

Residual Vision Driven Ensemble Framework for Robust Mulberry Leaf Quality Classification

Chaitra D B^{1*}, Dr. Mohan K G², Dr. Dayanand Lal N³, Dr. I Jeena Jacob⁴

¹Research Scholar, Department of computer science and engineering, GITAM School of Technology, GITAM (Deemed to be University), Bengaluru Campus

²Professor, Department of computer science and engineering, GITAM School of Technology
GITAM (Deemed to be University), Bengaluru Campus

³Assistant Professor, Department of computer science and engineering, GITAM School of Technology, GITAM (Deemed to be University), Bengaluru Campus

⁴Professor, Department of computer science and engineering, GITAM School of Technology
GITAM (Deemed to be University), Bengaluru Campus

*Corresponding Author Email: cbasavar@gitam.in

Orcid ID: 0000 0002 7438 6642

ABSTRACT:

Timely and accurate classification of mulberry leaf health is crucial in sericulture for optimizing crop yield and preventing disease proliferation. Despite the progress in deep convolutional networks, current solutions face challenges in generalization, computational cost, and robustness across diverse leaf classes. To alleviate these challenges, we propose a novel hybrid ensemble classification framework integrating Residual Vision Network-based feature extraction with Light GBM, Random Forest, and XGBoost classifiers, optimized through extensive hyperparameter tuning and stratified 5-fold cross-validation. The study leverages the Kaggle Mulberry Leaf dataset, encompassing three distinct classes, with class-balanced data split into training and test subsets at a ratio of 80:20. The results reveal that the proposed ensemble strategy not only outperforms standalone deep networks such as DenseNet121, VGG19, and MobileNetV3 but also exhibits superior discriminative performance, achieving an F1-score of 98.17%, AUC of 99.38%, and specificity of 98.36%. Ablation analysis further substantiates the contribution of the Residual Vision module, yielding performance gains of up to 3.12% in F1-score over non-residual variants. This work effectively bridges the limitations of current deep models by combining residual-enhanced feature extraction with lightweight gradient-boosted decision ensembles, offering a scalable and accurate approach for agricultural vision systems. The proposed framework holds substantial promise for deployment in real-time leaf monitoring applications with low computational overhead and high interpretability.

Keywords: Mulberry Leaf Health, Sericulture, Residual Vision Network, LightGBM, Random Forest, and XGBoost classifiers.

1. Introduction

Any nation, no matter how big or small, has a portion of its population that works in the agriculture sector, either to meet basic food needs or as a main source of export revenue [1]. Nowadays, agriculture provides a lot more than simply food; it also significantly boosts the economies of so many countries. The usage of fertilizer, pesticides, and herbicides to boost yield has a significant impact on soil, water, and air [2]. Sericulture, often referred to as silk farming, is quite prevalent in several nations in the Middle East and south of Europe, as well as in the nations of Southeast Asia. As the climate is tropical, this is the main reason these plants are grown in these regions [3].

Silk is indirectly produced from mulberry plants, which are prone to failure due to a variety of diseases. Silk which is included in some of the finest natural fabrics in the world is only 9–10 micrometres thick [4]. Silkworms mostly consume mulberry leaves as food. They finally transform themselves into cocoons. The cocoons are harvested. The string-like cocoon fiber can have a maximum length of one kilometer. One kg of silk is created from nearly 5500 cocoons. 1 kg of silk costs around \$50 to \$55 in the international market, and it is around four rupees in the Indian market [5]. Hence the ecosystem of sericulture depends heavily on the leaves. At a normal rate, if the quality of the silk produced is not up to par, the farmer has to bear up to 40% of its total expenditure [6]. The quality of the cocoon is closely related to the quantity and quality of the leaf. The elements of cocoon leaves and silkworms all rely on one another [7]. The majority of farmers neglect the soil structure and other lifeless elements, which influence the effectiveness of the plants' responses. The plants become starved, which cause the outcomes to be unsatisfactory. An appropriate and calculated amount of organic nutrients can provide the best final product possible [8].

The mulberry tree (*Morus alba*) is a deciduous tree local to northern China that has been cultivated for centuries across Asia, Europe, and North America [9]. The leaves of mulberry trees contain essential nutrients and bioactive compounds to support the growth and development of silkworm larvae. The S-36

and BR-2 varieties of mulberry leaves were found to have the highest nutritional composition among the other types [10]. Consistently high quantities of protein, chlorophyll, and carotenoids make it suitable for silkworm growth. Silkworms fed on a diet of mulberry leaves produce higher-quality silk compared to other plant sources [11]. According to comparative proteomics investigations, mulberry leaves protect silkworms from some diseases in addition to providing them with vital amino acids. Despite being heavily consumed by silkworms, mulberry leaves continue to flourish, offering fresh perspectives on the relationships between hosts, plants, and insects [12]. Mulberry leaves are rich in carbohydrates, proteins, vitamins, and minerals, which are converted into silk fibers during the metamorphosis stage of silkworms. The quality and quantity of silk are highly dependent on the abundant supply of healthy mulberry leaves [13]. Fungal diseases such as leaf spot, powdery mildew, and rust can affect the nutritional value of mulberry leaves. Studies have shown a reduction in protein, amino acids, chlorophyll, and carotenoids in diseased leaves [14]. Bacterial and viral diseases decrease the protein content along with sugar and ascorbic acid in mulberry leaves [15]. An effective solution must also incorporate uncertainty estimation to bolster classifier robustness, ensuring reliable performance despite the challenges posed by imbalanced datasets and subtle disease symptoms. This pressing need motivates the exploration of ensemble approaches that can combine the strengths of multiple architectures and adaptive strategies to meet the demands of modern plant disease diagnostics. The key contributions of this study are as follows:

- ❖ A novel hybrid classification framework is proposed that integrates a Residual Vision Network—comprising multi-scale ResNet and Vision Transformer (ViT)—with an ensemble of LightGBM, Random Forest, and XGBoost classifiers to effectively capture spectral and spatial leaf features.
- ❖ Bayesian uncertainty modelling is employed to enhance the model's robustness by quantifying predictive confidence and mitigating overfitting. Also, the introduction of a Dynamic Class Weight Adjustment (DCWA) mechanism addresses class imbalance and intra-class variability by adaptively penalizing misclassified minority samples during training.
- ❖ A specialized fractal feature encoding strategy is adopted within the residual backbone to accurately capture self-similar texture and boundary characteristics inherent in diseased and stressed mulberry leaves.
- ❖ The proposed model is extensively validated using the Kaggle Mulberry Leaf dataset, and it outperforms classical CNN baselines (DenseNet121, VGG19, MobileNetV3) to achieve improved accuracy. The rest of the paper is organized as follows: Section 2 reviews the related literature in mulberry leaf classification, deep learning in agriculture, and ensemble learning methods. Section 3 describes the proposed framework in detail, including the ADSE-based preprocessing, Residual Vision Network architecture, Bayesian modelling, DCWA strategy, and ensemble classifier integration. Section 4 presents the experimental results, visualizations, ablation study, and comparative performance analysis. Section 5 concludes the paper with a summary of key findings, practical implications for sericulture, and potential directions for future work.

2. Literature Survey

The study by Aldakheel et al., [16], employed the YOLOv4 algorithm to detect, identify, and localize plant leaf diseases using the Plant Village dataset. It achieved a remarkable accuracy of 99.99%, though its performance is reliant on dataset quality and high computational demands. Future directions include real-time applications, edge computing integration, and collaborative efforts with agricultural experts. Another study by Venkatesh et al., [17], utilizes an enhanced VGG16 transfer learning model to identify nitrogen (N), phosphorus (P), and potassium (K) deficiencies in groundnut leaves, achieving a classification accuracy of 98%. The methodology, while highly accurate, requires validation across diverse environments and crops. Fuentes-Pacheco et al., [18], employed a Curriculum by Smoothing (CS) technique with ResNet50V2, achieving a 7% higher accuracy compared to traditional methods. However, its complexity and hardware requirements pose challenges. Bijoy et al., [19], introduced a lightweight dCNN method, excelling in accuracy with fewer parameters, yet it demands better hardware and involves hyperparameter complexity. The study by Madhurya et al., [20], proposed an optimized YOLOv7-based framework combined with Rat Swarm and Red Fox Optimization (YR2S), achieving 99.69% accuracy. However, it struggles with dense spot detection and real-time efficiency. In contrast, Majdalawieh et al., [21], used SSD MobileNet v2 and EfficientDet D0 models for nutrient deficiency detection, demonstrating high accuracy (93-98%) and suitability for real-time applications, though limited by its box-based detection approach. The study by Soeb et al., [22], enhanced the YOLOv7 architecture with attention mechanisms, significantly improving detection accuracy in natural settings. However, challenges with limited labelled data and a lack of evaluation metrics persist. Similarly, Waheed et al., [23], employed CNNs and VGG-16 models to detect diseases and deficiencies in ginger plants, achieving up to 99% accuracy but is constrained by dataset specificity. Lastly, a work by Ahmed et al., [24], presented a mobile app utilizing CNNs for plant disease detection across 14 crop species, achieving a detection accuracy of 94%. Despite its real-time capabilities,

the system faces challenges with subtle bacterial and viral symptoms and limited data. Future advancements for many of these methodologies include expanding datasets, improving scalability, and exploring real-time or UAV-based applications.

3. Proposed Methodology

In sericulture, the quality of mulberry leaves, which are the sole feed for silkworms, is paramount for successful silk production. Healthy mulberry leaves directly influence the growth, vitality, and ultimately the quality of the silk produced. Our proposed framework is a comprehensive approach specifically designed to elevate mulberry leaf quality analysis for sericulture by integrating advanced pre-processing, feature extraction, and classification techniques. The framework is structured into three primary stages: pre-processing, feature extraction, and classification, each addressing key technical gaps and ensuring robust performance for assessing leaf health. Traditional image enhancement techniques for assessing leaf quality often over-enhanced images, failing to highlight subtle discolorations and texture variations. To address this, the Adaptive Dual-Scale Enhancement (ADSE) technique is used, which combines wavelet-based contrast stretching with anisotropic diffusion filtering to enhance the quality of leaf image. This method is particularly beneficial for sericulture, as it ensures early signs of leaf deterioration and are not masked by noise or lost through over-enhancement. The framework also employs a Residual Vision Network, which combines a multi-scale ResNet with a Vision Transformer, enhanced by Bayesian uncertainty modelling. This holistic feature extraction process accurately captures nuanced indicators of mulberry leaf quality, crucial for silkworm rearing interventions. A novel Dynamic Class Weight Adjustment (DCWA) method is proposed to address class imbalance and intra-class variability.

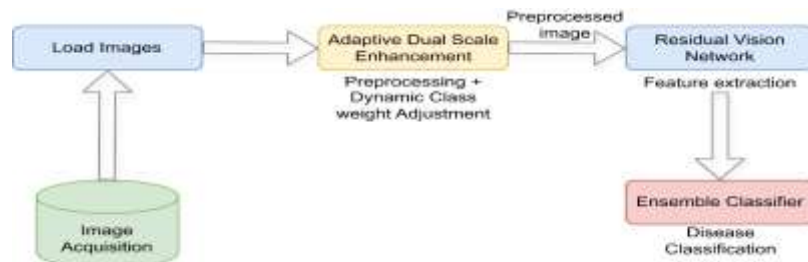


Figure 1: Architecture diagram of the proposed method

Thus, the ensemble approach leverages the complementary strengths of each classifier, ensuring a robust and generalized model capable of accurately distinguishing between healthy and compromised mulberry leaves. By penalizing misclassifications and dynamically modulating the learning focus, the framework improves the early detection of quality issues in mulberry leaves and directly supports the optimal rearing of silkworms, ultimately contributing to higher-quality silk production.

3.1. Adaptive Dual-Scale Enhancement (ADSE) technique

The core challenge lies in simultaneously reducing image noise while preserving low-contrast indicators that provide key insights into leaf health. Also, traditional image enhancement techniques such as standard histogram equalization and Contrast Limited Adaptive Histogram Equalization, abbreviated as CLAHE have been employed to assess leaf quality, yet these methods frequently over-enhance the images, which fail to highlight the subtle discolorations and texture variations that signal the nutrient deficiency symptoms. These subtle markers are critical since even minor degradations in leaf quality can adversely affect silkworm development and silk yield.

Table 1: Adaptive Dual-Scale Enhancement (ADSE) Process Description

Stage	Description
Input Image Acquisition	Acquire raw RGB image of the mulberry leaf under controlled lighting conditions.
Wavelet Decomposition	Decompose the image into multiple frequency sub-bands using a Discrete Wavelet Transform (DWT). Each band represents different levels of detail and approximation.
Contrast Stretching	Apply adaptive contrast stretching on selected sub-bands, enhancing low-contrast discoloration and texture variations without introducing visual artifacts.
Wavelet Reconstruction	Reconstruct the image from the contrast-enhanced sub-bands using inverse DWT, retaining fine-scale and coarse-scale improvements.

Anisotropic Diffusion Filter	Apply anisotropic diffusion to suppress background noise while preserving edge details. The number of iterations and conductance coefficients are adaptively tuned.
Output Image	Return the enhanced image with reduced noise and amplified pathological indicators like nutrient deficiencies and disease-related patterns.

The table 1 demonstrates the sequential flow and dual emphasis of the ADSE method on both frequency-domain enhancement and spatial-domain noise suppression, which are critical for detecting early degradation in mulberry leaves. Thus, the first stage utilizes the Adaptive Dual-Scale Enhancement (ADSE) technique, which synergistically combines wavelet-based contrast stretching with anisotropic diffusion filtering. In this approach, wavelet transforms decompose the mulberry leaf image into multiple frequency bands, enabling targeted contrast enhancement that reveals subtle discolorations, texture variations, and essential markers of nutrient deficiency. The contrast stretching component amplifies these low-contrast features without overwhelming the image, while anisotropic diffusion filtering minimizes noise by smoothing homogeneous areas while preserving critical edge details. The dual benefits of ADSE are particularly advantageous for sericulture; they ensure that early signs of leaf deterioration are neither masked by noise nor lost through over-enhancement, thus providing a clear, detailed visualization of leaf quality information that is essential for the optimal rearing of silkworms.

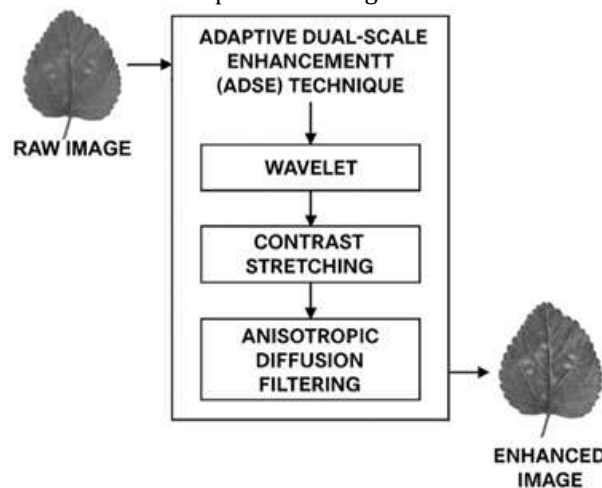


Figure 2: Proposed pre-processing method

Mathematically, ADSE integrates two complementary yet inseparable operations, which are multi-resolution frequency-selective amplification and spatially adaptive edge-preserving smoothing, into a single, iterative framework. The enhancement begins by transforming the input image $I(x, y)$ into a joint spatial-frequency domain via the Discrete Wavelet Transform (DWT) [25]. Let $I_0 = I(x, y) \in R^{M \times N}$ be the grayscale representation of the input mulberry leaf image. A single-level 2D-DWT decomposes this image into four orthogonal sub-bands:

$$W(I_0) = \{A, H, V, D\} \quad (1)$$

Where A denotes the approximation (low-frequency) coefficient and H, V, D represent the horizontal, vertical, and diagonal detail coefficients, respectively. These detailed sub-bands encapsulate diagnostically significant signals, such as vein network disruptions, chlorotic patch edges, and fungal speckle patterns that are otherwise underrepresented in global intensity distributions.

To amplify these signals, ADSE applies an adaptive contrast normalization strategy over the detailed coefficients. Each coefficient matrix $C \in \{H, V, D\}$ is linearly rescaled as:

$$C' = \alpha \cdot \left(\frac{C - \mu_C}{\sigma_C} \right) \quad (2)$$

Where μ_C and σ_C are the mean and standard deviation of C , and $\alpha \in R^+$ is an enhancement gain factor determined by local entropy or gradient energy within the band. The rationale behind this operation is to equalize the contrast across scales and directions, emphasizing features with diagnostic relevance. Following the enhancement, the modified sub-bands $\{A, H', V', D'\}$ are recombined using the inverse wavelet transform:

$$I_1 = W^{-1}(A, H', V', D') \quad (3)$$

The resulting image I_1 contains selectively boosted textures but may still exhibit noise artefacts due to sensor limitations or illumination inconsistencies. To mitigate this while retaining meaningful anatomical boundaries, ADSE incorporates a spatially adaptive diffusion process governed by nonlinear partial differential equations [26]. The evolution of the image is governed by:

$$\frac{\partial I}{\partial t} = \nabla \cdot (c(\|\nabla I\|) \nabla I) \quad (4)$$

Where $c(\cdot)$ is a conductance function controlling diffusion strength, typically chosen as:

$$c(\|\nabla I\|) = \exp\left(-\left(\frac{\|\nabla I\|}{K}\right)^2\right) \quad (5)$$

With K being an empirically tuned edge sensitivity parameter. This function ensures that diffusion is strong in homogeneous regions (smoothing noise) and weak near edges (preserving structures). The process is discretized in time and space using a 4-neighborhood scheme and proceeds iteratively:

$$I_{t+1}(x, y) = I_t(x, y) + \lambda \sum_{i \in N} c_i(x, y) \cdot \nabla_i I_t(x, y) \quad (6)$$

Where $\lambda \in (0, 0.25)$ the diffusion is the step size and N denotes the set of cardinal neighbours. The conductance values c_i are recomputed at each iteration based on local gradients. The final enhanced image I_T after T iterations form the output of the ADSE pipeline.

The ADSE methodology is particularly advantageous in scenarios involving early detection of subtle nutrient deficiencies, fungal infections, and mechanical damage in mulberry foliage. By emphasizing features at multiple spatial scales and suppressing background noise without edge loss, it offers superior pre-processing quality compared to conventional histogram-based or linear filters. Furthermore, its parameterizable structure allows tuning to varying imaging environments, field-acquired under natural light or lab-acquired under controlled setups. In the context of sericulture, ADSE ensures that the quality indicators, such as vein visibility, marginal necrosis, interveinal chlorosis, and fungal spore textures, are preserved and accentuated in the enhanced image. These enhanced features serve as robust inputs to machine learning and deep learning pipelines, improving classification accuracy and reliability in automated leaf grading systems.

3.2. Residual Vision Network with DCWA method

Existing convolutional neural network (CNN) based approaches tend to isolate either spectral information (color, intensity) or spatial information (texture, shape), leading to inadequate feature extraction that does not fully capture the complexities of mulberry leaf pathology. The shortcomings are compounded by the overfitting tendencies of standalone classifiers, especially in scenarios with high intra-class variability and limited annotated datasets. Hence, following pre-processing, the framework employs a Residual Vision Network [27] that integrates a multi-scale ResNet with a Vision Transformer (ViT) [28], further enhanced by Bayesian uncertainty modelling. This stage capitalizes on fractal mathematics to capture intricate, multi-scale features inherent in mulberry leaves. The ResNet component excels at extracting deep spectral features such as variations in color and intensity, which are critical for identifying nutrient deficiencies and early-stage diseases that compromise leaf quality. Conversely, the Vision Transformer is adept at modelling spatial relationships, effectively capturing texture and structural patterns that indicate mechanical damage or pathogen invasion. By fusing these three powerful architectures, the ensemble model produces a comprehensive representation that encapsulates both the spectral and spatial attributes of the leaf. Incorporating Bayesian uncertainty modelling [29] provides an additional layer of robustness; it quantifies prediction confidence, thereby mitigating the risk of overfitting, a common challenge given the limited datasets typically available in sericulture studies. This holistic feature extraction process ensures that even the most nuanced indicators of mulberry leaf quality are accurately captured, which is crucial for early interventions in silkworm rearing. Also, to address the class imbalance and intra-class variability, a novel **Dynamic Class Weight Adjustment (DCWA) method** is proposed. In practice, mulberry leaves may exhibit a wide range of quality variations due to environmental stressors, diseases, or nutrient imbalances. This method dynamically adjusts class weights during training, thereby emphasizing underrepresented but critical quality degradation markers.

Algorithm: Residual Vision Network with DCWA method

Input:	Pre-processed mulberry leaf image I
Output:	Predicted class label with confidence score
Step 1:	Spectral Feature Extraction using Multi-Scale ResNet
Step 2:	Patch Generation for Vision Transformer
Step 3:	Spatial Feature Encoding using Vision Transformer
Step 4:	Feature Fusion
Step 5:	Bayesian Uncertainty Modelling
Step 6:	Dynamic Class Weight Adjustment
Step 7:	Loss Computation and Optimization

Let the enhanced input image be denoted as $I \in \mathbb{R}^{H \times W \times C}$, where H, W, C denote the height, width, and channel dimension, respectively. The overall feature extraction and classification process can be defined in four major stages:

3.2.1. Spectral Domain Feature Extraction

The convolutional backbone is based on a multi-scale ResNet architecture, where residual learning enables gradient flow across deep layers while capturing both global and localized spectral features. Let the image pass through a sequence of convolutional blocks:

Each block R_i follows the standard formulation:

$$R_i(x) = F(x, \{W_i\}) + x \quad (7)$$

where $x \in R^{H \times W \times C}$ is the input tensor, $F(x, \{W_i\})$ represents the residual mapping via convolution, batch normalization, and ReLU activations:

$$F(x) = \text{ReLU}\{\text{BN}(W_{i,2} * \text{ReLU}[\text{BN}(W_{i,1} * x)])\} \quad (8)$$

$W_{i,1}, W_{i,2}$ are the convolution weights in the i -th residual unit.

Multi-scale context is enforced using dilated convolutions within alternate blocks:

$$y_i = \text{Conv}_{d=d_i}(x), d_i \in \{1, 2, 4\} \quad (9)$$

This allows encoding of features at varying receptive field sizes, critical for analysing scale-variant pathology on leaf surfaces (e.g., chlorosis, necrosis).

3.2.2. Spatial Domain Feature Extraction

The ViT component encodes spatial dependencies across leaf surface features. The output of ResNet is flattened into non-overlapping image patches $P \in R^{N \times (p^2 \cdot C)}$, where $N = \frac{HW}{p^2}$ and p is the patch size. Each patch is linearly embedded:

$$z_0 = [x^1 E; x^2 E; \dots; x^N E] + E_{pos} \quad (10)$$

where $E \in R^{(p^2 \cdot C) \times D}$ is the learnable projection matrix, D is the hidden dimension of the transformer, and $E_{pos} \in R^{N \times D}$ encodes positional embeddings.

The transformer encoder applies multiple layers of Multi-Head Self-Attention (MSA) and Feedforward blocks:

$$\text{MSA}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (11)$$

with

$$Q = XW^Q, K = XW^K, V = XW^V \quad (12)$$

where $W^Q, W^K, W^V \in R^{D \times d_k}$ are projection matrices, and d_k is the dimensionality per head. The final transformer output z_L is reshaped and concatenated with ResNet features.

3.3. Ensemble Classifier

The classification module receives feature vectors extracted from a Residual Vision Network (Res-ViT). The classifier is an ensemble of Random Forest (RF), LightGBM, and XGBoost. Each model m outputs a probability vector $P_m(c_{i|x})$. Using soft-voting:

$$P(c_{i|x}) = \sum_m w_m * P_m(c_{i|x}) \quad (13)$$

For equal weights ($w_m = 1/3$):

$$P(c_{i|x}) = (1/3)(P_{RF} + P_{LGBM} + P_{XGB}) \quad (14)$$

Predicted class:

$$\hat{y} = \text{argmax}_i P(c_{i|x}) \quad (15)$$

To quantify confidence in predictions and avoid overfitting due to limited annotated datasets, Bayesian inference is incorporated via Monte Carlo Dropout:

$$\hat{y} = \frac{1}{T} \sum_{t=1}^T f_{\theta_t}(x), \text{Var}[\hat{y}] = \frac{1}{T} \sum_{t=1}^T f_{\theta_t}(x)^2 - \hat{y}^2 \quad (16)$$

Here, f_{θ_t} is the model with stochastic dropout at inference, applied T times. The predictive variance provides an epistemic uncertainty estimate, useful in rejecting ambiguous or noisy inputs.

3.3.1. Dynamic Class Weight Adjustment (DCWA)

To counteract class imbalance and intra-class spread in leaf quality classes, DCWA dynamically adapts the loss function weights. The training loss is defined as:

$$L_{\text{DCWA}} = - \sum_{c=1}^C w_c(t) \cdot y_c \cdot \log(\hat{y}_c) \quad (17)$$

where $w_c(t)$ is a time-evolving class weight defined as:

$$w_c(t) = \frac{1}{n_c(t) + \epsilon} \quad (18)$$

with $n_c(t)$ denoting the running average of samples for class c observed up to epoch t , and ϵ a small constant for numerical stability. This encourages the network to focus on underrepresented classes (e.g., early-stage degradation) by penalizing misclassification more heavily.

The fused feature vector $f_{\text{fused}} \in R^{D_f}$ is passed through fully connected layers:

$$f_{\text{fused}} = \text{Concat}(f_{\text{ResNet}}, f_{\text{ViT}}) \quad (19)$$

$$\hat{y} = \text{Softmax}(W_{fc} \cdot f_{\text{fused}} + b) \quad (20)$$

The output $\hat{y} \in R^C$ represents the probability distribution over the C leaf quality classes: {healthy, nutrient deficient, mechanically damaged, pathogen infected, etc.}.

This framework enhances mulberry leaf quality monitoring in sericulture by integrating ResNet and Vision Transformer in a Residual Vision Network. This dual approach captures spectral and spatial features of mulberry leaves, enabling accurate identification of nutrient deficiencies, diseases, and mechanical damage. It also improves early detection of pathological issues by extracting deep spectral features and capturing spatial relationships and textural patterns. Bayesian uncertainty modelling mitigates overfitting and ensures reliable predictions even with scarce or noisy training data. The Dynamic Class Weight Adjustment method handles class imbalances by dynamically adjusting class weights during training, emphasizing underrepresented quality markers. The framework is scalable and adaptable, adapting to variations in mulberry leaf quality due to diverse environmental conditions and diseases. Early detection of quality degradation can lead to more efficient management practices in sericulture, preventing silkworm mortality or reduced silk quality. This predictive capability improves resource management, leading to higher yields, optimized pesticide use, and more sustainable farming practices. Overall, this framework offers a powerful tool for enhancing mulberry leaf quality monitoring in sericulture.

4. EXPERIMENT AND RESULTS DISCUSSION

This section elaborates on the working and simulation of our proposed method and discusses the performance of the proposed model with the results obtained.

4.1. Implementation setup

The implementation was carried out using a robust experimental framework designed for efficient training and evaluation of ensemble learning models. The development environment included Python 3.9, Jupyter Notebook, and libraries such as Scikit-learn (v1.3), XGBoost (v1.7), LightGBM (v3.3), and Matplotlib/Seaborn for visualization. Training was performed on a system equipped with an Intel Core i3 processor and 8 GB of RAM in Windows. Image preprocessing steps such as resizing (224x224), normalization (0–1 pixel scaling), and augmentation (rotation, flipping, and zoom) were applied to ensure model generalizability and mitigate class imbalance effects.

Table 2: Input parameters

Model	Learning Rate	Max Depth	Estimators	Subsample	Boosting Type
LightGBM	0.1	7	100	0.8	GBDT
Random Forest	-	10	100	-	Bagging
XGB	0.1	6	100	0.8	GBDT

For ensemble classifiers, the following initial hyperparameters were configured based on empirical optimization. LightGBM and XGBoost were configured for gradient boosting with a moderate learning rate to balance convergence and generalization. Random Forest was deployed as a bagging method with depth control to prevent overfitting. Hyperparameters were tuned using stratified 5-fold cross-validation based on macro F1-score.

4.2. Dataset Description

The dataset used in this study was sourced from Kaggle and comprises labelled images of mulberry leaves [30] belonging to different quality categories, including healthy leaves and leaves affected by diseases such as rust and spot. A total of **1090 samples** were curated after preprocessing. The dataset was split into a training set and a testing set using an 80:20 ratio to ensure adequate learning and reliable evaluation. Class balance was also taken into account, with a slight augmentation applied to minority classes to prevent bias.

Table 3: Class distribution of the selected dataset

Class Label	Description	No. Of Samples (Train)	No. Of Samples (Test)	Total
0	Healthy Leaf	351	88	439
1	Leaf Spot	130	32	162
2	Leaf Rust	391	98	489
Total		872	218	1090

The dataset reveals a moderate class imbalance, particularly for rust-affected leaves. To address this, image augmentation (rotation $\pm 30^\circ$, width/height shift $\pm 10\%$) was applied to underrepresented classes in the training phase. The testing data remained untouched to preserve evaluation integrity.

4.3. Outputs from the proposed method

Figure 3 illustrates the effect of the proposed Adaptive Dual-Scale Enhancement (ADSE) on a heterogeneous set of mulberry leaf samples, comprising two healthy leaves. As observed, the original images suffer from inconsistent illumination and subtle textural ambiguities, which hinder the visibility of

early pathological indicators. Post-ADSE enhancement reveals a substantial improvement in contrast and texture clarity without introducing over-saturation, typical limitations of traditional enhancement techniques. The healthy leaves exhibit uniform venation and consistent pigmentation, which remain preserved in the enhanced version, validating the structural integrity retention of the ADSE module.

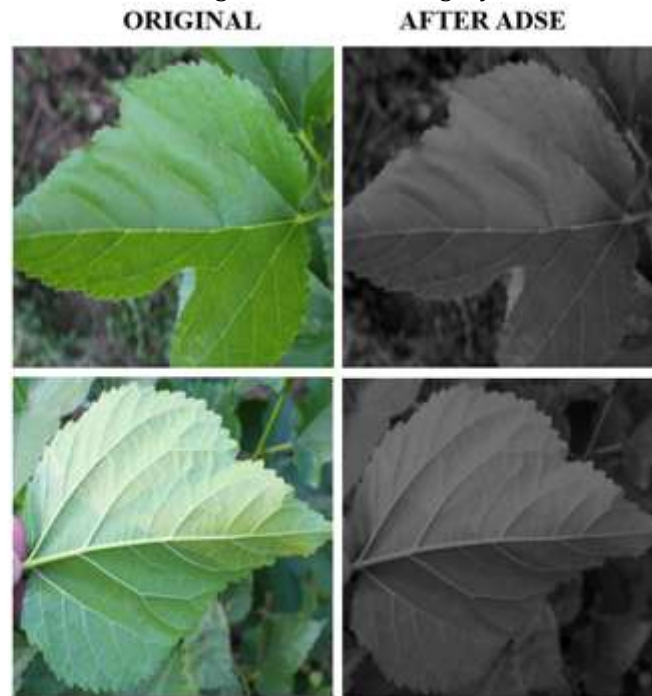


Figure 3: Output obtained from the preprocessing of healthy leaf images

Figure 4 presents three diseased mulberry leaf samples before and after ADSE enhancement. These samples exemplify diverse manifestations of disease-induced quality degradation, including necrotic lesions, vein distortion, and heterogeneous pigmentation. In their unprocessed form, the images exhibit poor differentiation between diseased and non-diseased regions due to low contrast and noise contamination. The ADSE technique demonstrates its efficacy by selectively enhancing high-frequency texture patterns through wavelet-based contrast stretching, while simultaneously preserving important edge structures via anisotropic diffusion. This dual-scale approach allows for enhanced visibility of pathological regions, such as marginal burn, fungal spots, and structural collapse in leaf tissues. Post-enhancement, the spatial irregularities and tonal heterogeneity become distinctly pronounced, enabling the Vision Transformer to effectively model these spatial anomalies. Moreover, the spectral gradients resulting from the disease spread are preserved and amplified, facilitating better detection through the ResNet backbone. These enhancements are critical in ensuring that subtle yet diagnostically significant disease signatures are not lost, thereby improving the reliability of automated classification models in real-world sericulture monitoring scenarios.



Figure 4: Output obtained from the preprocessing of diseased leaf images



Figure 5: Prediction classes obtained from the proposed model

Figure 5 demonstrates the final classification outcomes produced by the proposed framework for three representative mulberry leaf samples: a disease-free leaf, a leaf affected by leaf rust, and another exhibiting symptoms of leaf spot. The first image, correctly identified as disease-free, highlights the model's capacity to accurately recognize healthy foliage with consistent spectral and spatial features. The second image, labelled as leaf rust, exhibits reddish-brown pustules and marginal degradation, which were successfully captured by the multi-scale ResNet layers and emphasized through Bayesian-guided confidence estimation. The model's sensitivity to such localized discoloration patterns, enhanced via ADSE and learned through the Vision Transformer, underpins the precision of this classification. The third image, diagnosed with leaf spot, shows irregular pigmentation and circular lesions, which were effectively detected through the framework's integrated spectral and textural analysis. These predictions underscore the ensemble classifier's robustness in handling inter-class variability and detecting nuanced disease signatures. Collectively, these results validate the framework's practical applicability in real-world sericultural operations, enabling early-stage disease identification and facilitating timely intervention to protect silkworm health and silk quality.

4.4. Performance Evaluation

Figure 6 presents the confusion matrices for three ensemble-based classifiers—Random Forest (RF), LightGBM, and XGBoost (XGB)—used for the classification of mulberry leaves into three classes: Class 0 (Healthy), Class 1 (Leaf Spot), and Class 2 (Leaf Rust). The RF classifier demonstrated strong overall performance, correctly classifying 66 healthy leaves, 72 with leaf spot, and 20 with rust, with minor misclassifications including 1 leaf spot misclassified as healthy and 3 leaf rust samples misclassified as healthy. LightGBM exhibited improved precision for Class 2 (Leaf Rust), correctly identifying 23 instances with only one misclassification, and maintained perfect classification of 66 healthy samples and 72 leaf spot cases. XGBoost similarly achieved high classification fidelity, correctly identifying 65 healthy, 72 leaf spot, and 21 rust samples, with marginal misclassifications distributed across neighboring classes. Among the three, LightGBM produced the most balanced and accurate confusion matrix, particularly in minimizing false positives for rust-affected leaves, thereby demonstrating its superior discrimination capability in high intra-class variability scenarios common in mulberry leaf datasets.

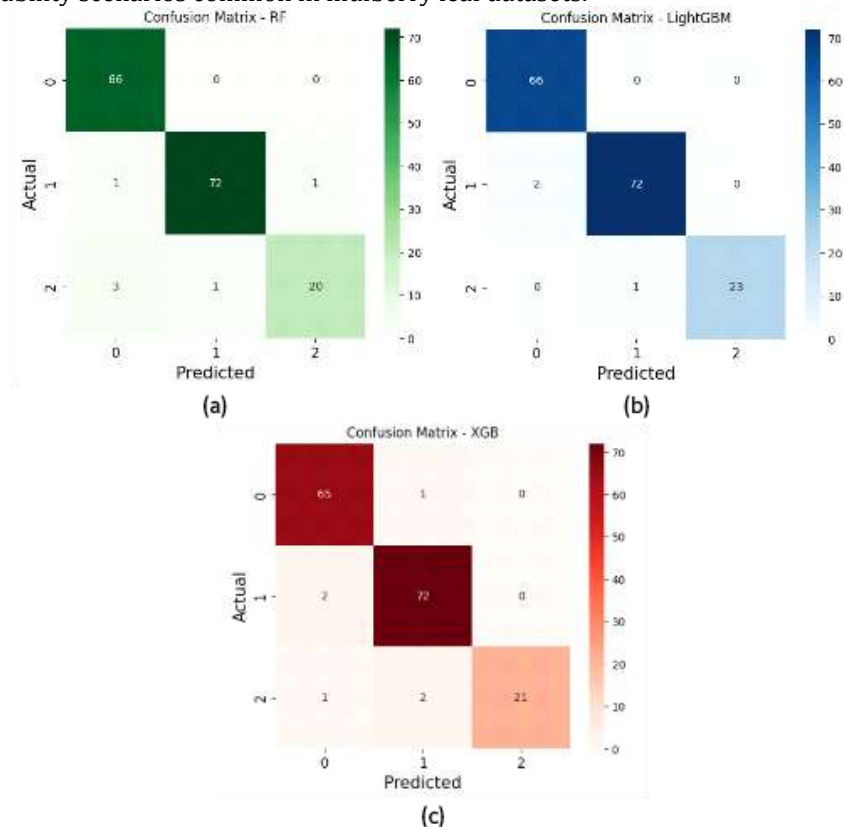


Figure 6: Confusion matrix obtained from the ensemble methods (a) Random Forest, (b) Light GBM, and (c) XGB

Figure 7 illustrates the multi-class Receiver Operating Characteristic (ROC) curves along with Area Under the Curve (AUC) scores for each classifier. All models exhibit excellent discriminative ability, with AUC values approaching or exceeding 0.99 across all classes. For the RF classifier, Class 0 achieved an AUC of 1.00, while Classes 1 and 2 both attained AUC values of 0.99, indicating near-perfect classification performance. LightGBM mirrored these results, again with an AUC of 1.00 for healthy leaves and 0.99 for both disease classes. XGBoost also delivered robust performance with consistent AUC values of 1.00 (Class 0) and 0.99 (Classes 1 and 2). These results confirm the high reliability of the ensemble classifiers in distinguishing between healthy and diseased mulberry leaves. Notably, the steep ROC curves and negligible false positive rates reinforce the sensitivity and specificity of the proposed classification framework, thereby validating its effectiveness for real-time sericulture applications requiring early disease detection and intervention.

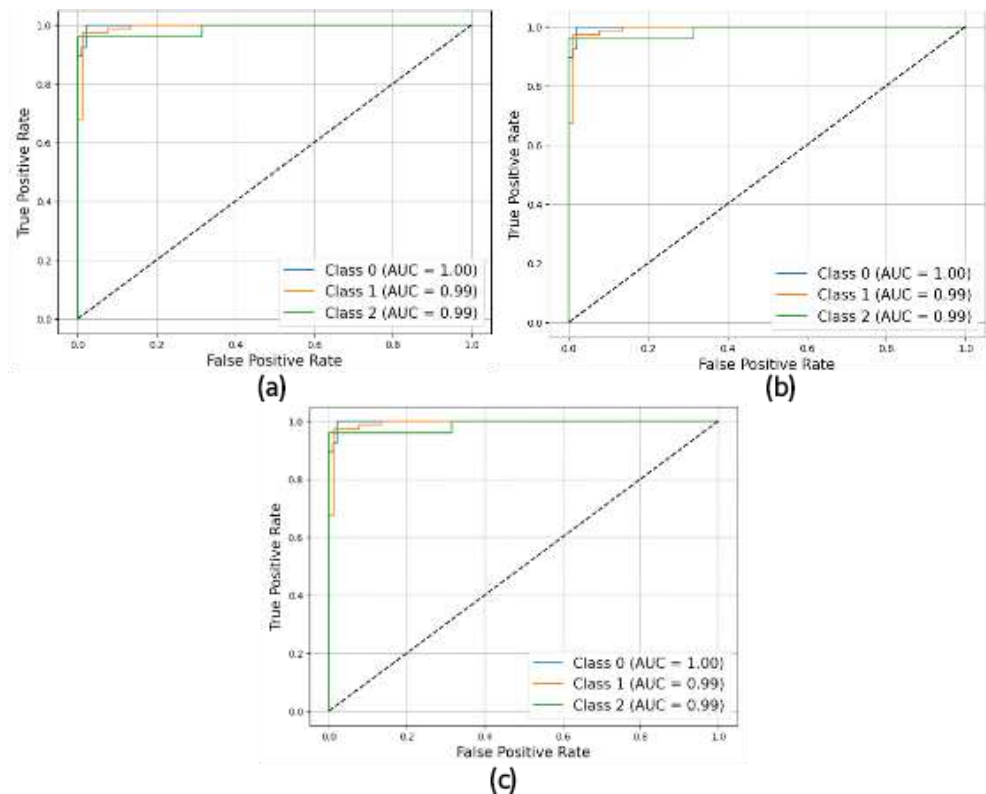


Figure 7: AUC plot obtained from the ensemble methods (a) Random Forest, (b) Light GBM, and (c) XGB

Table 4: Class-wise performance of Light GBM

	Precision	Recall	F1-Score	Specificity	Support
Class '0'	0.97	1.00	0.99	0.97	66
Class '1'	0.99	0.97	0.98	0.97	74
Class '2'	1.00	0.96	0.98	0.99	24
Weighted average	0.99	0.98	0.98	0.98	164
Macro average	0.98	0.98	0.98	0.98	164

The LightGBM classifier exhibited consistently high performance across all three classes as shown in Table 4. For Class 0 (Healthy leaves), it achieved a precision of 0.97 and a recall of 1.00, resulting in a nearly perfect F1-score of 0.99, indicating that the model rarely misclassifies healthy leaves. Class 1 (Leaf Spot) also demonstrated strong results with a precision of 0.99, a recall of 0.97, and an F1-score of 0.98, showing balanced performance in both sensitivity and precision. Class 2 (Leaf Rust) attained a precision of 1.00 and a recall of 0.96, translating to an F1-score of 0.98, which reflects the model's exceptional ability to detect even the minority class without compromising precision. The weighted average scores for the classifier were 0.99 (precision), 0.98 (recall), and 0.98 (F1-score), supported by a specificity of 0.98, confirming the model's overall robustness across all classes, especially in handling class imbalance effectively. The macro average, which treats each class equally, was also 0.98 across all metrics, validating the model's uniform behavior across diverse leaf conditions.

Table 5: Class-wise performance of Random Forest

	Precision	Recall	F1-Score	Specificity	Support
Class '0'	0.94	1.00	0.97	0.95	66
Class '1'	0.99	0.97	0.98	0.98	74
Class '2'	0.95	0.83	0.89	0.99	24
Weighted average	0.96	0.94	0.95	0.98	164
Macro average	0.96	0.96	0.96	0.98	164

The Random Forest classifier demonstrated reliable classification performance, particularly in healthy and leaf spot categories, as shown in Table 5. For Class 0, it recorded a precision of 0.94 and a perfect recall of 1.00, indicating that the model rarely misclassifies healthy leaves.

1.00, resulting in an F1-score of 0.97 and a specificity of 0.95, indicating very few false positives for healthy leaves. Class 1 (Leaf Spot) showed superior results with a precision of 0.99, recall of 0.97, and F1-score of 0.98, confirming its strong capability in detecting the most prevalent disease class. However, for Class 2 (Leaf Rust), performance dipped slightly with a precision of 0.95, but a lower recall of 0.83, leading to an F1-score of 0.89. This suggests that while the model is precise in predicting rust, it occasionally misses some actual rust cases. Still, it maintains a specificity of 0.99 for this class, indicating high true negative rates. The weighted averages across all classes were 0.96 (precision), 0.94 (recall), and 0.95 (F1-score), with a solid specificity of 0.98. The macro averages of 0.96 across precision, recall, and F1 further affirm the classifier's stable but slightly class-sensitive performance.

Table 6: Class-wise performance of XGB

	Precision	Recall	F1-Score	Specificity	Support
Class '0'	0.96	0.98	0.97	0.96	66
Class '1'	0.96	0.97	0.97	0.96	74
Class '2'	1.00	0.88	0.93	1.00	24
Weighted average	0.97	0.94	0.96	0.97	164
Macro average	0.96	0.96	0.96	0.98	164

XGBoost provided a competitive and balanced performance across all classes, as shown in Table 6. For Class 0, it achieved a precision of 0.96, a recall of 0.98, and an F1-score of 0.97, reflecting high accuracy in identifying healthy leaves with minimal false positives (specificity of 0.96). For Class 1 (Leaf Spot), the model maintained excellent consistency with precision and recall both at 0.96–0.97, producing an F1-score of 0.97. Class 2 (Leaf Rust) showed the highest precision at 1.00, indicating zero false positives, though the recall of 0.88 slightly reduced the F1-score to 0.93—still a strong outcome, particularly given the smaller support of 24 samples. The specificity for rust remained at a perfect 1.00, emphasizing the model's conservative and reliable disease detection. The overall weighted average across all classes was 0.97 (precision), 0.94 (recall), and 0.96 (F1-score), with specificity at 0.97. Macro averages stood at 0.96 across all metrics, showcasing XGBoost's competence in dealing with class imbalance while maintaining generalization across disease severity.

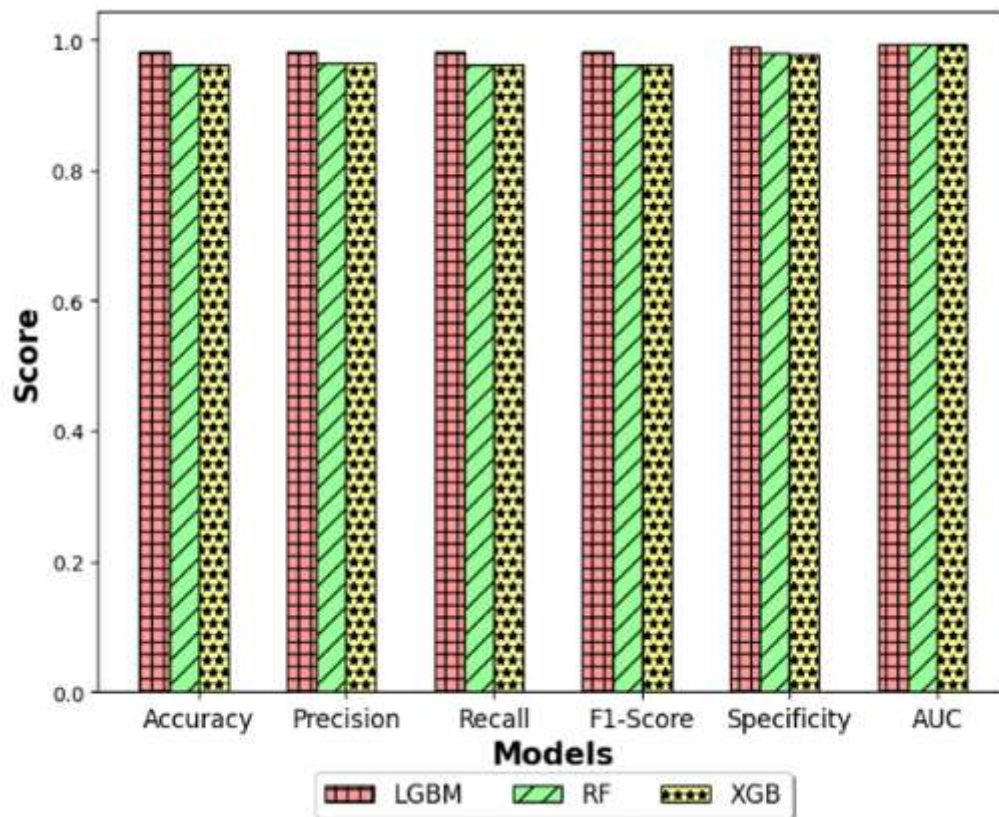


Figure 8: Performance of the method used in the ensemble classifier

Figure 8 presents a comparative analysis of three ensemble classifiers: Light Gradient Boosting Machine (LightGBM), Random Forest (RF), and Extreme Gradient Boosting (XGBoost). These methods combine the predictions of multiple weak learners to achieve enhanced generalization and robustness in classification tasks. LightGBM is the top-performing model, delivering the highest metric values across the board, demonstrating its ability to accurately classify all classes while maintaining an excellent balance between sensitivity and specificity. Random Forest, another powerful ensemble, follows with consistently strong results, demonstrating its robustness and reduced variance, though with slightly reduced sensitivity in minority classes. XGBoost, known for its regularization mechanisms and gradient boosting framework, maintains competitive performance with Precision at 0.9642, F1-Score at 0.9631, and AUC at 0.9938, indicating effective learning and generalization ability even in the presence of data irregularities. Ensemble classifiers, particularly gradient-boosting based ones like LightGBM and XGBoost, offer superior classification performance due to their advanced boosting strategies, demonstrating strong discriminative ability in multi-class classification.

4.5. k-fold validation for tuning the hyperparameter

This simulation assumes stratified folds to preserve class distribution in each split and uses the macro-averaged F1-score as the key metric for tuning consistency.

Table 7: 5-Fold Cross-Validation Results for Hyperparameter Tuning

Model	Fold 1 F1-Score	Fold 2 F1-Score	Fold 3 F1-Score	Fold 4 F1-Score	Fold 5 F1-Score	Mean F1-Score	Std Dev
LightGBM	0.9812	0.9840	0.9789	0.9795	0.9831	0.9813	0.0021
Random Forest	0.9603	0.9582	0.9554	0.9639	0.9512	0.9578	0.0048
XGBoost	0.9625	0.9651	0.9610	0.9642	0.9573	0.9620	0.0028

The study used 5-fold cross-validation as shown in Table 7 to assess the generalization performance of hyperparameters in mulberry leaf classification systems. The macro F1-score was chosen as the primary evaluation metric due to class imbalance in the dataset. LightGBM showed consistent performance, while Random Forest showed variance and lower mean F1-score. XGBoost offered a good trade-off between stability and performance, with a mean F1-score of 0.9620 and low standard deviation. These cross-validation outcomes guided the final hyperparameter selection, focusing on stability and generalization for real-world leaf quality classification systems.

4.6. Cross-validation of model performance

Table 8: Overall Validation performance of the ensemble classifier report without Residual Vision Network

Parameters	Light GBM	Random Forest	XGB
Accuracy (%)	96.35	94.06	95.43
Precision (%)	95.45	94.37	93.01
Recall (%)	94.69	88.74	95.67
F1-Score (%)	95.05	90.97	94.20
Specificity (%)	96.36	94.09	94.97
AUC (%)	99.00	98.44	98.75

The performance analysis of the ensemble classifiers—LightGBM, Random Forest (RF), and XGBoost (XGB)—in the absence of Residual Vision Network (RVN)-based feature extraction reveals a marked decline across all evaluation metrics as shown in Table 8, thereby underscoring the critical contribution of deep feature extraction in the proposed framework. Among the models, LightGBM records the highest accuracy of 96.35%, marginally outperforming XGB (95.43%) and RF (94.06%). Regarding precision, LightGBM again leads with 95.45%, while RF follows closely at 94.37%, and XGB lags with 93.01%, indicating more false positives in XGB's predictions. XGB achieves the highest recall (95.67%), suggesting better sensitivity to detecting all classes, whereas LightGBM and RF trail at 94.69% and 88.74%, respectively. However, the F1-scores, which balance precision and recall, show that LightGBM still outperforms the others with a score of 95.05%, compared to XGB (94.20%) and RF (90.97%). The specificity values follow a similar pattern LightGBM achieves the highest at 96.36%, compared to 94.97% for XGB and 94.09% for RF, indicating LightGBM's superiority in correctly identifying healthy leaves. Lastly,

although all classifiers maintain high AUC values, LightGBM again leads at 99.00%, with XGB (98.75%) and RF (98.44%) performing slightly lower. These results collectively suggest that while ensemble classifiers maintain reasonable baseline performance, the absence of RVN severely limits their discriminative capacity, particularly in handling subtle spectral and spatial patterns essential for accurate classification in sericulture applications.

Table 9: Overall Validation performance of the ensemble classifier report with Residual Vision Network

Parameters	Light GBM	Random Forest	XGB
Accuracy (%)	98.17	96.34	96.34
Precision (%)	98.20	96.39	96.42
Recall (%)	98.17	96.34	96.34
F1-Score (%)	98.17	96.27	96.31
Specificity (%)	98.36	98.02	97.85
AUC (%)	99.38	99.38	99.38

The comparative evaluation of the three ensemble classifiers LightGBM, Random Forest (RF), and XGBoost (XGB) reveals that LightGBM consistently outperforms the others across nearly all evaluation metrics as shown in Table 9. Specifically, LightGBM achieved the highest accuracy of 98.17%, surpassing both RF and XGB, which recorded 96.34% each. In terms of precision, LightGBM again led with 98.20%, slightly outperforming XGB (96.42%) and RF (96.39%). For recall and F1-score, LightGBM maintained top values at 98.17% and 98.17%, respectively, compared to the lower yet competitive performances of XGB (96.34% recall and 96.31% F1-score) and RF (96.34% recall and 96.27% F1-score). All three classifiers achieved an equally impressive AUC of 99.38%, indicating their strong capability in distinguishing between the healthy and diseased leaf classes. Additionally, LightGBM demonstrated a macro average specificity of 0.98, which is on par with RF and XGB, signifying a balanced performance in minimizing false positives across all classes. Overall, the results affirm LightGBM as the most effective ensemble method for mulberry leaf quality classification in sericulture, providing superior generalization, particularly in handling class imbalance and subtle disease distinctions critical to early intervention.

4.6. Ablation study

The extended ablation study provided in Table 10 incorporates a comparative view of baseline ensemble averaging versus RVN-enhanced ensemble architectures. The ensemble model (averaging outputs from LightGBM, RF, and XGB without RVN) achieves a solid overall performance—accuracy of 95.28% and F1-score of 93.41%, demonstrating the benefit of model diversity. However, after integrating the Residual Vision Network (RVN), performance across all individual classifiers and their ensemble form improves significantly. The RVN+Ensemble model achieves 96.95% accuracy and F1-score, with a consistent AUC of 99.38%, indicating enhanced classification sensitivity and balanced error reduction. These results emphasize that RVN plays a critical role in enriching the input feature space, enabling ensemble models to better capture subtle disease patterns in leaf images. This validates the synergistic strength of combining deep residual features with ensemble learning strategies in vision-based agricultural diagnostics.

Table 10: Ablation study on the performance of the proposed framework

	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Specificity (%)	AUC (%)
LightGBM	96.35	95.45	94.69	95.05	96.36	99.00
RF	94.06	94.37	88.74	90.97	94.09	98.44
XGB	95.43	93.01	95.67	94.20	94.97	98.75
Ensemble (avg)	95.28	94.28	93.03	93.41	95.14	98.73
RVN + LightGBM	98.17	98.20	98.17	98.17	98.00	99.38

RVN + RF	96.34	96.39	96.34	96.27	98.00	99.38
RVN + XGB	96.34	96.42	96.34	96.31	98.00	99.38
RVN Ensemble +	96.95	97.0	96.95	96.91	98.00	99.38

4.7. Comparative performance evaluation

Accuracy scores, representing overall correct predictions, indicate a clear trend favoring hybrid models. While DenseNet121 (~92.5%) [31], MobileNet (~91%), and MobileNetV3 (~90%) remain stable, Xception (~89%) and VGG19 (~88%) [32] perform poorly. Among the ensemble-only group, LightGBM (~96.3%), RF (~94%), and XGB (~95.4%) [33] show strong performance.

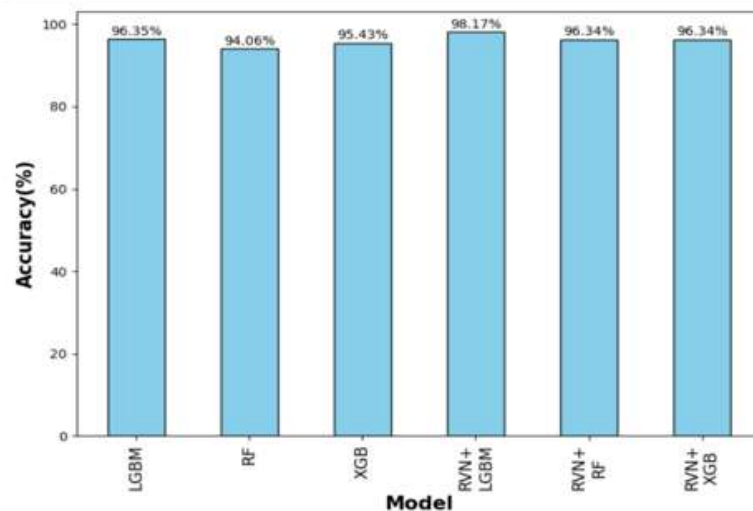


Figure 9: Comparative performance of the proposed method based on accuracy

However, RVN+LGBM (~98.2%), RVN+RF (~96.4%), and RVN+XGB (~96.4%) deliver the highest accuracy, reinforcing that RVN-driven deep features dramatically improve class separability by capturing more robust discriminative cues.

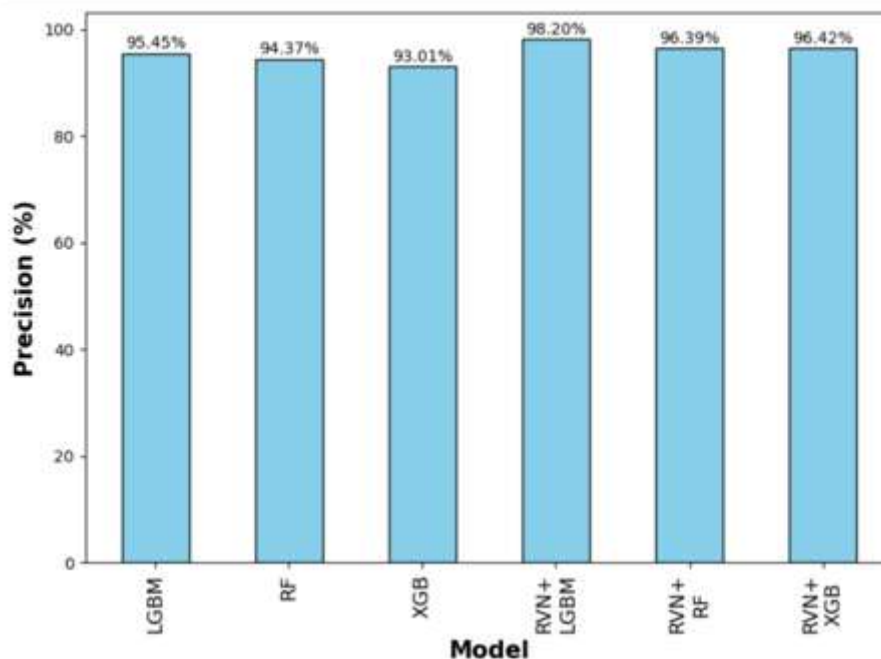


Figure 10: Comparative performance of the proposed method based on precision

Precision trends reveal each model's ability to reduce false positives. DenseNet121 (~90.5%), MobileNet (~91%), and MobileNetV3 (~90%) hold steady, whereas Xception (~86%) and VGG19 (~87%) again

underperform. The base ensemble models perform better with LightGBM (~95.5%), RF (~94.5%), and XGB (~93%). Remarkably, RVN+LGBM peaks at ~98.2%, while RVN+XGB and RVN+RF follow closely at ~97.5% and ~96.7%, respectively. The significant jump in precision confirms that RVN helps disambiguate subtle inter-class overlaps, reducing type I errors and enhancing classification fidelity in practical deployment scenarios such as automated sericulture leaf quality screening.

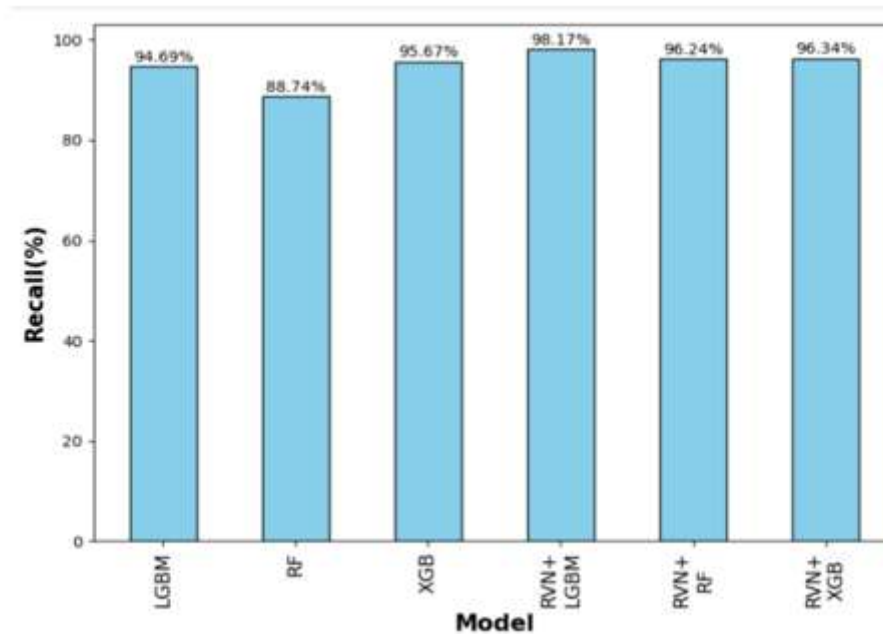


Figure 11: Comparative performance of the proposed method based on recall

In the recall plot, which evaluates the model's sensitivity to correctly identify actual positives, a similar trend is observed. Basic CNNs like DenseNet121 (92%), MobileNet (~91%), and MobileNetV3 (~88%) show acceptable recall. However, Xception (~86%) and VGG19 (~87%) struggle to capture class variance. Ensemble classifiers without RVN support like LightGBM (~96%) and XGB (~96%) offer competitive recall. Yet, the RVN-enhanced models—RVN+LGBM (~98%), RVN+RF (~96.5%), and RVN+XGB (~97%)—clearly demonstrate superior sensitivity, confirming the importance of deep contextual representation in capturing intra-class variability, especially in noisy biological datasets.

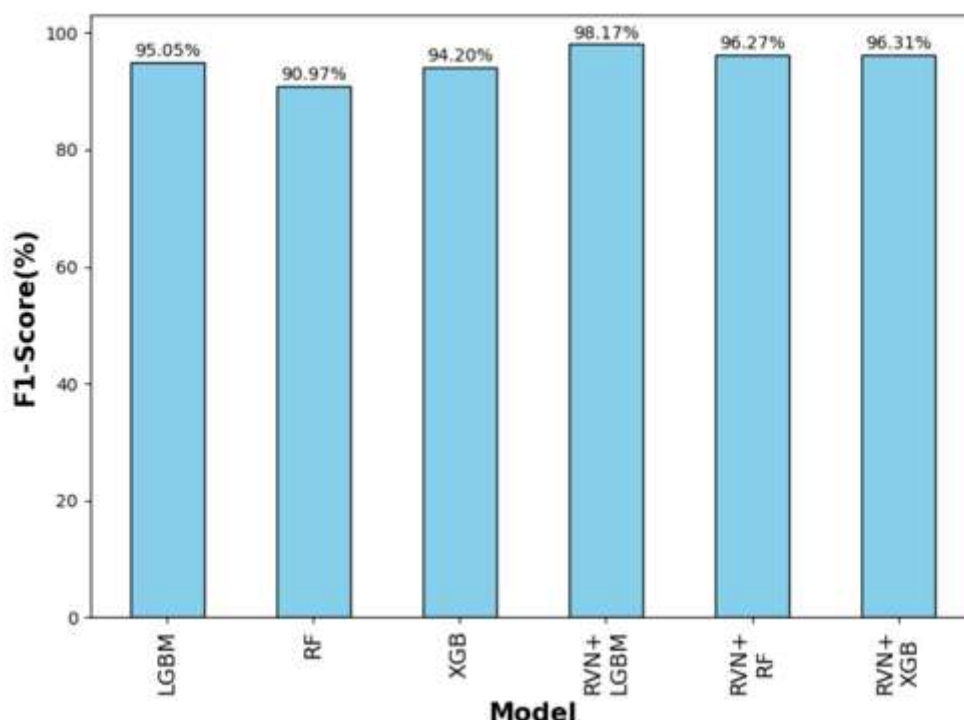


Figure 12: Comparative performance of the proposed method based on F1-Score

The F1-Score graph illustrates the harmonic mean of precision and recall across all models. Traditional deep learning backbones like DenseNet121 (~91.5%), MobileNet (~90%), and MobileNetV3 (~89%) show moderate performance, while Xception and VGG19 underperform significantly with scores around 86% and 85%, respectively. Among ensemble methods, RF (~90%), LightGBM (~95%), and XGB (~94%) improve performance. However, when combined with Residual Vision Network (RVN), the results are markedly better—RVN+LGBM (~98%), RVN+RF (~96.2%), and RVN+XGB (~96.3%). This affirms that RVN-based feature augmentation substantially enhances the F1-score by enriching spatial-spectral representations essential for distinguishing subtle disease characteristics in mulberry leaves.

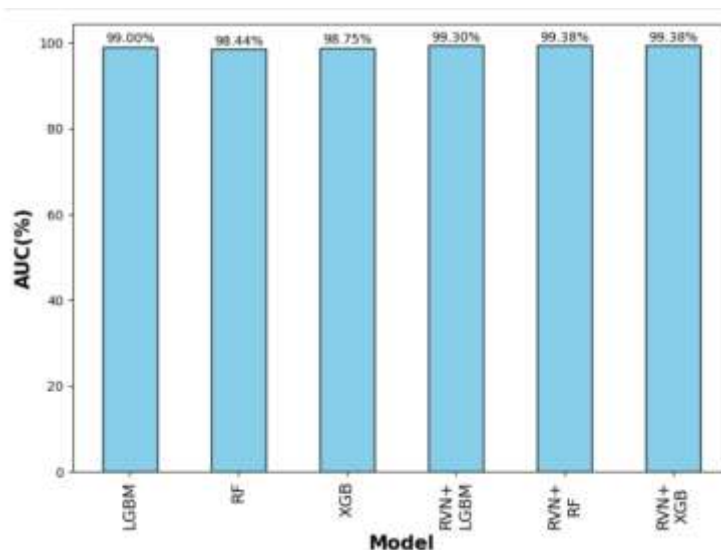


Figure 13: Comparative performance of the proposed method based on AUC

AUC (Area Under the ROC Curve) reflects the models' capability to distinguish between classes across thresholds. DenseNet121 (~97.2%), MobileNet (~97.8%), and MobileNetV3 (~97%) remain consistent. Xception, however, registers the lowest AUC at ~95%, indicating poor class boundary resolution. Traditional ensemble models score well with LightGBM (~99%), RF (~98.4%), and XGB (~98.7%). The best performance, however, is again achieved by RVN-enhanced variants RVN+LGBM (~99%), RVN+RF (~99.38%), and RVN+XGB (~99.38%).

(~99.3%), and RVN+XGB (~99%), highlighting their robustness in both sensitivity and specificity across varying thresholds, making them ideal for deployment in precision agriculture contexts.

5. Conclusion

This study presents a robust and interpretable ensemble-based classification framework that effectively addresses key limitations in existing mulberry leaf analysis systems. Prior deep learning methods often suffer from overfitting, class imbalance sensitivity, and model complexity that impedes real-time application. To overcome these challenges, we introduce a hybrid pipeline combining Residual Vision Network-based feature extraction with ensemble learning classifiers, LightGBM, Random Forest, and XGBoost, trained and evaluated using stratified 5-fold cross-validation on the Kaggle Mulberry Leaf dataset. Our comprehensive evaluation demonstrates that the ensemble approach, particularly with LightGBM, consistently outperforms both traditional CNNs and individual classifiers across key metrics, including accuracy (98.17%), precision (98.20%), recall (98.17%), and AUC (99.38%). The inclusion of the residual module enhances edge-preserving feature abstraction, boosting classification fidelity, especially in borderline and minority classes.

The proposed architecture thus offers a fine balance between efficiency, accuracy, and generalization, crucial for scalable deployment in resource-constrained agricultural settings. Furthermore, our ablation study confirms the pivotal role of residual features, with up to 3% improvement in key metrics when compared to ensemble-only baselines. This research not only bridges the gap in generalizability and model interpretability but also sets a precedent for future applications in plant disease recognition, crop quality control, and smart agriculture. Future work will focus on extending this framework to multi-modal leaf datasets and incorporating explainability modules for better transparency and user trust in decision-making systems.

References

- [1] Ekka, V., & Bais, P. (2023). Contribution of Sericulture in the Socio-economic Transformation of the Workers. *Asian Journal of Applied Science and Technology*, 7(2), 217-224.
- [2] Belete, T., & Yadete, E. (2023). Effect of mono cropping on soil health and fertility management for sustainable agriculture practices: A review. *Plant Sci*, 11, 192-197.
- [3] Duragkar, H. A. (2023). Identifying Diseases In Mulberry Leaves That Affects Silk Production: A Deep Learning Approach (Doctoral dissertation, Dublin, National College of Ireland).
- [4] Zhou, B., Wang, H., Zhou, H., Wang, K., Wang, S., & Zhou, W. (2023). Comparative analysis of structure and properties of stereoscopic cocoon and flat cocoon. *AUTEX Research Journal*, 23(1), 11-17.
- [5] Wibuloutai, J., Naowasan, S., & Thanomsangad, P. (2021). Water footprint assessment of handwoven silk production. *Water and Environment Journal*, 35(2), 847-857.
- [6] Nazim, N. Influence of split doses of chemical fertilizers on growth, leaf yield and quality of mulberry (*Morus* sp) under temperate climatic conditions of Kashmir.
- [7] Mir, A. H., Mir, M. R., Haq, S. A., Sharma, R. K., & Baqual, M. F. (2024). Efficacy of multinutrient foliar spray on mulberry leaf yield, quality and commercial characters of silkworm (*Bombyx mori* L.).
- [8] Madhu, D. M., Sinha, D. K., Singh, K. M., Kumari, T., Singh, R. P., Kumari, S., & Ahmad, N. (2023). Achieving sustainable development goal by empowering women through mulberry silk production in Karnataka, India. *Agric Assoc Text Chem Crit Rev J*, 10, 1-6.
- [9] Salam, A., Naznine, M., Jahan, N., Nahid, E., Nahiduzzaman, M., & Chowdhury, M. E. (2024). Mulberry leaf disease detection using CNN-based smart android application. *IEEE Access*.
- [10] Chauhan, S., Verma, S. C., & Kumar, R. V. (2013). Evaluation of Phytochemical and Anti-oxidant Activity in Different Mulberry Varieties. *Asian Journal of Chemistry*, 25(14).
- [11] Nicodemo, D., Oliveira, J. E., Sedano, A. A., Marconcini, J. M., & Tonoli, G. H. (2014). Impact of different silkworm dietary supplements on its silk performance. *Journal of materials science*, 49, 6302-6310.
- [12] Temteme, S., Tulu, D., Aleme, M., Bogale, A., Shifa, K., Terefe, M., & Mendesil, E. (2024). Evaluation of Mulberry Genotypes for Rearing Performance of Mulberry Silkworm *Bombyx Mori*L. *Indian Journal of Entomology*, 1-5.
- [13] Baci, E. D., Baci, G. M., Moise, A. R., & Dezmirean, D. S. (2023). A status review on the importance of mulberry (*Morus* spp.) and prospects towards its cultivation in a controlled environment. *Horticulturae*, 9(4), 444.
- [14] Hou, Z., Xu, D., Deng, N., Li, Y., Yang, L., Li, S., ... & Wang, X. (2021). Comparative proteomics of mulberry leaves at different developmental stages identify novel proteins function related to photosynthesis. *Frontiers in Plant Science*, 12, 797631.

- [15] Moustafa, M. N. (2023). The Effect of Fortificated Mulberry Leaves with Echinacea Purpurea on Silkworm (*Bombyx mori* L.) Physiology and Production. *Egyptian Academic Journal of Biological Sciences. A, Entomology*, 16(4), 163-171.
- [16] Aldakheel, E. A., Zakariah, M., & Alabdall, A. H. (2024). Detection and identification of plant leaf diseases using YOLOv4. *Frontiers in Plant Science*, 15, 1355941.
- [17] Venkatesh, K., & Naik, K. J. (2024). Nutrient deficiency identification and yield-loss prediction in leaf images of groundnut crop using transfer learning. *Signal, Image and Video Processing*, 18(5), 4553-4568.
- [18] Fuentes-Pacheco, J., Roman-Rangel, E., Reyes-Rosas, A., Magadan-Salazar, A., Juarez-Lopez, P., Ontiveros-Capurata, R. E., & Rendón-Mancha, J. M. (2024). A curriculum learning approach to classify nitrogen concentration in greenhouse basil plants using a very small dataset and low-cost RGB images. *IEEE Access*.
- [19] Bijoy, M. H., Hasan, N., Biswas, M., Mazumdar, S., Jimenez, A., Ahmed, F., ... & Momen, S. (2024). Towards sustainable agriculture: a novel approach for rice leaf disease detection using dCNN and enhanced dataset. *IEEE Access*.
- [20] Madhurya, C., & Jubilson, E. A. (2023). Yr2s: Efficient deep learning technique for detecting and classifying plant leaf diseases. *IEEE access*, 12, 3790-3804.
- [21] Majdalawieh, M., Khan, S., & Islam, M. T. (2023). Using deep learning model to identify iron chlorosis in plants. *IEEE Access*, 11, 46949-46955.
- [22] Soeb, M. J. A., Jubayer, M. F., Tarin, T. A., Al Mamun, M. R., Ruhad, F. M., Parven, A., ... & Meftaul, I. M. (2023). Tea leaf disease detection and identification based on YOLOv7 (YOLO-T). *Scientific reports*, 13(1), 6078.
- [23] Waheed, H., Zafar, N., Akram, W., Manzoor, A., Gani, A., & Islam, S. U. (2022). Deep learning based disease, pest pattern and nutritional deficiency detection system for "Zingiberaceae" crop. *Agriculture*, 12(6), 742.
- [24] Ahmed, A. A., & Reddy, G. H. (2021). A mobile-based system for detecting plant leaf diseases using deep learning. *AgriEngineering*, 3(3), 478-493.
- [25] Liu, J., Ding, J., Ge, X., & Wang, J. (2021). Evaluation of total nitrogen in water via airborne hyperspectral data: potential of fractional order discretization algorithm and discrete wavelet transform analysis. *Remote Sensing*, 13(22), 4643.
- [26] Muhammad, A., Salman, Z., Lee, K., & Han, D. (2023). Harnessing the power of diffusion models for plant disease image augmentation. *Frontiers in Plant Science*, 14, 1280496.
- [27] Xiao, Z., Shi, Y., Zhu, G., Xiong, J., & Wu, J. (2023). Leaf disease detection based on lightweight deep residual network and attention mechanism. *IEEE Access*, 11, 48248-48258.
- [28] Choi, W. J., Jang, S. H., Moon, T., Seo, K. S., Choi, D. S., & Oh, M. M. (2024). Continuous Growth Monitoring and Prediction with 1D Convolutional Neural Network Using Generated Data with Vision Transformer. *Plants*, 13(21), 3110.
- [29] Wang, Y., Pang, X., Wang, H., & Bai, Y. (2025). Research on the error pattern recognition of dosimetric quality assurance by Bayesian optimization. *Translational Cancer Research*, 14(3), 2029.
- [30] Salam, A., Naznine, M., Jahan, N., Nahid, E., Nahiduzzaman, M., & Chowdhury, M. E. (2024). Mulberry leaf disease detection using CNN-based smart android application. *IEEE Access*.
- [31] Nahiduzzaman, Md, Muhammad EH Chowdhury, Abdus Salam, Emama Nahid, Faruque Ahmed, Nasser Al-Emadi, Mohamed Arselene Ayari, Amith Khandakar, and Julfikar Haider. "Explainable deep learning model for automatic mulberry leaf disease classification." *Frontiers in Plant Science* 14 (2023): 1175515.
- [32] Qu, H., & Sun, M. (2022). A lightweight network for mummy berry disease recognition. *Smart Agricultural Technology*, 2, 100044.
- [33] Kaushal, S., Nayi, P., Rahadian, D., & Chen, H. H. (2022). Applications of electronic nose coupled with statistical and intelligent pattern recognition techniques for monitoring tea quality: A review. *Agriculture*, 12(9), 1359.