \sim p_{z(z)}} [\log (1 - D(G(Z)))] + \lambda \|M \odot x - G(z)\|_1 \right] \quad (3)$$

Where G specifies the Generator network, D denotes the Discriminator network, $x \sim p_{data(x)}$ refers to the real cotton leaf image samples, $z \sim p_{z(z)}$ express the random noise input, M indicates the spatial attention mask focusing on disease-affected regions, $\odot$ specifies the element-wise multiplication, λ defines the weighting factor for spatial attention regularization.

3.3. Segmentation

AGMS UNet and TriAttnNet are applied for different purposes in the proposed pipeline. AGMS UNet has conducted a process of segmentation, which isolates areas of cotton leaves that are pest and disease-infected to allow proper localization. After identifying these regions, TriAttnNet serves as the supporting element of feature extraction and classification, whereby spatial, channel, and contextual attention are used to take in the fine-grained disease features and identify the disease type. Convolutional U-Net networks typically become blurred, and thus fail to address the multi-scale context; hence, the introduced AGMS UNet advances detection of the lesions by controlling the local, the global, and attention-generating refinement. Figure 3 displays the architecture of the AGMS-Unet segmentation model.

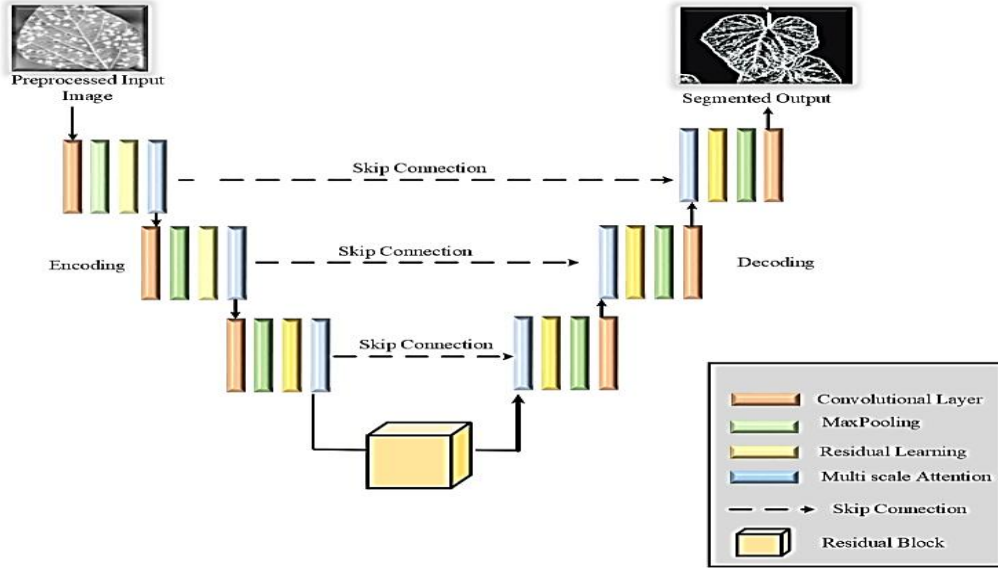


Fig. 3: Proposed Architecture of AGMS-Unet

The Residual Learning is also incorporated to combat the vanishing gradients, allow deeper architecture, and save the necessary low-level information (mildew spots, blight edges). In contrast to simple U-Net blocks, in residual connections, the propagation of features is possible, which increases the convergence and stability. Let the input image be expressed in Eq. (4), l refers to the convolutional feature map, and $H(\cdot)$ specifies the residual transformation with parameters w_l . This formulation preserves spatial consistency and stabilizes gradient flow. Residual Encoding, at each encoding stage l . The feature representation is represented in the equation. (5).

$$I \in R^{H*W*C} \quad (4)$$

$$R_l = F_l + H(F_l; w_l) \quad (5)$$

Multi-Scale Fusion is used because pest and disease symptoms have various spatial scales, like some (aphids, mildew) are fine-grained, whereas others (blight, target spot) are large-scale. The conventional single-scale encoders cannot encode such variability, and our design adaptively combines several receptive fields. Let $U_s(\cdot)$ refers to the down-sampling to a uniform resolution and β_s denotes the adaptive fusion weights. This ensures simultaneous capture of local lesion patterns and global structural spread. Features from different scales s are illustrated in eqn. (6).

$$F_{ms} = \sum_{s=1}^S \beta_s \cdot U_s(F_s) \quad (6)$$

Attention Guidance makes the model concentrate on the areas of the disease, but not on the irrelevant background (soil, veins, or artifacts of the illumination). AGMS-UNet uses channel and spatial attention to dynamically focus on the areas of greatest relevance and the features of greatest discriminative value

to make a correct segmentation. Let $\odot$ denotes element-wise multiplication. This ensures that segmentation highlights lesion boundaries and infected patches, while suppressing redundant background. Channel Attention is defined in Eqn. (7) and Spatial Attention is illustrated in eqn. (8). The refined feature is expressed in eqn. (9).

$$\alpha_c = \sigma(W_c \cdot \text{GAP}(F_{ms}) + b_c) \quad (7)$$

$$\alpha_s = \sigma(W_s * F_{ms} + b_s) \quad (8)$$

$$F' = \alpha_c \odot F_{ms} + \alpha_s \odot F_{ms} \quad (9)$$

Decoder and Segmentation Mask, Skip connections link encoder features with the decoder to recover spatial resolution. Let $\psi(\cdot)$ refers the decoding transformation. The mask M denotes the segmented diseased or pest-infected regions on the cotton leaf. The segmentation mask is derived in Eqn. (10).

$$M = \psi(F'; \theta_d), M \in [0,1]^{H*W} \quad (10)$$

BCE Loss Optimization, BCE Loss is a loss that guarantees accuracy in the classification of pixels, Dice Loss is a loss that is more resistant to imbalance in the classes, and sharpens the boundaries of the lesions. In order to trade off pixel-wise accuracy and structural overlap, a hybrid loss is expressed in eqn. (11).

$$L_{seg} = \lambda_1 \cdot L_{BCE} + \lambda_2 \cdot L_{Dice} \quad (11)$$

3.4. Proposed TriAttnNet for Feature Extraction

To achieve a discriminative and powerful feature representation of cotton leaf images, a new TriFusionNet is proposed, a fusion of the hierarchical features of ResNet-152 and InceptionV3, based on an attention-guided fusion mechanism. As opposed to traditional single-network extractors. TriFusionNet is a single architecture that combines the multi-scale convolutional representations with residual learning, which guarantees depth and diversity of the feature representation. Learns deep hierarchical characteristics using residual connections, which reduces vanishing-gradient problems and preserves fine-grained structural characteristics as a form of expression, as in eqn (12).

$$f_{res} = \phi_{ResNest}(I_{seg}) \quad (12)$$

Provides multi-scale features through asymmetric convolutions and factorization techniques that reduce computational overhead while retaining discriminative power is explained in Eqn (13).

$$F_{inc} = \phi_{Inception}(I_{seg}) \quad (13)$$

Instead of direct concatenation, an attention mechanism is introduced to weigh the importance of features from each backbone. The fusion is expressed in eqn (14), Let value α, β indicates the adaptive attention

weights learned during training. This ensures that only the most informative features are emphasized. The optimized deep feature set is specified in the eqn (15). Figure 4 illustrates the architecture of the developed TriAttnNet feature extraction model.

$$F_{deep} = \alpha \cdot F_{res} + \beta \cdot F_{inc} \quad (14)$$

$$F_{TriFusionNet} = \Phi_{fusion}(F_{res}, F_{inc}) \quad (15)$$

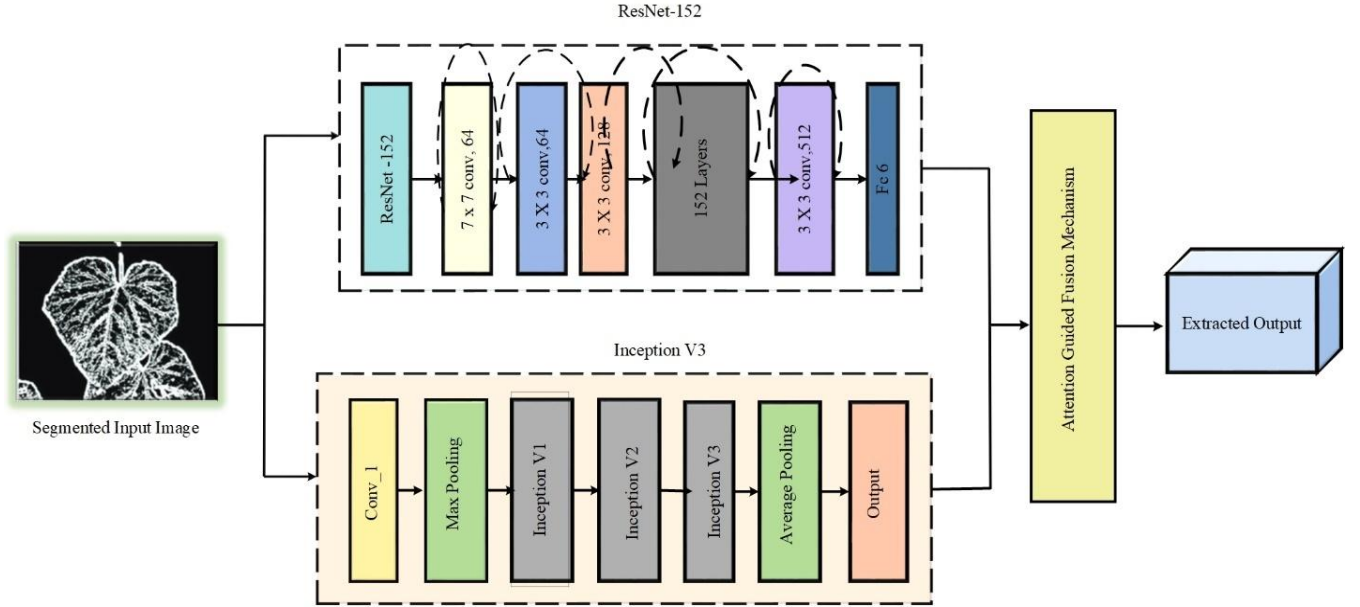


Fig. 4: Architecture of the developed TriAttnNet feature extraction model

The resulting fused representation of $F_{TriFusionNet}$ It is a small but highly discriminative feature space, and it is a good balance between depth and diversity. The network reduces redundancy by adaptively combining the residual and multi-scale features, and is resistant to changes in illumination, scale, and background noise. The given mechanism enables TriFusionNet to detect both the small-scale lesion patterns and the larger structural deformations of cotton leaves, which are essential to secure disease detection. The attention-guided fusion not only increases the accuracy of the classification but also increases the interpretability by indicating which stream of features is playing a bigger role in a particular case. The dropout and normalization layers also provide regularization, and overfitting is minimized when these layers are trained on small agricultural datasets. The experimental performance of TriFusionNet proves that it is better than individual backbones and simple fusion techniques, and proves its effectiveness in working with complex, high-dimensional data. Moreover, it can be adapted to fit other crop diseases beyond cotton, highlighting its generalizability. On the whole, TriFusionNet provides a well-developed and generalizable model that fills the gap between inefficient features by hand and DL overfitting, and becomes a new standard of smart agriculture disease detection.

3.5. Feature Selection using HMRCO

To enhance the interpretability and performance of the model, we combined the Dwarf Mongoose optimization (DMO) with the suggested Elite Chaotic Manta Ray Algorithm (ECMRA) to optimally select the features. The algorithm is designed to optimize a population of potential solutions represented by a matrix S , where each row corresponds to a solution, and each column represents a trait or decision variable. The algorithm consists of three main groups: Alpha Group, Scout Group, and Babysitters. As stated in Eq. (16): Initializing the population for mongoose potential solutions (S) marked the beginning of the DMOA.

$$S = \begin{bmatrix} S_{1,1} & \dots & S_{1,Q} \\ \vdots & \ddots & \vdots \\ S_{N_p,1} & \dots & S_{N_p,Q} \end{bmatrix} \quad (16)$$

Q is the number of dwarf mongoose traits, N_p is the size of the population across all traits, and Q The number of decision variables is used for the parameter estimation problem. The population was created randomly using Eq. (17). The problem's lower and upper limits are denoted by L_b and U_b .

$$S_{u,v} = \text{unifrnd}(L_b, U_b, Q) \quad (17)$$

Three groups were created throughout the DMOA optimization process, as shown below.

Alpha Group, following initialization, Eq. (18) was used to determine the population fitness of each solution. The female alpha was selected based on fitness, as shown in Eq. (18).

$$\alpha = \frac{Fit_j}{\sum_{j=1}^{N_p} Fit_j} \quad (18)$$

Because the vocalization of the dominant female ρ and the number of babysitters bb are connected to the number of mongooses in α , the revised mechanism of the solution was computed using Eq. (19):

$$S_{j+1} = S_j + \emptyset * \rho \quad (19)$$

The distributed random number is denoted by $\emptyset$. For each iteration, the sleeping mound was determined using Eq. (20), and Eq. (21) was used to get the average of ε_j :

$$\varepsilon_j = \frac{Fit_{j+1} - Fit_j}{\max\{|Fit_{j+1}, Fit_j|\}} \quad (20)$$

$$\sigma = \frac{\sum_{j=1}^{N_p} \varepsilon_j}{N_p} \quad (21)$$

Scout Group, Elite chaotic searching was then used to increase the exploitation potential of the algorithm. In this method, elite people undergo chaotic mutation in order to achieve the additional renewal of elite individuals. After computing and sorting the population's fitness values in ascending order, the first n ($n = p \cdot M$) people are chosen as elite individuals; p is the selection proportion and falls between $[0, 1]$. The elite individuals are logged as $\{es_1(x), es_2(x), \dots, es_m(x)\} \in \{s_1(x), s_2(x), \dots, s_m(x)\}$. The j^{th} An elite individual is $es_j(x) = es_j^1(x), es_j^2(x), \dots, es_j^D(x)$ with the following upper and lower bounds as per Eq. (22).

$$\begin{cases} eb^d(x) = \max\{es_1^d(x), es_2^d(x), \dots, es_n^d(x)\} \\ ea^d(x) = \min\{es_1^d(x), es_2^d(x), \dots, es_n^d(x)\} \end{cases} \quad (22)$$

Following the mapping of the elite individual S_j to the interval $[0, 1]$, the chaotic individual $c_j(x) = c_j^1(x), c_j^2(x), \dots, c_j^D(x)$ is obtained. Its computation is as follows, as per Eq. (23).

$$c_j^d = \frac{es_j^d(x) - L_b^d(x)}{U_b^d(x) - L_b^d(x)}, \quad i = 1, 2, \dots, n \quad (23)$$

After that, a chaotic individual $c_i(x)$ logistic chaotic maps are executed, as per Eq. (24).

$$c_j^d(K + 1) = \mu \cdot c_j^d(K) [1 - c_j^d(K)] \quad (24)$$

The iteration number of the chaotic maps is represented by K , with the constant $\mu = 4$. Relocating the chaotic individuals into the interval $[ea^d(x), eb^d(x)]$ occurs when the maximum number of chaotic iterations, K_{max} , is achieved. Let set K_{max} to X in this study. The following is the process to acquire the j^{th} new elite individual $ec_j^d(x)$, as per Eq. (25).

$$ec_j^d(x) = c_j^{d, K_{max}}(x) \cdot [eb^d(x) - ea^d(x)] + ea^d(x) \quad (25)$$

Ultimately, a greedy decision is taken between $es_j(x)$ and $ec_j(x)$, that is expressed as per Eq. (26).

$$s_j(x + 1) = \begin{cases} es_j(x), & f(es_j(x)) \leq f(ec_j(x)) \\ ec_j(x), & f(es_j(x)) > f(ec_j(x)) \end{cases} \quad (26)$$

The boundaries of elite individuals converged towards the intended solution with increasing iterations, enhancing the local search capability of the MRFOs. During this phase, the family may discover a new sleeping mound if they forage sufficiently, as established by Eqs (27) and (28).

$$\text{If } \theta_{j+1} > \theta_j: s_{j+1} = s_j - DF * rand * [s_j - \vec{V}] \quad (27)$$

$$\text{Else: } s_{j+1} = s_j + DF * rand * [s_j - \vec{V}] \quad (28)$$

In this case, the random value falls between $[0, 1]$. The parameters DF and $\vec{V}$ were computed using Eq. (29) and (30) to regulate the mongoose group's collective violative movement and movement vector, respectively.

$$DF = \left(1 - \frac{m}{max_G}\right)^{\left(2 * \frac{m}{max_G}\right)} \quad (29)$$

$$\vec{V} = \sum_{j=1}^{N_p} \frac{S_j \times \varepsilon_j}{S_j} \quad (30)$$

The second category of people who remain with their children are the babysitters. As the rest of the team goes on daily hunting trips, babysitters are rotated to the alpha female. Algorithm 1 shows the HMRCO for feature selection.

Algorithm 1: HMRCO for Feature Selection	
Input:	N_p : Population size Q : Number of decision variables G_{max} : Maximum number of generations L_b, U_b : Lower and Upper bounds of the search space p : Selection proportion $\mu = 4$: Chaos constant K_{max} : Maximum number of iterations
Hyperparameters of feature selection and classifier parameter tuning	η : Learning rate (1e-3) B : Batch size (64) d : Dropout rate (0.4) C_{dim} : Capsule dimension: 8 α_t, γ : Focal loss ($\alpha_t = 0.25, \gamma = 2$)
Steps 1: Population Initialization	Initialize population $S \in \mathbb{R}^{N_p \times Q}$ using as per Eq. (16).
Step 2: Fitness Evaluation	Evaluate fitness Fit_j for each individual S_j Select dominant female (alpha) using Eq. (17)
Step 3: Alpha Group Update	Update alpha position as per Eq. (18) Compute fitness difference as per Eq. (19) Average fitness difference as per Eq. (20)
Step 4: Scout Group – Elite Chaotic Search	Select $n = p \cdot N_p$ elite individuals Determine bounds as Eq. (22) Normalize the elite individual as Eq. (23) Apply logistic chaos mapping for K_{max} as per Eq. (24) Re-scale chaotic output to search bounds as Eq. (25) Greedy selection as Eq. (26)
Step 5: Position Update with Movement Vector	Update solution position as per Eq. (27) and (28), whereas per (29) and (30)
Step 6: Babysitter Role	Retain some individuals as babysitters to preserve population diversity and avoid local minima.
Step 7: Repeat the steps	Repeat Steps 2–6 for $m = 1$ to G_{max}

Step 8: Return the best individual	Return the best individual S^* with optimal fitness value.
Output:	S^* Optimal feature subset

3.6. Cotton plant pest and disease Detection using TriAttnNet for classification

Using an 8-dimensional vector, each class may be represented using an eight-dimensional vector. The activation function Softmax is used in the output layer, as per Eqs (31) and (32)

$$y_i^{(l)} = f(z_i^l) \quad (31)$$

$$\text{where } z_i^l = \sum_{j=1}^{m_i^{(l-1)}} w_{i,j}^l y_j^{l-1} \quad (32)$$

Where $w_{i,j}^l$ are the weights that should be tuned by the complete fully connected layer to form the representation of each class, and f is the transfer function that represents the nonlinearity.

The performance of the model was enhanced using focal loss as the loss function. Focal loss is a type of categorical loss that addresses data imbalances. Because of the special concentrated loss that focuses on training on a limited set of difficult situations, the vast number of simple negatives is prevented from overwhelming the detector during training. The concept is quantified by Eq. (33). For each ground-truth class, p represents the expected probability, and the hyperparameters of the loss functions are α_t and γ .

$$FL(p_t) = -\alpha_t * (1 - p_t)^\gamma * \log(p_t) \quad (33)$$

Hard samples enhance the loss values more than easily classified samples do when there is concentrated loss. Thus, the model assigns a greater weight to these samples, thereby improving the accuracy of the hard samples. The TriAttnNet model ran at 3.2 GFLOPs on the forward pass and 12.5 ms on average on NVIDIA RTX 3060 at a batch size of 1 and 224x224 resolution. These findings confirm that the model is lightweight and suitable for practical agricultural applications. Figure 5 shows the disease classification diagram.

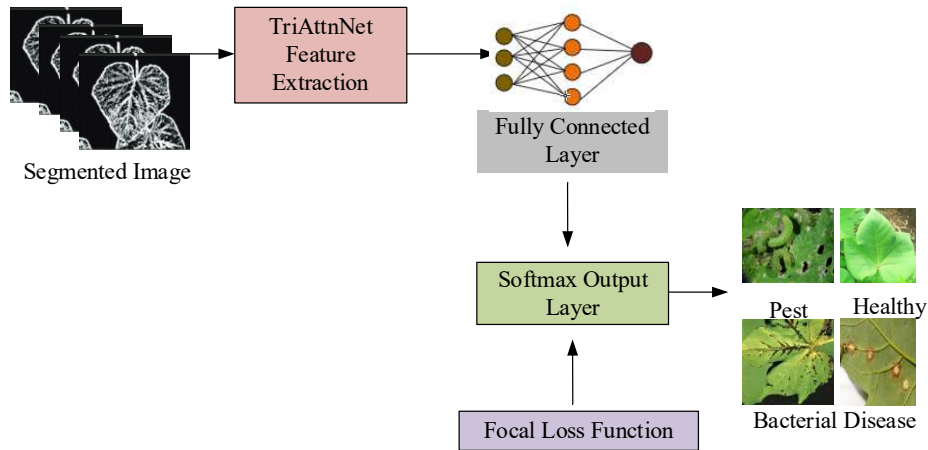


Fig. 5: Cotton plant pest and disease classification

4. SIMULATION RESULTS

4.1. Simulation Setup

The TriAttnNet model is inscribed in Python with TensorFlow and PyTorch frameworks and executed on an Intel(r) Core i7 processor with 4.2 GHz and 32 GB RAM with a 64-bit operating system. The experimental platform took advantage of the power of GPUs (NVIDIA RTX 3080, 12 GB) to train and infer faster. The cotton plant disease data, which is obtained in Kaggle, is partitioned into training (70%), validation (15%), and testing (15%) datasets. To evaluate performance, the proposed TriAttnNet is compared against four state-of-the-art models, namely RF-CottonNet [28], YOLOv9c [29], BERT-ResNet-PSO [30], and EfficientNetB1-CBAM [31]. Measurements like Accuracy, Precision, Recall, F-score, Sensitivity, Specificity, Negative Predictive Value (NPV), Matthews Correlation Coefficient (MCC), False Positive Rate (FPR), and False Negative Rate (FNR) are used to ensure that there is a thorough evaluation.

4.2. Algorithmic Analysis

The results of the comparative analysis are summarized in Tables 1 and 2, which clearly demonstrate the superior performance of the proposed TriAttnNet over existing approaches. Table 1 (70:20 split) and Table II (80:20 split) show the relative performance of various models in detecting pests and classifying diseases in cotton plants. In both data splits, the proposed TriAttnNet performs better than all the baseline models, with the highest accuracy (98.42% and 98.66%), precision, recall, F-score, and MCC, and the FPR and FNR. EfficientNetB1-CBAM and BERT-ResNet-PSO are competitive, and the accuracy is more than 95, whereas RF-CottonNet and YOLOv9c are comparatively less effective in all metrics. The 70:20 to 80:20 ratio is also similar in all models, but the TriAttnNet proposed has better generalization, with a balance between sensitivity and specificity. This shows that it is more effective in managing skewed disease patterns and subtle pest symptoms as compared to other state-of-the-art methods.

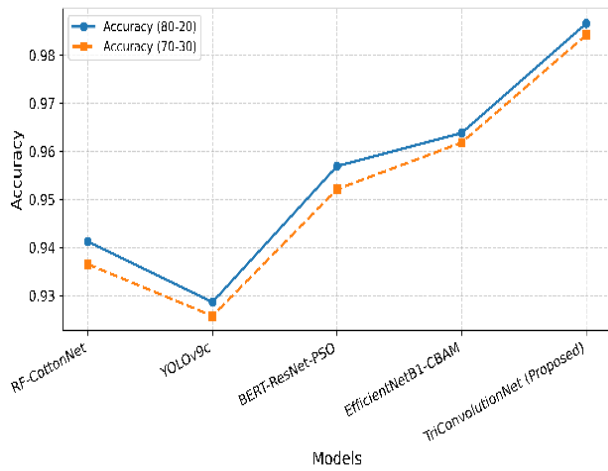
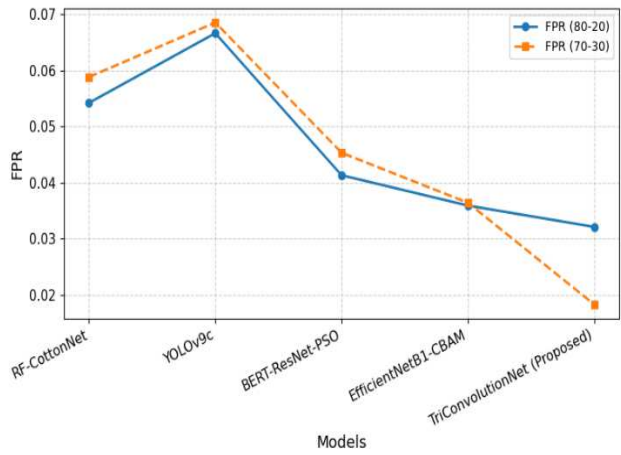
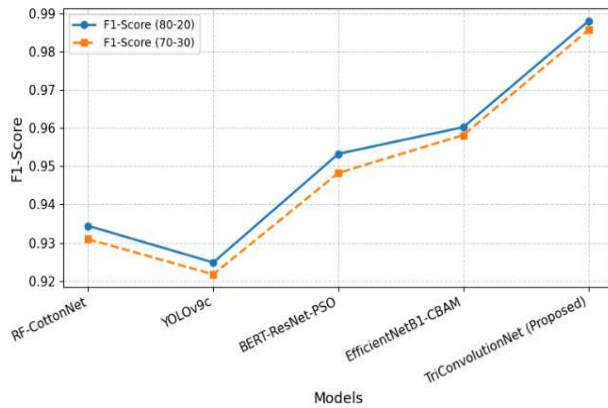
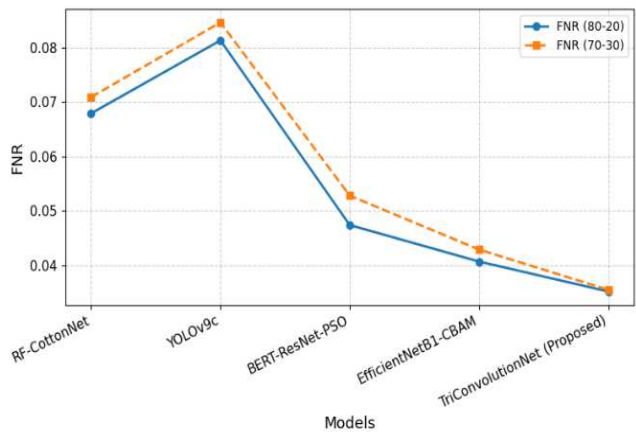
Table 1: Performance Analysis Of The Proposed Model over Other Models For 70:30% Learning Samples

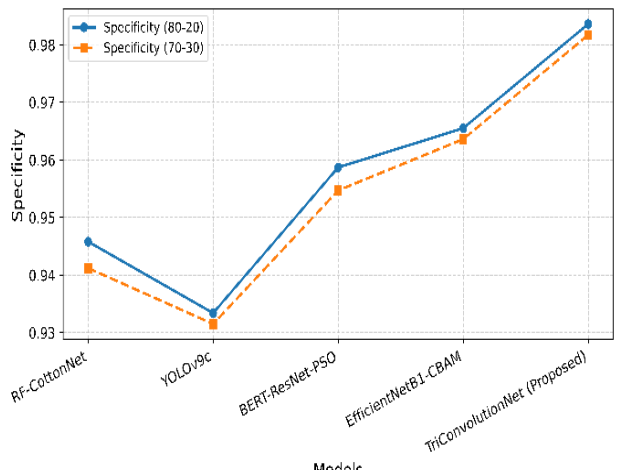
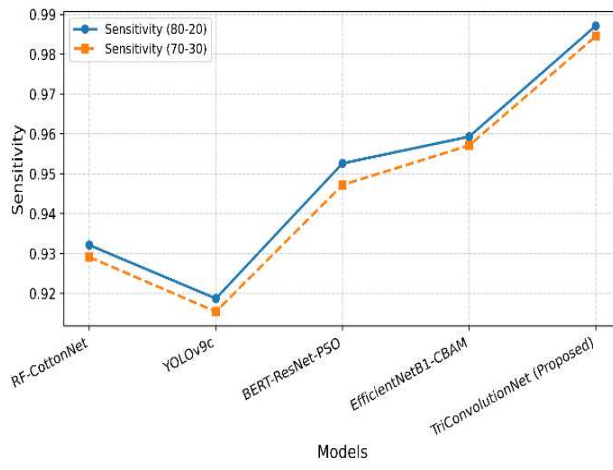
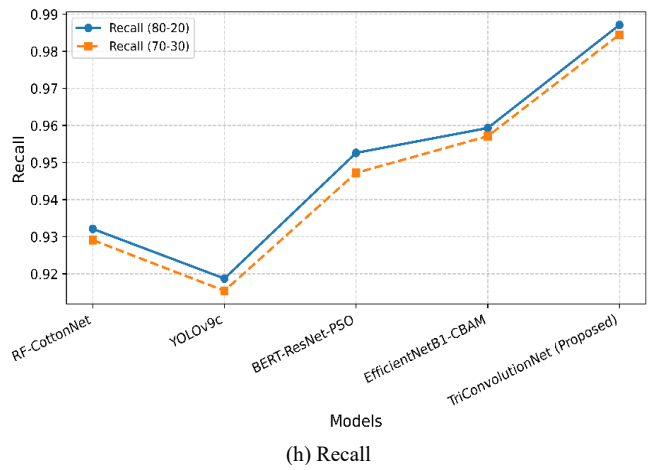
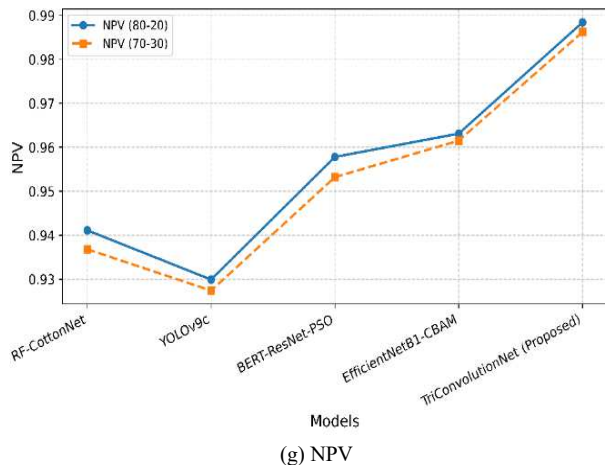
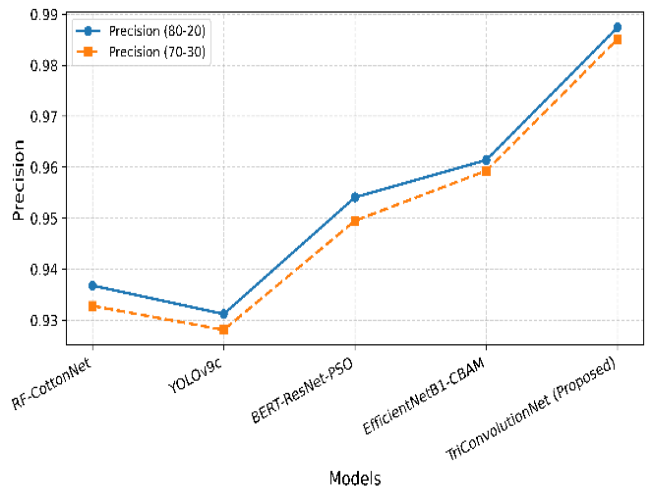
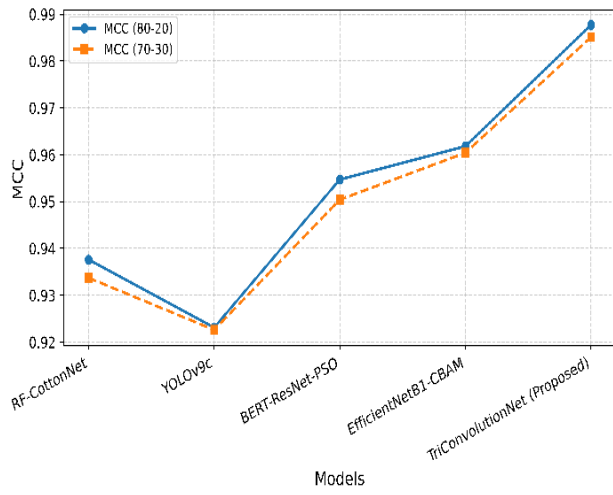
Model	Accu- racy	Preci- sion	Recall	F- Score	Sensi- tivity	Speci- ficity	NPV	MCC	FPR	FNR
RF-CottonNet	0.9365	0.9328	0.9291	0.9309	0.9291	0.9412	0.9368	0.9337	0.0588	0.0709
YOLOv9c	0.9257	0.9281	0.9154	0.9217	0.9154	0.9315	0.9274	0.9226	0.0685	0.0846
BERT-ResNet- PSO	0.9521	0.9495	0.9472	0.9483	0.9472	0.9547	0.9532	0.9504	0.0453	0.0528
EfficientNetB1- CBAM	0.9618	0.9593	0.9571	0.9582	0.9571	0.9636	0.9615	0.9604	0.0364	0.0429
Proposed	0.9842	0.9851	0.9845	0.9858	0.9845	0.9817	0.9862	0.9851	0.0183	0.0355

Table 2 : Performance Analysis Of The Proposed Model over Other Models For 80:20% Learning Samples

Model	Accuracy	Preci- sion	Recall	F- Score	Sensi- tivity	Speci- ficity	NPV	MCC	FPR	FNR
RF-CottonNet	0.9412	0.9368	0.9321	0.9344	0.9321	0.9458	0.9411	0.9376	0.0542	0.0679
YOLOv9c	0.9286	0.9312	0.9187	0.9248	0.9187	0.9334	0.9299	0.9231	0.0666	0.0813
BERT-ResNet- PSO	0.9569	0.9541	0.9526	0.9533	0.9526	0.9587	0.9578	0.9547	0.0413	0.0474
EfficientNetB1- CBAM	0.9638	0.9614	0.9593	0.9603	0.9593	0.9655	0.9631	0.9618	0.0345	0.0407
Proposed	0.9866	0.9874	0.9871	0.9881	0.9871	0.9836	0.9884	0.9877	0.0341	0.0452

Figure 6 shows a comparative performance evaluation of the proposed TriAttnNet with the benchmark models RF-CottonNet, YOLOv9c, BERT-ResNet-PSO, and EfficientNetB1-CBAM on several evaluation measures. Accuracy, FNR, FPR, F1-score, MCC, NPV, Precision, Recall, Sensitivity, and Specificity are represented by subfigures (a) to (j), respectively. The findings show that the proposed model is always obtained with better values in Accuracy, F1-score, MCC, NPV, Precision, Recall, Sensitivity, and Specificity, and has the lowest FPR and FNR than other models.

**(a) Accuracy****(b) FPR****(c) F1-score****(d) FNR**



(i) Sensitivity

(j) Specificity

Fig. 6: Performance of Proposed Model over Existing Models for (a) Accuracy, (b) FPR, (c) F1-score, (d)FNR, (e) MCC, (f) Precision, (g) NPV, (h) Recall, (i) Sensitivity, (j) Specificity

Table 3: Performance comparison of The Proposed Model against Transformer-based approaches for plant disease detection

Model	Accuracy	Precision	Recall	F1-Score
Ensemble Heterogeneous Transformer [32]	0.975	0.976	0.974	0.975
Vision Transformer [33]	0.978	0.979	0.977	0.978
Efficient Swin Transformer [34]	0.981	0.982	0.981	0.981
Proposed	0.986	0.987	0.987	0.988

Table 3 presents the comparison between the proposed model and Transformer-based methods of detecting plant diseases. Although the Ensemble Heterogeneous Transformer[32], Vision Transformer[33], and Efficient Swin Transformer[34] all scored highly with varying accuracies of 0.975 and 0.981, respectively, and better precision, recall, and F1-score, TriAttnNet always outperforms the other models significantly at 0.9866 with a higher precision, recall, and F1-score.

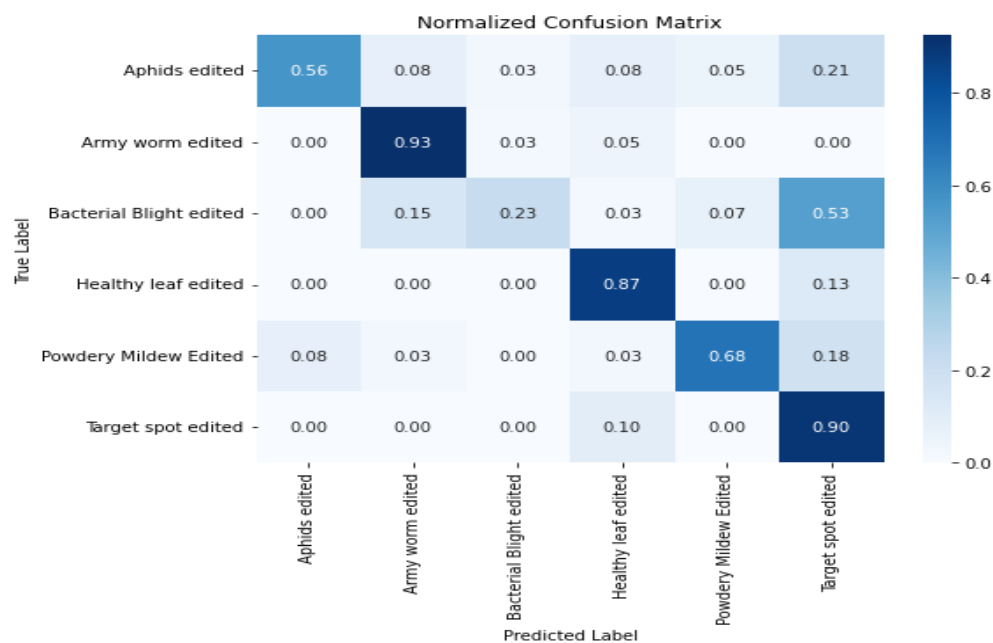


Fig. 7: Confusion Matrix

Figure 7 explains that the confusion matrix reveals that the proposed model performs well for most classes, with Class 1 (93%), Class 3 (87%), and Class 5 (90%) achieving high correct prediction rates. Class 4 also shows moderate accuracy (68%), though with some misclassifications. However, Class 0 (56%) and especially Class 2 (23%) exhibit weaker performance, with significant confusion toward Class 5. This indicates that while the model effectively distinguishes major categories, certain classes with overlapping visual features require further refinement, possibly through data augmentation and enhanced attention mechanisms.

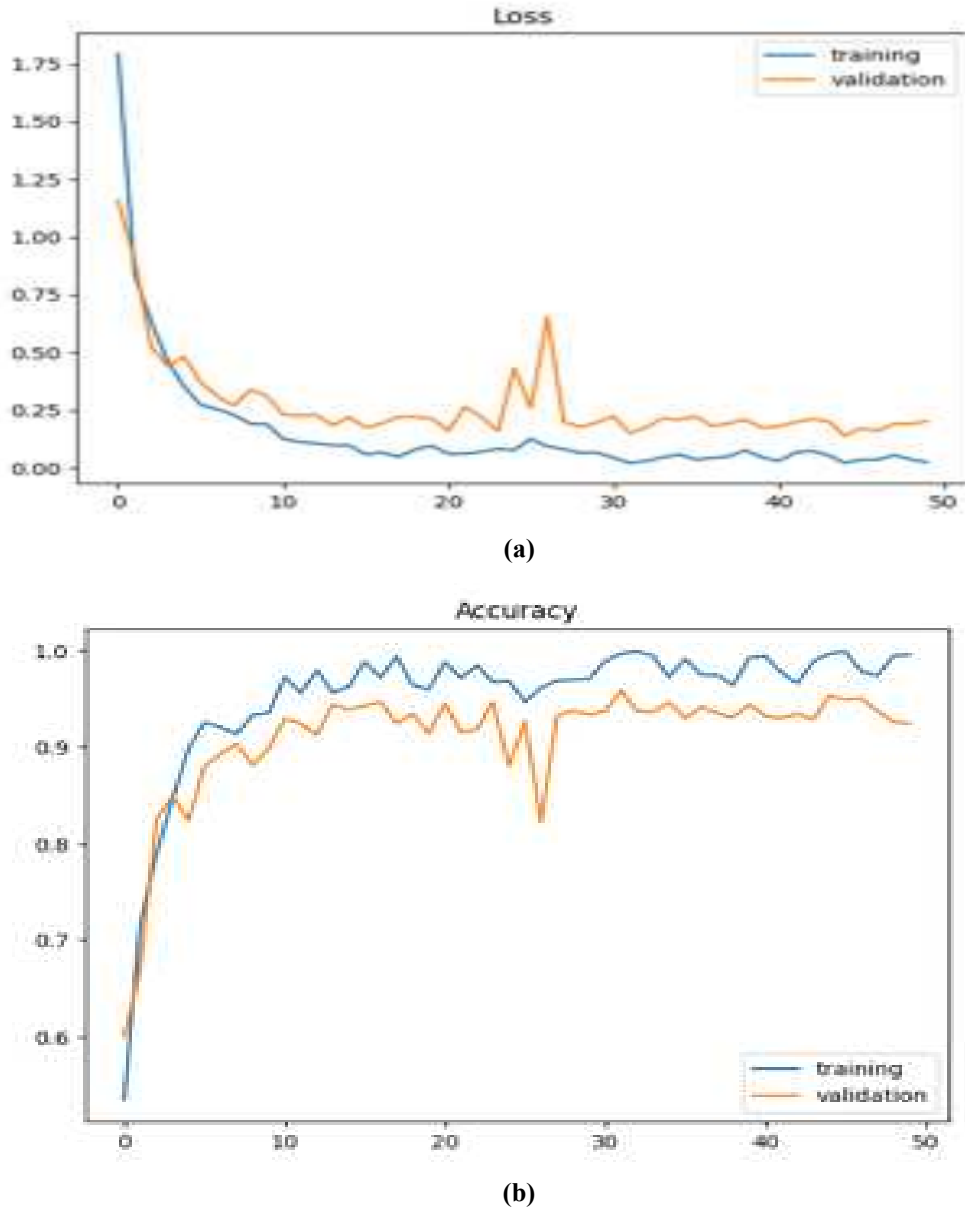


Fig. 8: a) Training and Validation Loss over Epochs, b) Training and Validation Accuracy

Figure 8 illustrates the training dynamics of the proposed TriAttnNet model with 50 epochs. It shows the training loss reduces gradually with a high value at the beginning, and approaches almost zero at approximately 40 epochs, which means that the model successfully learns and reaches convergence. The loss on validation has a similar downward trend, albeit with minor variations, as is anticipated because of the difference in unknown data. The fact that the gap between training and validation losses is relatively low indicates that the model does not overfit significantly and generalizes. Figure 8b displays that the training accuracy increases quickly in the first few epochs and reaches above 90 percent in the first 10 epochs and almost 100 percent in the last few epochs. The validation accuracy also increases rapidly and even levels off at approximately 90 percent by the 15th epoch. The fact that the training and validation accuracy curves are very close indicates that the model has a high generalization ability. Minor variations in the

accuracy of validation at some points are natural variations that occur during training but do not affect the overall trend.

Table 4: Ablation Study Of The Metrics

Ablation	Accuracy
Without Augmentation	0.8912
Without Segmentation	0.9024
Without Optimization	0.9176
Developed TriAttnNet	0.9866

Table 4 shows that the various components play a role in the performance of the developed TriAttnNet. Data augmentation is not used; the model attains an accuracy of 89.12, which shows that augmentation is very important in enhancing the capacity of the network to generalize, through exposure to different variations of leaf images. In the absence of segmentation, the accuracy is 90.24, indicating that segmentation helps the model focus on relevant leaf regions and reduces background noise, thereby improving feature extraction. Without optimization methods, the accuracy reaches 91.76, and it demonstrates that optimization additionally stabilizes the training and increases convergence. The TriAttnNet is fully developed, and the highest accuracy of 98.66 is confirmation that the effect of augmentation, optimization, and the proposed architecture, combined with each other, positively influences the classification. This development shows that every design decision adds value to the robustness of models, and their combination results in the state-of-the-art accuracy of cotton leaf disease detection.

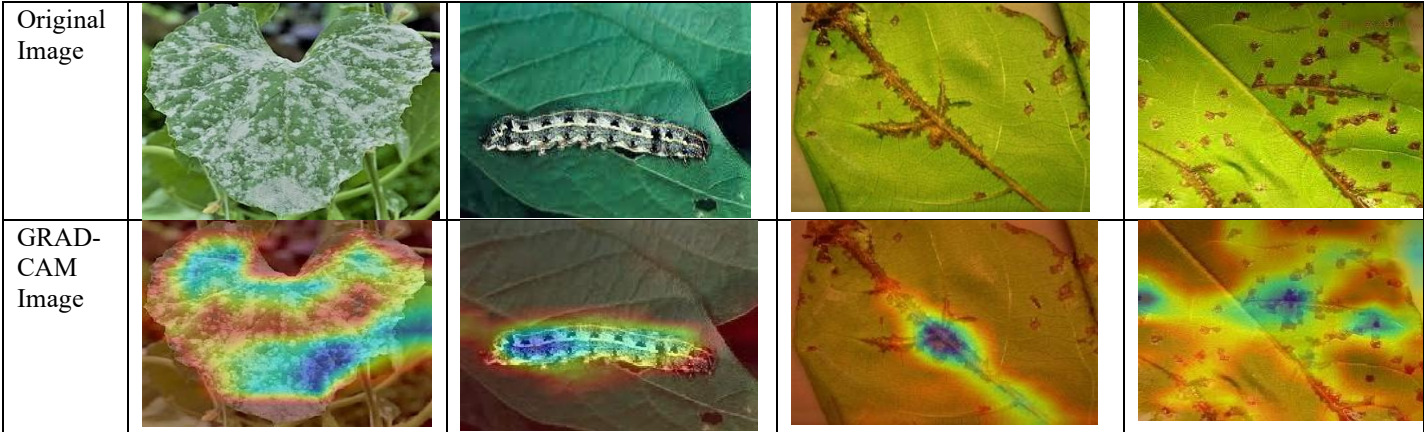


Fig. 9: Original cotton leaf images with corresponding Grad-CAM visualizations highlighting disease-affected regions.

Figure 9 shows the cotton leaf diseases (Powdery mildew, Armyworm, Bacterial Blight, Aphids) images with their corresponding Grad-CAM visualizations, highlighting the regions most influential in Tri-AttnNet’s predictions. The heatmaps are a clear indication that the model targets the diseased areas of the leaves, meaning that the network pays attention to the features that matter, and that the network interprets to substantiate the argument on explainability.

5. CONCLUSION

The suggested framework presents a powerful and integrated pipeline of cotton pest detection and disease classification. The model uses Gaussian blur filtering and CLAHE as the preprocessing steps, and SpaGAN-based data augmentation to guarantee the high quality and variety of training samples. To perform segmentation, an AGMS-UNet is utilized to learn fine-scale disease localities. The TriAttnNet, a combination of spatial, channel, and contextual attention, enables feature extraction with rich and discriminative representations. In order to further optimize the model, the HMRCO algorithm is employed to optimize features and parameters. Lastly, the TriAttnNet classification framework, which is an integration of DenseNet, Improved MobileNet, and Classification with focal loss, is the best in terms of accuracy and robustness. This framework shows some continuous improvements in accuracy, precision, recall, F1-score, and MCC, and error-rate metrics compared to recent benchmark models such as RF-CottonNet, YOLOv9c, BERT-ResNet-PSO, and EfficientNetB1-CBAM. TriAttnNet, which is used to classify features robustly. To combine explainable AI (XAI) methods with the future study to achieve a better level of interpretability, sample the data using a multi-geographical and multi-seasonal sample to achieve a better generalization, and implement lightweight deployment on edge devices to provide real-time agricultural decision support.

DECLARATIONS:

FUNDING

NO

CONFLICTS OF INTEREST

NO

COMPETING INTERESTS

NO

DECLARATION OF INTERESTS

Author states that they have no conflicting, financial or personal interests or relationships that might affect the work presented in this paper.

DATA AVAILABILITY STATEMENT

All data used for this study are available within the manuscript. Model/code will be made available upon request.

References

- [1]. S. Wang, Y. Li, J. Yuan, L. Song, X. Liu, and X. Liu, "Recognition of cotton growth period for precise spraying based on convolution neural network," *Information Processing in Agriculture*, vol. 8, no. 2, pp. 219–231, 2021.
- [2]. Y. Lu, K. A. Wyckhuys, L. Yang, B. Liu, J. Zeng, Y. Jiang, and K. Wu, "Bt cotton area contraction drives regional pest resurgence, crop loss, and pesticide use," *Plant Biotechnology Journal*, vol. 20, no. 2, pp. 390–398, 2022.
- [3]. Z. Mellick, A. Payne, and L. Buys, "From fibre to fashion: Understanding the value of sustainability in global cotton textile and apparel value chains," *Sustainability*, vol. 13, no. 22, p. 12681, 2021.
- [4]. G. Silesh, A. Zemedkun, and S. Nurhussien. "Assessment of the pesticide use, practice and hazards on cotton growing farmers in the Awash Valley, Ethiopia," *AgriRxiv*, 2021.
- [5]. Y. Kurmi and S. Gangwar, "A leaf image localization-based algorithm for different crops disease classification," *Information Processing in Agriculture*, vol. 9, no. 3, pp. 456–474, 2022.
- [6]. G. Storey, Q. Meng, and B. Li, "Leaf disease segmentation and detection in apple orchards for precise smart spraying in sustainable agriculture," *Sustainability*, vol. 14, no. 3, p. 1458, 2022.
- [7]. O. Bashir, S. A. Bangroo, N. B. Naikoo, N. A. Ganie, S. S. Kukal, D. Amin, and K. R. Hakeem, "Deep learning approach for detecting, classifying, and managing insect pests of various crops," in *Soil, Agriculture, and Ecosystem Modeling*, pp. 171–196, Apple Academic Press, 2024.
- [8]. S. Yogindran, M. Kumar, L. Sahoo, K. Sanatombi, and S. Chakraborty, "Occurrence of cotton leaf curl Multan virus and associated betasatellites with leaf curl disease of Bhut-Jolokia chillies (*Cap-sicum chinense* Jacq.) in India," *Molecular Biology Reports*, vol. 48, no. 3, pp. 2143–2152, 2021.
- [9]. B. Binyameen, Z. Khan, S. H. Khan, A. Ahmad, N. Munawar, M. S. Mubarik, and S. H. Qari, "Using multiplexed CRISPR/Cas9 for suppression of cotton leaf curl virus," *International Journal of Molecular Sciences*, vol. 22, no. 22, p. 12543, 2021.
- [10]. K. Bananej, M. Shafiq, and M. S. Shahid, "Association of cotton leaf curl Gezira virus with tomato leaf curl betasatellite infecting Carica papaya in Iran," *Australasian Plant Disease Notes*, vol. 16, no. 1, p. 4, 2021.
- [11]. Y. Chi, L. L. Pan, S. S. Liu, S. Mansoor, and X. W. Wang, "Implication of the whitefly protein Vps twenty associated 1 (Vta1) in the transmission of cotton leaf curl Multan virus," *Microorganisms*, vol. 9, no. 2, p. 304, 2021.
- [12]. P. Johri, S. Kim, K. Dixit, P. Sharma, B. Kakkar, Y. Kumar, and M. F. Ijaz, "Advanced deep transfer learning techniques for efficient detection of cotton plant diseases," *Frontiers in Plant Science*, vol. 15, p. 1441117, 2024.
- [13]. Z. A. Haider, F. M. Khan, I. U. Khan, M. A. Khan, and R. Khan, "Early detection and prediction of pests in field crops using transfer learning," *VFAST Transactions on Software Engineering*, vol. 12, no. 3, pp. 98–113, 2024.
- [14]. J. Karthika, M. Santhos, and T. Sharan, "Disease detection in cotton leaf spot using image processing," in *Journal of Physics: Conference Series*, vol. 1916, no. 1, p. 012224, IOP Publishing, 2021.
- [15]. M. R. Latif, M. A. Khan, M. Y. Javed, H. Masood, U. Tariq, and S. Kadry, "Cotton leaf diseases recognition using deep learning and genetic algorithm," *Computers, Materials & Continua*, vol. 69, no. 3, 2021.
- [16]. M. Zekiwo and A. Bruck, "Deep learning-based image processing for cotton leaf disease and pest diagnosis," *Journal of Electrical & Computer Engineering*, 2021.
- [17]. R. F. Caldeira, W. E. Santiago, and B. Teruel, "Identification of cotton leaf lesions using deep learning techniques," *Sensors*, vol. 21, no. 9, p. 3169, 2021.
- [18]. N. Zhang, X. Zhang, P. Shang, R. Ma, X. Yuan, L. Li, and T. Bai, "Detection of cotton verticillium wilt disease severity based on hyperspectrum and GWO-SVM," *Remote Sensing*, vol. 15, no. 13, p. 3373, 2023.

- [19]. A. I. Jajja, A. Abbas, H. A. Khattak, G. Niedbała, A. Khalid, H. T. Rauf, and S. Kujawa, "Compact convolutional transformer (CCT)-based approach for whitefly attack detection in cotton crops," *Agriculture*, vol. 12, no. 10, p. 1529, 2022.
- [20]. X. Liang, "Few-shot cotton leaf spots disease classification based on metric learning," *Plant Methods*, vol. 17, no. 1, p. 114, 2021.
- [21]. M. S. Memon, P. Kumar, and R. Iqbal, "Meta deep learn leaf disease identification model for cotton crop," *Computers*, vol. 11, no. 7, p. 102, 2022.
- [22]. Z. Wang, H. W. Zhang, Y. Q. Dai, K. Cui, H. Wang, P. W. Chee, and R. F. Wang, "Resource-efficient cotton network: A lightweight deep learning framework for cotton disease and pest classification," *Plants*, vol. 14, no. 13, p. 2082, 2025.
- [23]. P. Bishshash, A. S. Nirob, H. Shikder, A. H. Sarower, T. Bhuiyan, and S. R. H. Noori, "A comprehensive cotton leaf disease dataset for enhanced detection and classification," *Data in Brief*, vol. 57, p. 110913, 2024.
- [24]. S. Kumar, R. Ratan, and J. V. Desai, "Cotton disease detection using TensorFlow machine learning technique," *Advances in Multimedia*, vol. 2022, no. 1, p. 1812025, 2022.
- [25]. Dataset is taken from <https://www.kaggle.com/datasets/dhamur/cotton-plant-disease> dated on 17/01/2024.
- [26]. A. Kaushal, R. Kumar, and S. Kumar, "SobelMNet: A lightweight CNN-based approach for median filtering detection in compressed images," in *Proceedings of the 2025 27th International Conference on Advanced Communications Technology (ICACT)*, pp. 244–248, 2025.
- [27]. F. F. Khan, J. H. Kim, C. S. Park, J. I. Kim, and G. R. Kwon, "A resource-efficient 3D U-Net for hippocampus segmentation using CLAHE and SCE-3DWT techniques," *IEEE Access*, 2025.
- [28]. S. Zhang, Y. Zhang, Z. Yu, S. Yang, H. Kang, and J. Xu, "MCA-GAN: A multi-scale contextual attention GAN for satellite remote-sensing image dehazing," *Electronics*, vol. 14, no. 15, p. 3099, 2025.
- [29]. R. He, F. Zhang, J. Zhu, Y. Wang, D. Yang, T. Zhang, and P. Li, "YOLOv9-LSBN: An improved YOLOv9 model for cotton pest and disease," *IEEE Access*, 2025.
- [30]. C. Singh, S. Wibowo, and S. Grandhi, "A hybrid deep learning approach for cotton plant disease detection using BERT-ResNet-PSO," *Applied Sciences*, vol. 15, no. 13, p. 7075, 2025.
- [31]. M. Rahman, M. Ullah, R. Devnath, T. Chowdhury, G. Rahman, and M. Rahman, "Cotton leaf disease detection: an integration of CBAM with deep learning approaches," *International Journal of Computer Applications*, vol. 975, no. 8887, 2025.
- [32]. P. Kalpana, P. Gera, E. Alabdulkreem, M. T. Quasim, J. Baili, Y. Cho, and Y. Nam, "An ensemble heterogeneous transformer model for an effective diagnosis of multiple plant diseases," *Frontiers in Plant Science*, vol. 16, p. 1693095, Dec. 2025.
- [33]. M. Clement, "Vision Transformer-Based Systems for Crop Disease Detection and Monitoring in Precision Agriculture," 2025.
- [34]. W. Liu and A. Zhang, "Plant disease detection algorithm based on Efficient Swin Transformer," *Computers, Materials & Continua*, vol. 82, no. 2, 2025.

