NL-FuRBe: Precision Diagnosis of Citrus Leaf Diseases using Image Enhancement and Non-Linear Fuzzy Ranking Ensemble Approach

Bobbinpreet Kaur
Lincoln University College
Shashi Kant Gupta
Lincoln University College
Midhunchakkaravarthy Janarthan
Lincoln University College
Deema Mohammed Alsekait
Princess Nourah bint Abdulrahman University
Diaa Salama AbdElminaam
diaa.salama@miuegypt.edu.eg

Benha University

Article

Keywords:

Posted Date: July 3rd, 2025

DOI: https://doi.org/10.21203/rs.3.rs-6898815/v1

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Scientific Reports on September 2nd, 2025. See the published version at https://doi.org/10.1038/s41598-025-16923-4.

NL-FuRBe: Precision Diagnosis of Citrus Leaf Diseases using Image Enhancement and Non-Linear Fuzzy Ranking Ensemble Approach

Bobbinpreet Kaur[1], Shashi Kant Gupta[2], Midhunchakkaravarthy Janarthan[3], Deema Mohammed Alsekait[4], and Diaa Salama AbdElminaam[5,6,*]

[1]Lincoln University College Malaysia
[2]Lincoln University College Malaysia
[3]Lincoln University College Malaysia
[4]Department of Information Technology, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia
[5]Faculty of computers and Artificial Intelligence, Benha University, Benha, Egypt
[6]Jadara Research Center, Jadara University, Irbid, 21110, Jordan
[*]diaa.salama@miuegypt.edu.eg
[+]these authors contributed equally to this work

ABSTRACT

Citrus fruits, especially lemons, play a vital economic and nutritional role worldwide but are increasingly threatened by a wide range of diseases that diminish yield quality and quantity. Traditional manual and automated methods for disease detection requires domain expert, ample observation time, and is often ineffective during early infection stages. This paper presents a novel automated approach for the symptom based detection and classification of citrus leaf diseases using a Non-Linear Fuzzy Rank-Based Ensemble (NL-FuRBE) methodology, enhanced by image quality improvement techniques. The study emphasizes the significance of timely disease diagnosis in citrus crops, which are vital for global food security and economic stability. The methodology begins with image quality enhancement through Vector-Valued Anisotropic Diffusion (VAD) and morphological filtering, evaluated using PSNR, SSIM, and NIQE metrics to ensure optimal visual clarity for classifier input. The core ensemble integrates three deep learning architectures—VGG19, AlexNet, and Xception—using a fuzzy rank-based scoring mechanism built on non-linear transformations (exponential, tanh, and sigmoid functions) to address prediction uncertainty and model bias. A comprehensive dataset of lemon leaf diseases, consisting of 1354 images across nine classes, was utilized for training and evaluation. Experimental results using five-fold cross-validation demonstrate that the proposed model achieves superior performance with an avearge accuracy of 96.51%, outperforming conventional ensemble and state-of-the-art approaches. The results validate the proposed NL-FuRBE as an effective, automated, and cost-efficient tool for precision agriculture and early disease diagnosis in citrus farming.

Introduction

Agricultural and horticultural research aims towards enhancing the food and fruit production and quality while reducing costs and maximizing profits. Fruit trees significantly contribute to a state's economic prosperity. The citrus plant, rich in vitamin C and widely used in the Indian subcontinent, is among the best-recognized fruit species in the citrus family. Citrus fruits and leaves are utilized to manufacture jams, candies, confectionary, and other agricultural goods that promote human health. Timely diagnosis of infections is crucial for enhancing plant yield, a job that presents significant challenges[1].

Table 1. Statistics of Citrus Production Globally (Production in Million Tons, Loss in USD Billions)[10–12]

Country	Prod.	Major Diseases	Infected Fruits (%)	Loss
USA (Florida, California)	15.8	HLB (Citrus Greening), Citrus Canker	40–50% (HLB in FL)	$4.5
Brazil	18.7	Citrus Black Spot, HLB	25–30%	$3.2
China	43.0	HLB, Citrus Yellow Shoot	20–25%	$5.0
India	12.5	CTV, HLB	15–20%	$1.8
Mexico	8.2	Citrus Canker, HLB	18–22%	$1.4
Spain	6.9	HLB, Black Spot	10–15%	$1.0

With around 150 million metric tons yearly production, citrus fruits covers the major commercially important fruit crops grown worldwide. The statistics of citrus production are shown in table 1.Especially important in China, Brazil, the United States, India, and nations in the Mediterranean basin, citrus farming is a vital aspect of their agricultural economy. Many viruses and pathogens able to destroy whole orchards affect the viability and output of citrus trees, therefore compromising food security and resulting significant financial losses. Timely interventions and crop health management depend on correct diagnosis and fast disease identification[2,3]. Because of their quick spread and sometimes asymptomatic first phases, citrus diseases—including black spot, citrus canker, Huanglongbing (HLB or citrus greening), scab, and many fungal infections—cause great difficulties. Conventional methods of disease detection have mostly relied on laboratory-based molecular diagnostics instruments or hand-examining skilled experts. While these approaches have helped the sector for decades, they are labor-intensive, time-consuming, subjective, and often find diseases only after significant development, therefore causing lasting crop harm[4,5]. The prevention and treatment of crop diseases have consistently been a critical concern in agriculture, evolving from traditional manual techniques to instrumental detection and, more recently, to advanced intelligence technologies. Disease recognition technology has experienced multiple phases of advancement[6]. Contemporary research emphasizes the identification, prevention, and management of prevalent diseases, although there is an absence of systematic research data and efficient identification techniques for developing or geographically unique disease species.

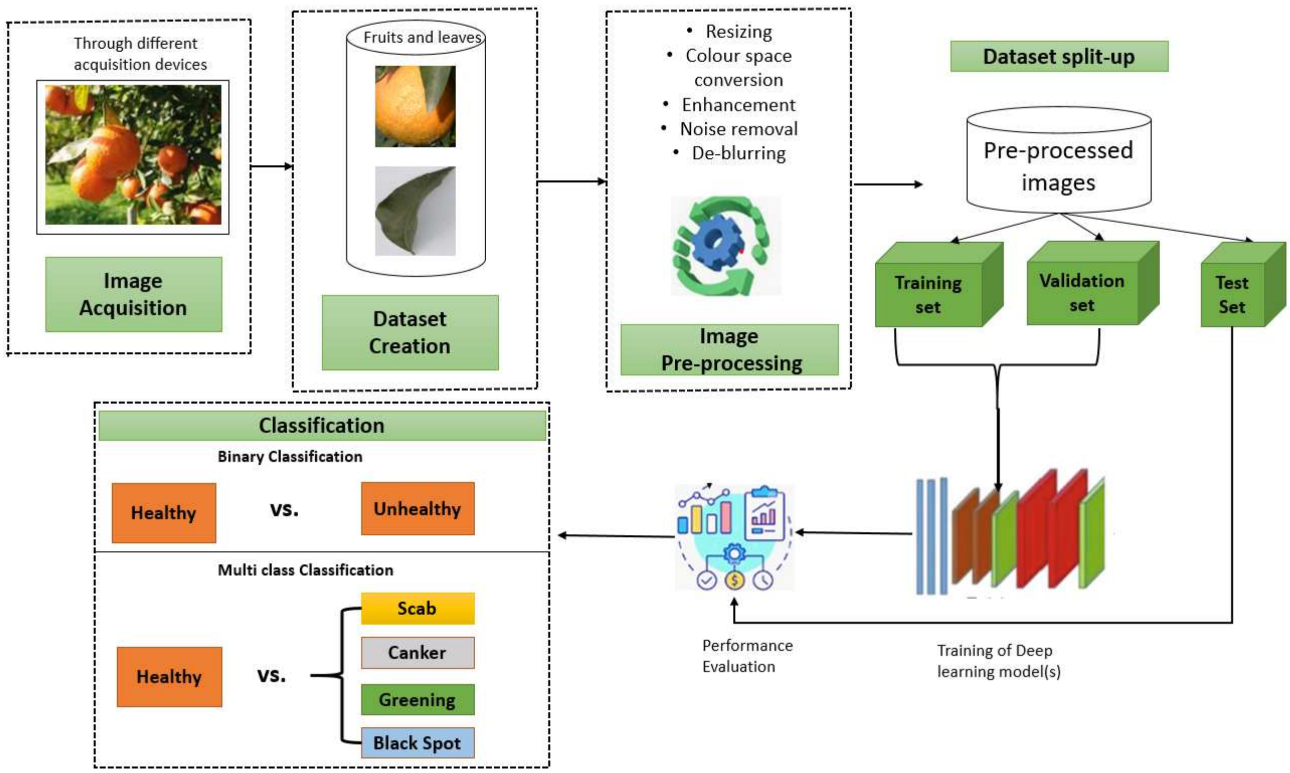

Figure 1. Layout for Citrus Disease Detection Using Deep Learning

The diagnosis of categories of citrus diseases presently depends on symptomatic observation (either through images or manual field examination) and laboratory testing; nevertheless, these procedures frequently exhibit delays. Laborious, time-consuming, and requiring professional involvement are conventional disease detection methods that are more prone to erroneous detection. Deep learning is a disruptive technology in agriculture that, via data-driven decision-making and automation, is changing many aspects of farming. Deep learning, applied with artificial intelligence (AI), enables researchers, agronomists, and farmers to improve crop quality, lower costs, and increase output[7]. Deep learning techniques have lately been a reliable and exact choice for early identification of citrus illnesses. Image-based disease detection has shown remarkable performance for convolutional neural networks (CNNs), transfer learning, and hybrid deep learning models[8], therefore enabling automated, scalable, accurate diagnosis. These systems can effectively identify patterns and symptoms using large databases of images of citrus leaf, therefore helping farmers and agricultural experts to quickly reduce disease spread[9]. The generalized layout of citrus disease detection using deep learning is shown in figure 1.

The development of artificial intelligence and computer vision has revolutionized plant pathology by offering efficient means of automated, quick, exact disease diagnosis. Deep learning systems have become powerful tools for picture analysis-based plant disease diagnosis since they are rather good in identifying small visual patterns connected to different diseases. From traditional Convolutional Neural Networks (CNNs) to advanced Vision Transformers and self-supervised learning models, the evolution of deep learning architectures has steadily raised the precision and dependability of citrus disease detection systems[13].

1 Major Contributions

The major contributions of the work presented in this article include:

- The quality of images available in the dataset is enhanced in order to improve the classifier accuracy.
- Design of an efficient denoising mechanism for better PSNR and improved image quality.
- Developed an ensemble architecture that utilizes various base classifiers and consolidates the individual decisions through a novel fuzzy rank assignment logic.
- Implementation of a Non-Linear Fuzzy Ranking Mechanism for class probability scoring, where the class probabilities of each base learner undergo non-linear transformations based on the exponential and hyperbolic tangent (tanh) functions. These transformed scores are then mapped to corresponding rank values.
- The proposed approach is resistant to biased decisions of individual classifiers, as it effectively provides final predictions based on a robust ensemble strategy.
- The proposed method is automated, computationally efficient, and cost-effective, ensuring the ecological and economic significance of citrus plants and their yields remains maintained.
- The proposed work outperforms existing methods on a newly developed Lemon disease dataset [31].

2 Literature Review

Accurate and automated detection techniques have been developed in response to citrus crops' increasing sensitivity to diseases such citrus canker, black spot, and Huanglongbing (HLB). Conventional manual examination methods produce inconsistent and delayed diagnosis by tedious, time-consuming, often subjective nature. Because deep learning can independently extract hierarchical features from complex image data, it has become a powerful tool in agricultural image analysis in recent years. Emphasizing model designs, datasets, pre-processing techniques, and performance measures, this section reviews present development in citrus disease detection with deep learning models. The convolutional neural network (CNN)[14] can autonomously extract image features that are more expressive and resilient. By autonomously acquiring hierarchical feature representations inside the image, it can enhance comprehension of the image structure and augment classification accuracy. Diverse IoT-enabled gadgets and sensors are extensively employed to capture images of fruit. Typically, hyperspectral functionality yields 2D statistics over time and examines three-dimensional data comprehensively. Based on the filtering mechanism, multispectral imagers and imaging systems are categorized into two types: optical filtration and digitally adjustable filters[15]. The authors in[16] employed transfer learning on the PlantVillage dataset to identify 26 types of illnesses. They attained a classification accuracy of 99.35% with AlexNet and GoogleNet. Citrus fruits such as lime and orange, exhibiting pharmacological degradation, possess diminished economic viability. The scientist presents a methodology for image analysis to identify disease kinds in X-ray images of lemons and oranges. The authors in[17] employed various networks, such as AlexNet, VGG16, GoogLeNet, MobileNetV2, and SqueezeNet, which were trained by transfer learning, in their investigation of tomato leaf disease identification. The experimental findings indicated that the VGG16 network achieved the greatest recognition accuracy of 99.17%. Image enhancement can be utilized to identify and categorize diseases more economically. The suggested approach implements an advanced Image Principal Component Analysis to reduce the dimensionality of orange images. The Edge identification module in the proposed model primarily pulls seven features from the orange fruit picture dataset. A deep learning algorithm is subsequently utilized to identify the circumstances of orange fruit at the initial stages of growth[18]. The authors in[19] employed fine-tuned Caffenet networks to identify several crop leaf diseases, attaining an average accuracy of 96.3%. The authors in[20] employed a CNN model with eight hidden layers, training it on the PlantVillage dataset and ultimately attaining an accuracy of 98.4%. The authors in[21] employed a transfer learning approach to identify six illness categories and healthy tomato photos utilizing pre-trained AlexNet and VGG16 models, achieving accuracies of 97.29% and 97.49%, respectively. The authors proposed[22] enhanced parameter utilization by introducing a weakly densenet-16 network and used a cross-channel feature fusion technique, achieving 93.33% accuracy on a citrus pest and disease dataset. Recurrent Neural Networks (RNNs) are a category of neural network models that incorporate the immediately preceding step in the process output as direct input in the first stage. The predominant applications of RNNs are sequence classification, sentiment analysis, image processing, and video analysis[23]. The authors in[24] created a transformer model that effectively classified tomato leaf images (including both regular and complex backgrounds) into 13 categories, with an accuracy of 93.51%. It necessitates reduced storage capacity and abbreviated training duration, and can be integrated with IoT, drones, and other apparatus for real-time surveillance. The authors proposed[25] discuss many types of convolutional neural networks (CNNs) that were presented. Methods such as Self-Structured Convolutional Neural Networks (SSCNN) and MobileNet were employed for citrus leaf segmentation. The database was initially created utilizing smartphones in this study. Subsequently, both comprehensive models underwent training utilizing identical datasets of citrus plants. The Mobile Net CNN training achieves a peak accuracy of 98%, with a verification accuracy of 92% at the 10th iteration. The authors proposed[26] proposed a method for identifying citrus diseases with deep metric learning, specifically tailored for sparse data. This framework analyzes data utilizing resource-limited devices such as mobile phones. It integrates a network patch utilizing class action with distinct modules (focus, collection, and fundamental neural network classes) to identify different citrus diseases. Implement a deep metric-based composition by partitioning the leaf into segments. The precision of both categories was 95.04%. Classification accuracy was employed as a metric for evaluation, with findings given following five-fold validation. The authors proposed[27] thoroughly assessed the utilization of IoT and DL models for the monitoring and categorization of plant diseases. The benefits and drawbacks of various architectures were examined, and deficiencies in current research were recognized. Their performance on publicly accessible datasets was analyzed to establish a benchmark for identifying the ideal model. In summary, deep learning, characterized by automatic feature extraction, advanced semantic comprehension, robust generalization capabilities, streamlined processing, and excellent scalability, has demonstrated exceptional effectiveness in picture classification. Deep learning, as a prevalent contemporary approach, offers robust theoretical foundations and practical insights for the categorization of citrus leaf diseases[28]. A contemporary AlexNet architecture employing deep learning has been effectively applied for disease detection. A 2 stage Deep CNN model was proposed in[29]. The major novelty s extraction of target area using region proposal network followed by classification into a subsequent disease class. The value of achieved accuracy in detection is 94.37%.Particularly targeting Greening disease detection, the authors in[30] developed a faster RCNN architecture which achieved an Average Precision of 84.13%

2.1 Reseacrh Gaps

Although basic machine learning and deep learning approaches have demonstrated efficacy and widespread application in crop disease prediction, several prior studies encountered challenges in enhancing classification accuracy rates to a significant degree. Furthermore, most research employs standard CNNs (e.g., VGG, ResNet) without customizing them to the distinct visual characteristics of citrus diseases in terms of agricultural imagery.

Table 2. Research Questions and Corresponding Rationales

Research Question	Rationale for Question Design
RQ1. What is the impact of image noise on the performance of deep learning-based citrus disease detection systems?	Images captured for citrus disease detection are often affected by various types of noise and distortions. Investigating the impact of such noise is essential to understand how model performance degrades.
RQ2. How effective are deep learning and ensemble models in accurately detecting and classifying citrus leaf diseases from real-world images?	Individual deep learning models (such as CNNs, ResNet, or MobileNet) have demonstrated good accuracy. Ensemble models, which combine predictions from multiple classifiers, offer the potential for improved performance.
RQ3. To what extent does the proposed model outperform state-of-the-art deep learning methods in detecting citrus diseases?	To determine the comparative efficacy of the proposed ensemble approach by benchmarking it against recognized models using quantitative metrics.

Also inadequate parameter and layer selection in the neural network model has resulted in performance loss. Noise present in the images still remains an under-explored challenge thereby poisng an impact on model accuracy. The proposed approach in the first stage caters the impact of noise by developing noise-aware architectures or denoising models for pre-processing The proposed Fuzzy rank based ensemble model employs integration of four base learners based upon their individual rankings for the classification of citrus diseases in both fruit and leaf images. A comparison with base learner's performance and proposed ensemble model is also presented.

2.2 Rationale for Research question design

The study conducted in this work aims to thoroughly examine the practical obstacles and technical deficiencies in the application of deep learning for citrus disease identification. Considering the economic importance of citrus crops and the rising incidence of diseases like HLB, Citrus Canker, and Black Spot, it is essential to establish reliable and precise detection systems. Considering the survey conducted for the existing state of the art techniques to detect diseases prevailing in citrus fruits the questions were designed and are listed in table 2

3 Methods and Dataset

3.1 Dataset and its acquisition

The prevalence of diseases in fruits may lead to widespread loss of the yields. To control the spreading of disease it is important to detect the disease at its onset.

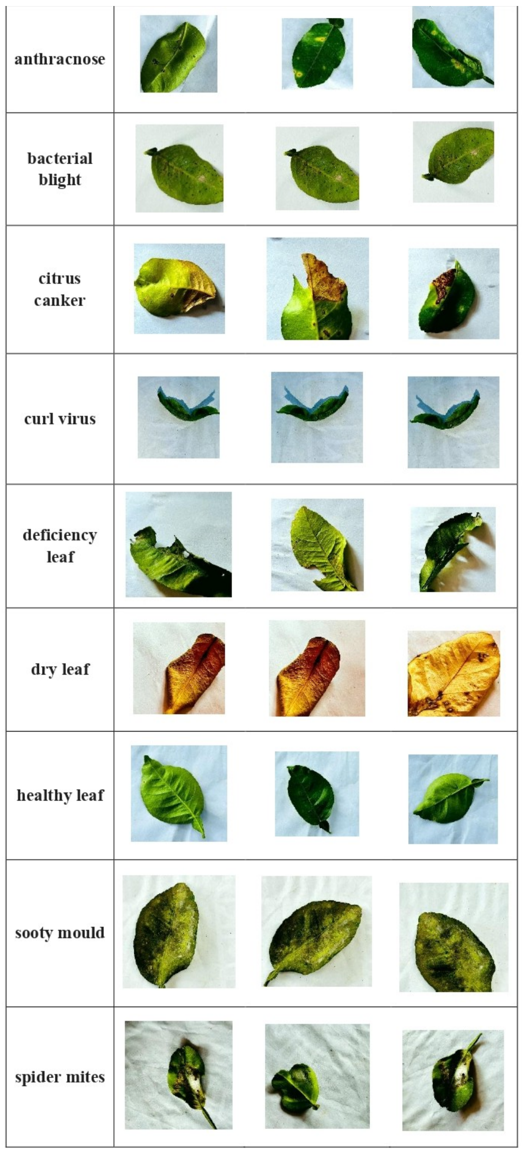

Figure 2. Sample images from different categories of diseases[31]

The design of an automated system for disease detection requires the image dataset to train the deep learning model. The proposed model is trained using public available lemon disease dataset[31] released online in February 2025, consisting of 1354 raw images with 9 class- 8 types of diseases (anthracnose, bacterial blight, citrus canker, curl virus, deficiency leaf, dry leaf, healthy leaf, sooty mould, and spider mites) and Healthy class.This latest dataset chosen for the work reflects the current trends and broader disease types coverage and is captured under wider range of conditions. The sample images belonging to each disease category are shown in figure 2.

3.2 Data Preprocessing and Balancing

Pre-processing and data balancing are essential elements of deep learning, guaranteeing effective model training and applicability in diverse situations. Some classes may be outnumbered by datasets in comparison to others. Data balancing strategies, such oversampling minority classes, under sampling majority classes, or employing SMOTE to create synthetic data, can address this issue. Resizing, normalizing, flipping or rotating images, and noise reduction are all operations that enhance the quality of input during pre-processing. The amalgamation of these operations improves the model's performance, reduces its bias, and enables accurate identification of all classes. The principal benefits of these strategies include adaptability to new data, less overfitting, enhanced consistency in performance across all classes, and increased model accuracy. This research involves the oversampling of minority classes.Figure 3 illustrates a strong class imbalance regarding disease prevailing in citrus fruits, with a highly varying number of images per class. The preconfigured models employed on this imbalanced dataset will exhibit suboptimal performance regarding disease detection in citrus fruits, resulting in diminished accuracy and skewed predictions. Figure 4 illustrates the balanced count for each class, facilitating equitable learning across all classes to enhance prediction and accuracy. The images from each class are reduced to a resolution of 128x128x3 due to varying original resolutions. Random transformations like as orientation, scaling, and flipping are applied to the training dataset. Images in the training dataset are randomly rotated by ±40 degrees and moved horizontally and vertically by 20%, respectively. The training dataset has a 20% variance in magnification adjustments. Furthermore, horizontal flipping of images is executed while preserving their intrinsic significance. During the execution of operations on images, such as rotation, scaling, and flipping, certain pixels may stay unoccupied. The nearest pixel value is utilized to address these gaps.

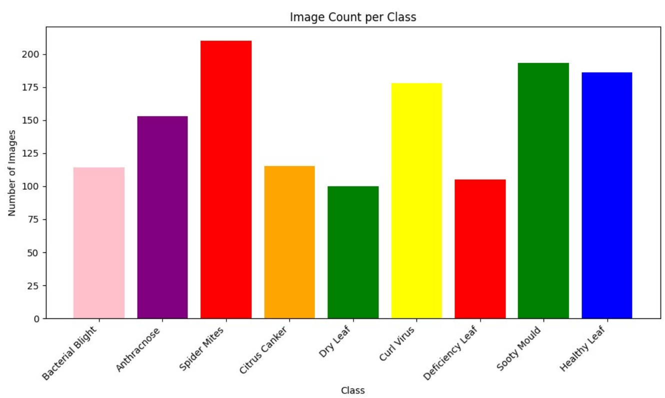

Figure 3. Sample images from different categories of diseases[31]

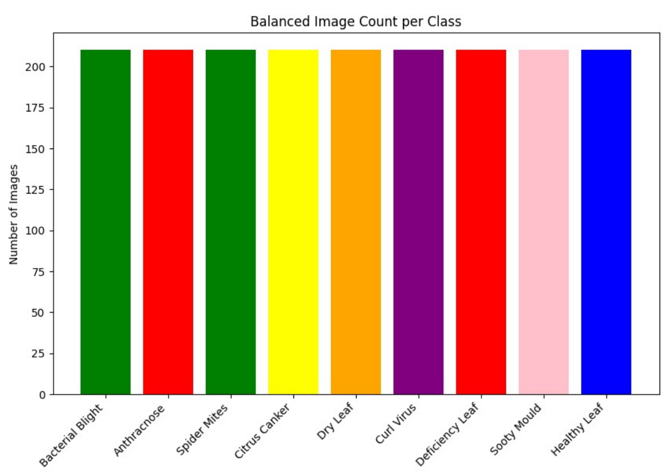

Figure 4. Sample images from different categories of diseases [31]

3.3 Edge preserving Image noise removal

3.3.1 Vector valued-Anisotropic Diffusion(VAD)

In the diagnosis of citrus diseases by deep learning, image quality is crucial for the precise identification and classification of illness symptoms. Images obtained in actual agricultural settings frequently exhibit numerous types of noise, including motion blur, variable lighting, shadows, and background clutter, which might mask essential disease indicators such as lesions, spots, or discolouration. Denoising is an essential pre-processing technique that improves image quality by eliminating undesired distortions. Denoising approaches enhance the visual quality of input data, allowing deep learning models to extract more significant and consistent features, thus augmenting classification accuracy and robustness. Furthermore, denoising enhances the generalizability of models across various field circumstances and imaging devices, such as smartphones and drones. Incorporating classical filters (e.g., Gaussian or median filters) into the pre-processing stage can substantially improve model performance, particularly in resource-constrained or noisy settings. These conventional filters smoothen the image globally thereby blurring of edges. Figure 5 shows the output using VAD for sample image taken from dataset.

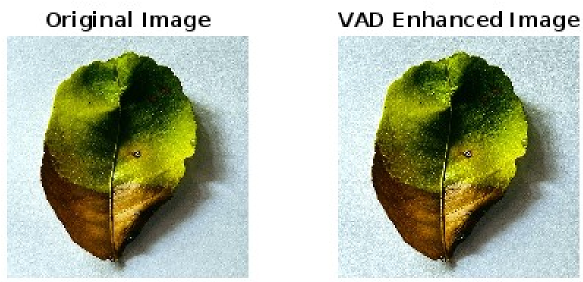

Figure 5. Original and Enhanced image

Also each colour channel is treated separately leading to colour artefacts. While applying filters for denoising the major factor to be considered is preservance of edges. In such context Vector Valued Anisotropic Diffusion method minimizes the noise and also preserve the edges in the image[32]. Each pixel is treated as an independent vector comprising of R, G, B vectors in combined form. The smoothing process is guided using the combined gradient magnitude. Diffusion denotes a method that enhances an image by reducing the disparity in intensity levels among adjacent pixels[33]. Isotropic diffusion uniformly smooths in all directions, resulting in edge blurring. This can be controlled by applying modified approach where the diffusion is varied along different directions. The process is thus known as Anisotropic diffusion[34].
The operation of this filter is governed by the following key components:

- **Diffusion Tensor**:
  This tensor represented through a matrix provides the information of rate and direction of diffusion at a particular point. Expressed using the anisotropic diffusion PDE:

$$\frac{\partial I}{\partial t} = \nabla \cdot (\mathbf{D}(I)\nabla I) \tag{1}$$

  $\mathbf{D}(I)$ represents diffusion tensor a symmetric square matrix, $\frac{\partial I}{\partial t}$ represents derivative of Image intensity value, ∇I represents intensity gradient

- **Structure Tensor**:
  This tensor stores the local gradient information across all channels on the similar instant of time. The information in this vector can be utilised to extract information of edge strength and edge direction.

$$\mathbf{S}_t = \sigma * \left(\sum_{c=1}^{f} \begin{bmatrix} I_{c,x}^2 & I_{c,x}I_{c,y} \\ I_{c,x}I_{c,y} & I_{c,y}^2 \end{bmatrix} \right) \tag{2}$$

  Where: $\mathbf{S}_t$ represents structure tensor, σrepresents Gaussian convolution kernel, $I_{c,x}, I_{c,y}$: Spatial gradients of channel c. In case of R, G,B image the number of channels c=3.

The amount of diffusion is varied across the pixels over the image and the amount is varied as soon as an edge is encountered. This process reduces the diffusion effect on the sharp edges thereby preserving the edge information[35,36]. The gradient for each individual channel is calculated through equation (1). The resultant image can be calculated by equation (3) by combining R G B channels.

$$I_c(x,y) \leftarrow I_c(x,y) + t \cdot \nabla \cdot (\mathbf{D}(I)\nabla I_c) \tag{3}$$

where $I_c(x,y)$ represents intensity value of individual channel.

Algorithm 1 Vector Valued Anisotropic Diffusion
vad.filter()

1.5

1: **Input:** Color image $I \in \mathbb{R}_{m\times n\times c}$ I_{diff} has size $m \times n \times c$, where m and n are spatial dimensions, and c is the number of color channels.
2: **Parameters:** num_iter=3, $\Delta t = 1/7$, κ=30, σ, option= exponential or inverse
3: **Initialize:** $I_{\text{diff}} \leftarrow I$
4: Determine the size of the image. each iteration from 1 to num_iter each color channel $c \in \{R,G,B\}$ each pixel (i,j)
5: Compute gradients:
6: North gradient:

$$\nabla_N I_{\text{diff}}(i,j,c) = I_{\text{diff}}(i+1,j,c) - I_{\text{diff}}(i,j,c)$$

7: South gradient:

$$\nabla_S I_{\text{diff}}(i,j,c) = I_{\text{diff}}(i-1,j,c) - I_{\text{diff}}(i,j,c)$$

8: East gradient:

$$\nabla_E I_{\text{diff}}(i,j,c) = I_{\text{diff}}(i,j+1,c) - I_{\text{diff}}(i,j,c)$$

9: West gradient:

$$\nabla_W I_{\text{diff}}(i,j,c) = I_{\text{diff}}(i,j-1,c) - I_{\text{diff}}(i,j,c)$$

10: Compute combined gradient magnitude:

$$Mag = \sqrt{(\nabla_N I)^2 + (\nabla_S I)^2 + (\nabla_E I)^2 + (\nabla_W I)^2}$$

11: Compute conductance C:
12: **if** option == "exponential" **then** $C = \exp\left(-\frac{Mag}{\kappa}\right)$
13: **else**

$$C = \frac{1}{1 + \left(\frac{Mag}{\kappa}\right)^2}$$

14: **end if**
15: Update pixel value:
16: $I_{\text{diff}}(i,j,c) \leftarrow I_{\text{diff}}(i,j,c) + \Delta t \times (C \times (\nabla_N + \nabla_S + \nabla_E + \nabla_W))$
17: **Output:** Diffused image $I_{\text{diff}}\}$

3.3.2 Morphological- filter module

Top-hat and bottom-hat filters(THBH) are morphological methods employed to enhance features in an image by highlighting minor aspects and details such as bright spots, dark spots, and textures. For image quality enhancement considerations, morphological filter fromed by amalgamation of two filters using elementary morphological operations i.e. Top hat filter and Bottom hat filter is a suitable option. The formation of morphological filter module is elaborated in figure 7. The top hatted version is computed by subtracting the resultant of opening operation between original image and kernel from the original image. Figure 6 shows the output using THBH for sample image taken from dataset.

Figure 6. Original Vs Morphological filtered image

This solves the problems relating to background illumination and perform enhancement of the bright objects with respect to a dark background[37]. The bottom-hatted variant is a dual filter of top-hat filtering. The filtered picture is derived by subtracting the original image from the result of applying a closure operation on the original image using a kernel. This will enhance the picture contrast for effective lesion segmentation. Moreover, the sharp noise peaks are removed. The structuring element(K) plays an important role in the operation of this filter[38] . For top hatted image the features lesser than K are removed from the image and the features greater than k are enhanced. Whereas in bottom hat version the smaller feature elements are preserved. The combined resultant image is thus enhanced[39]. The erosion process entails the removal of specific pixels based on the attributes of the kernel. K, and is implemented as given in Equation (4).

$$I(m,n) \ominus K = \{Z \mid K_Z \subseteq I(r,c)\} \tag{4}$$

Where $I(r,c)$ is the original input image and K is the kernel or structuring element.

Through the process of dilation, more pixels are added to the object boundaries depending upon the characteristics of the kernel K, and is implemented as given in Equation (5).

$$I(r,c) \oplus K = \{Z \mid (K_Z^{\wedge} \cap I(m,n)) \subseteq I(r,c)\} \tag{5}$$

An operation termed **Opening** is constituted by the sequential application of erosion and dilation, utilizing a same kernel element. The outcome is the refinement of contours and the elimination of narrow peaks. It is implemented as:

$$I(r,c) \circ K = (I(r,c) \ominus K) \oplus K \tag{6}$$

A composite operation termed **Closing** is constituted by a sequence of dilation succeeded by erosion, employing an identical kernel element. The outcome is the refinement of contours and the eradication of minor depressions. It is implemented as:

$$I(r,c) \bullet K = (I(r,c) \oplus K) \ominus K \tag{7}$$

3.3.3 Selection of an efficient Image enhancement filter

In the previous section, we have discussed the two categories of filters in order to remove noise from the images as well as preserving or enhancing the image quality. The subjective evaluation of image may not guarantee the effectiveness of the proposed image enhancement filter. Therefore, an objective evaluation is required to be conducted so as validate the performance of the Image enhancement filter[40]. The objective evaluation has been conducted by evaluating the standard image quality metrics listed in table 3 The choice of an effective image enhancement filter is vital for enhancing visual quality and maintaining key image attributes. To objectively measure and compare the efficacy of enhancement techniques, three widely recognized image quality evaluation metrics are utilized: Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index Measure (SSIM), and Natural Image Quality Evaluator (NIQE). PSNR measures the ratio of the highest possible signal power to the noise power impacting the image, with elevated values signifying superior noise reduction. SSIM evaluates perceptual

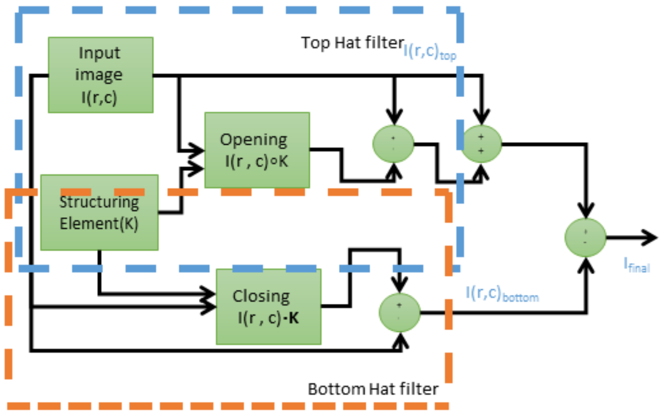

Figure 7. Morphological filter module

Algorithm 2 Top-hat and Bottom-hat Filtering with Disk-shaped Structuring Element
`thbh_filter()`

1: **Input:** I: Input image I(r,c); $radius = 3$; Z: set of positions where structuring element K can fit

2: Create a structuring element:

$$K = \texttt{strel}(\text{'disk'}, \text{radius})$$

3: **Erosion:**

$$I(r,c) \ominus K = Z \mid (K)_Z \subseteq I(r,c)$$

4: **Dilation:**

$$I(r,c) \oplus K = Z \mid \left[(K)_Z^{\wedge} \cap I(r,c)\right] \subseteq I(r,c)$$

5: **Opening (Erosion followed by Dilation):**

$$I(r,c) \circ K = (I(r,c) \ominus K) \oplus K$$

6: **Closing (Dilation followed by Erosion):**

$$I(r,c) \bullet K = (I(r,c) \oplus K) \ominus K$$

7: **Apply Top-hat Filtering:**

$$I_{\text{top}}(r,c) = I(r,c) - (I(r,c) \circ K)$$

8: **Apply Bottom-hat Filtering:**

$$I_{\text{bottom}}(r,c) = (I(r,c) \cdot K) - I(r,c)$$

9: **Output:** Top-hat and Bottom-hat filtered image

$$I_{\text{thbh}} = I(r,c) + I_{\text{top}}(r,c) - I_{\text{bottom}}(r,c)$$

Table 3. Comparison of Image Quality Metrics with Their Mathematical Formulations and Interpretations

Name of Metric	Mathematical Equation	Significance
PSNR (Peak Signal-to-Noise Ratio)	$\text{PSNR (dB)} = 10\log_{10}\left(\frac{R_I^2}{\text{MSE}}\right)$ *Where MSE is the mean squared error.*	Higher value indicates better image strength over noise, implying better enhancement.
SSIM (Structural Similarity Index)	$\text{SSIM}(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}$ *Where:* μ_x, μ_y are means; σ_x^2, σ_y^2 are variances; σ_{xy} is covariance; constants: $C_1 = (K_1L)^2, C_2 = (K_2L)^2$.	Measures luminance, contrast, and structure similarity. Higher SSIM indicates better perceptual image quality.
NIQE (Natural Image Quality Evaluator)	$\text{NIQE} = \sqrt{(\mu_d - \mu_p)^T \left(\frac{\Sigma_d + \Sigma_p}{2}\right)^{-1} (\mu_d - \mu_p)}$ *Where:* μ_d, Σ_d and μ_p, Σ_p are the means and covariances of distorted and pristine image features respectively.	No-reference metric. Lower NIQE scores imply higher perceptual image quality.

similarity of pictures by analyzing brightness, contrast, and structure, so effectively assessing structural fidelity[41]. NIQE, a no-reference metric, evaluates image quality by analyzing statistical deviations from natural scene statistics, with lower values indicating superior perceptual quality[42]. Through the concurrent analysis of these measures, filters may be evaluated with more reliability, facilitating the selection of the enhancing technique that minimizes distortion while maximizing perceptual clarity and naturalness.

Table 4. Comparison of Image Quality Metrics for THBH Filter and VAD method

Disease Image Category	Top Hat Bottom Hat Filter			Vector-Valued Anisotropic Diffusion		
	PSNR	SSIM	NIQE	PSNR	SSIM	NIQE
Anthracnose00005	28.2277	0.9588	8.2514	29.7000	0.9589	5.0717
bacterial_blight00003	21.8589	0.7443	9.3133	22.5717	0.7776	6.6862
Citrus_Canker00009	25.6710	0.9186	7.5630	24.2778	0.9047	5.5003
curl_virus00007	32.4988	0.9516	7.4150	31.8224	0.9444	4.4148
Deficiency_leaf00009	22.9131	0.8605	6.3553	22.8427	0.8660	3.8767
Dry_leaf00007	23.3556	0.8858	7.6135	22.4526	0.8739	6.0716
sooty_mould00002	19.8858	0.8211	8.0138	22.3137	0.8542	4.4302
spider_mites00004	24.5846	0.7980	6.7255	24.8009	0.8241	4.6680
Healthy_leaf00006	29.3078	0.9491	7.3364	29.2762	0.9526	5.4856

To obtain the enhanced images, we have done implementation through MATLAB. The rigorous testing of variants of VAD and THBH filters in order to obtain the optimal values of quality metrics including PSNR, SSIM and NIQE[43]. For implementation of VAD the testing was done considering varying number of iterations and the optimal values of metrics were obtained with iterations=3. At each iteration, directional gradients (North, South, East, West) were computed for each color channel, followed by gradient magnitude calculation and application of the diffusion process to preserve edges while smoothing noise.For conductivity function, the value is calculated using exponential function. Similarly, the testing of THBH filter was done considering different shapes of structuring elements like Disk, Diamond, cube and rectangle with different sizes of radii. The optimal performance was obtained through the disk shaped kernel with radii=3. Table 4 shows the comparitive analysis metric wise for VAD and THBH filter. The implemntation results demosterated in this table shows the values obtained for sample image taken from each image class.

Algorithm 3 Enhanced Image Selection Based on Image Quality Metrics

1.5

Input: VAD image I_{diff}, Top Hat Bottom Hat Filtered Image I_{thbh}
Get VAD image (call `vad.filter()`)
Get Top Hat Bottom Hat image (call `thbh_filter()`)
Compute PSNR(I_{diff}), PSNR(I_{thbh})

$$\text{PSNR} = 10 \times \log_{10}\left(\frac{255^2}{\text{MSE}}\right)$$

Compute SSIM(I_{diff}), SSIM(I_{thbh})

$$\text{SSIM}(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}$$

Compute NIQE(I_{diff}), NIQE(I_{thbh})

$$\text{NIQE} = \sqrt{(\mu_d - \mu_p)^T \left(\frac{\Sigma_d + \Sigma_p}{2}\right)^{-1} (\mu_d - \mu_p)}$$

if PSNR(I_{diff}) $>$ PSNR(I_{thbh}) **and** SSIM(I_{diff}) $>$ SSIM(I_{thbh}) **then**
    if NIQE(I_{diff}) $<$ NIQE(I_{thbh}) **then**
        $I_{enhanced} \leftarrow I_{diff}$
    end if
else
    $I_{enhanced} \leftarrow I_{thbh}$
end if
Output: $I_{enhanced}$

From the value of metrics obtained through implementation, the results obtained for VAD filter satisfies the selection criteria designed in algorithm 3. The testing has been done on sample images belonging to each class of disease and healthy images. In numerous disease categories, the VAD filter produces either elevated or equivalent PSNR and SSIM, while consistently attaining reduced NIQE values, indicating enhanced perceptual quality. For example, in the instance of Anthracnose00005, VAD attained a PSNR of 29.7 and NIQE of 5.0717, in contrast to 28.2277 and 8.2514 for the Top Hat Bottom Hat technique, signifying superior signal fidelity and naturalness. Comparable trends are observed in bacterial_blight00003, curl_virus00007, and sooty_mould00002, wherein VAD markedly enhances NIQE scores while preserving competitive PSNR and SSIM. In several instances, such as Citrus_Canker00009 and Dry_leaf00007, the Top Hat Bottom Hat technique demonstrated marginally superior PSNR or SSIM; nonetheless, it remained inferior in NIQE. This suggests that although both techniques can improve image features, VAD consistently provides a superior equilibrium of objective quality and perceptual realism, rendering it a more reliable option for enhancing images of plant diseases.

4 Proposed Methodology: NL-FuRBE

Fusion in ensemble learning is the technique for aggregating predictions from several base learners into a single output. This stage is essential, as how well the outputs of the base models are merged determines the usefulness of an ensemble in addition to their diversity and correctness. While fusing the predictions of multiple base learners, the underlying problem is giving uniform priority to all the base learners, which results in lower classification accuracy. This section presents a brief overview of the utilized base learners and their related technical aspects, which are later utilized to develop the proposed model based on fuzzy-based ranking. Each base learner generates the value of a confidence score indicating the probability of disease presence. Multiple customized base learners ensure robustness and increased accuracy. The fuzzy-ensemble method subsequently integrates these confidence scores through fuzzy set theory, considering the uncertainty and variability inherent in the scores[45]. The resultant fuzzy ensemble calculated score serves as the deciding criterion for disease diagnosis. This method seeks to enhance the accuracy and reliability of citrus disease detection by leveraging the advantages of various classifiers and integrating uncertainty into the final decision-making process. The weights of the early layers of three base learners are fixed, while the subsequent layers are optimized on the Lemon leaf disease dataset utilized.

1. VGG19

   VGG19 A high-in-depth CNN model, introduced by Simonyan and Zisserman, consisting of 19 weight layers. Out of 19 layers- 16 are the convolutional layers while three are the fully connected layers. Each convolutional block is followed by a max-pooling layer, and the entire network utilizes ReLU activation functions to introduce non-linearity. The uniform architecture, combined with its depth, enables VGG-19 to capture rich hierarchical features, making it highly suitable for tasks such as object recognition, feature extraction, and transfer learning.

2. AlexNet

   AlexNet is a renowned neural network architecture primarily developed by Alex Krizhevsky. It was renowned for competing in the ImageNet Large Scale Visual Recognition Challenge in 2012 and securing the championship. The network attains a 15.3% error rate, which is remarkable since it is 10.8 percentage points lower than the second-place competitor.

   The principle of AlexNet is that a model's depth is crucial for its performance. Consequently, it was developed in CUDA to leverage GPU capabilities for enhanced training speed. Although AlexNet is not the inaugural network trained on a GPU, it has had a significant impact in this domain. It stimulates greater research than training on GPUs and is regarded as one of the most significant articles in the domain of image recognition.

3. Xception

   The Xception architecture, designed by Chollet et al., is derived from the Inception v3 design, maintaining an equivalent number of model parameters while utilizing them more efficiently. They demonstrated that pointwise convolutions and depthwise separable convolutions represent the two extremes of a discrete spectrum, with inception modules positioned centrally. Consequently, they substituted the inception modules with depthwise separable convolutions, resulting in enhanced classification performance without increasing computational costs.

To leverage the strengths of multiple base models and enhance the robustness and accuracy of our predictions, we introduce a novel NL-FuRBE (Non-Linear Fuzzy Rank-Based Ensemble). A Fuzzy Rank-Based Ensemble is an advanced fusion methodology that integrates the advantages of many classifiers or models by allocating fuzzy ranks according to their efficacy or confidence, instead of depending exclusively on definitive voting or averaging. It is particularly advantageous in situations such as disease detection, where model uncertainty and overlapping class borders frequently occur. This approach considers the prediction scores of each base classifier for every individual test case independently. The overall framework of the proposed methodology demonsterated through Figure 8 and step by step process is represented through Algorithm 4.

The NL-FuRBE methodology comprises the following key steps:

4.1 Image quality enhancement

The Images available in the dataset suffers from noise and may impair classifier performance. In order to suppress the impact of noise images were enhanced using a VAD filter explained in Algorithm 1. These images were subjected to image denoisning and edge blurring was well taken care. This forms the enhanced datset and will be used for base models training. The entire dataset is divided into five segments to facilitate 5-fold cross-validation. The initial four folds of the dataset were utilized for training in the first fold. The final fold evaluated all the foundational learners and the proposed fuzzy-rank-based ensemble model.

4.2 Base Model Training

The diverse base learners employed are denoted as $BM_1, BM_2, \ldots, BM_N$. In our proposed methodology, we have employed VGG-19, AlexNet, and Xception as the base learners. Each base model BM_i is trained independently on the training dataset for citrus leaf diseases (CD_{train}) to learn the underlying patterns and relationships within the data. Upon completion of the training phase, each base model BM_i generates prediction scores for a given input sample S_x belonging to the test dataset (CD_{test}). the complete steps are shown in Algorithm 4.

Let the prediction score of each base model BM_i for individual disease or healthy class c be denoted as $P_{i,c}$, where $c \in \{1, 2, \ldots, C\}$ and C is the total number of classes. I represents the number of base models. In our dataset, there are a total of nine different classes. Therefore, the total prediction sets will include values as:

$$\begin{bmatrix} BM_1 \\ BM_2 \\ BM_3 \end{bmatrix} \rightarrow \begin{bmatrix} P_{1,1}, P_{1,2}, P_{1,3}, \ldots, P_{1,9} \\ P_{2,1}, P_{2,2}, P_{2,3}, \ldots, P_{2,9} \\ P_{3,1}, P_{3,2}, P_{3,3}, \ldots, P_{3,9} \end{bmatrix}$$

4.3 Fuzzy Ranking of Predictions

For each test sample S_x, the prediction scores from each base model BM_i are subjected to a fuzzy ranking process through three different nonlinear functions. Let $R_{i,c}^{(1)}$, $R_{i,c}^{(2)}$, and $R_{i,c}^{(3)}$ denote the fuzzy ranks generated by the nonlinear functions for each base model BM_i corresponding to all classes in the dataset. The equations for the nonlinear functions are given as:

$$R_{i,c}^{(1)} = 1 - \tanh\left(\frac{(P_c - 1)^2}{2}\right) \tag{8}$$

$$R_{i,c}^{(2)} = 1 - \exp\left(-\frac{(P_c - 1)^2}{2}\right) \tag{9}$$

$$R_{i,c}^{(3)} = \frac{1}{1 + e^{(-P_c)}} \tag{10}$$

Equation 8 representa a hyperbolic tangent function which is a symmetric function centered around 1. The rank value is calculated based on prediction value- lower the prediction score lower is the rank allocated.

Equation 9 represents an exponential function which assigns rank based on the prediction scores of each base classifier. As soon as the predicted score approach towards 1, the minimum rank value is assigned to a particular classifier for a particular class label.

Equation 10 represents a sigmoid function. The sigmoid is suitable for smoother transitions from higher to lower weights. This will also generates the ranks for each individual classifier based on prediction score. These ranks combined together forms the base for aggregate rank generation.

4.4 Ensemble Aggregation using Fuzzy Ranks

To obtain the final ensemble prediction for a test sample S_x, the fuzzy ranks generated by all base models BM_i are aggregated. For each class c, we calculate an aggregated fuzzy rank R_c^{fused} by combining the ranks from all N base models.

The fused rank score is given by:

$$R_c^{\text{fused}} = \sum_{i=1}^{N} \left(R_{i,c}^{(1)} \times R_{i,c}^{(2)} \times R_{i,c}^{(3)}\right) \tag{11}$$

Algorithm 4 Multi-Class Citrus Leaf Disease Detection

Require: Image dataset with classes: Anthracnose, Bacterial Blight, Canker, curl virus, Deficiency leaf, Dry leaf, Sooty mould, Spider mites, Healthy

Ensure: Predicted class label for each image

Load the multi-class Citrus disease (CD_train) image dataset with C classes.
Balance the dataset using techniques like SMOTE or undersampling.
for each image in dataset **do**
  Call Algorithm 3
  Find Enhanced image $I_{enhanced}$ based on selection criteria
end for
Add $I_{enhanced}$ to the enhanced dataset
Models ← {AlexNet,VGG19, Xception}
for each model in Models **do**

$$layer \leftarrow \texttt{pre\_trained\_model2.layers[i]}$$

$$layer.trainable \leftarrow \texttt{False}$$

$$Predictions[model] \leftarrow model.predict(TestData)$$

$$Scores[model] \leftarrow Evaluate(Predictions[model])$$

end for
for each model score P_c **do** Compute:

$$R_{i,c}^{(1)} = 1 - \tanh\left(\frac{(P_c - 1)^2}{2}\right)$$

$$R_{i,c}^{(2)} = 1 - \exp\left(-\frac{(P_c - 1)^2}{2}\right)$$

$$R_{i,c}^{(3)} = \frac{1}{1 + e^{(-P_c)}}$$

end for
Compute overall RankScore:

$$R_c^{\text{fused}} = \sum_{i=1}^{N} R_{i,c}^{(1)} \cdot R_{i,c}^{(2)} \cdot R_{i,c}^{(3)}$$

Calculate the fused score

$$\text{fusedScore} = \text{rankSum}^{\top} \cdot \text{scoreSum}$$

Select the class with minimum fused score

$$\hat{y} = \min_{\forall C}\left(R_c^{\text{fused}}\right)$$

Output: Final Score

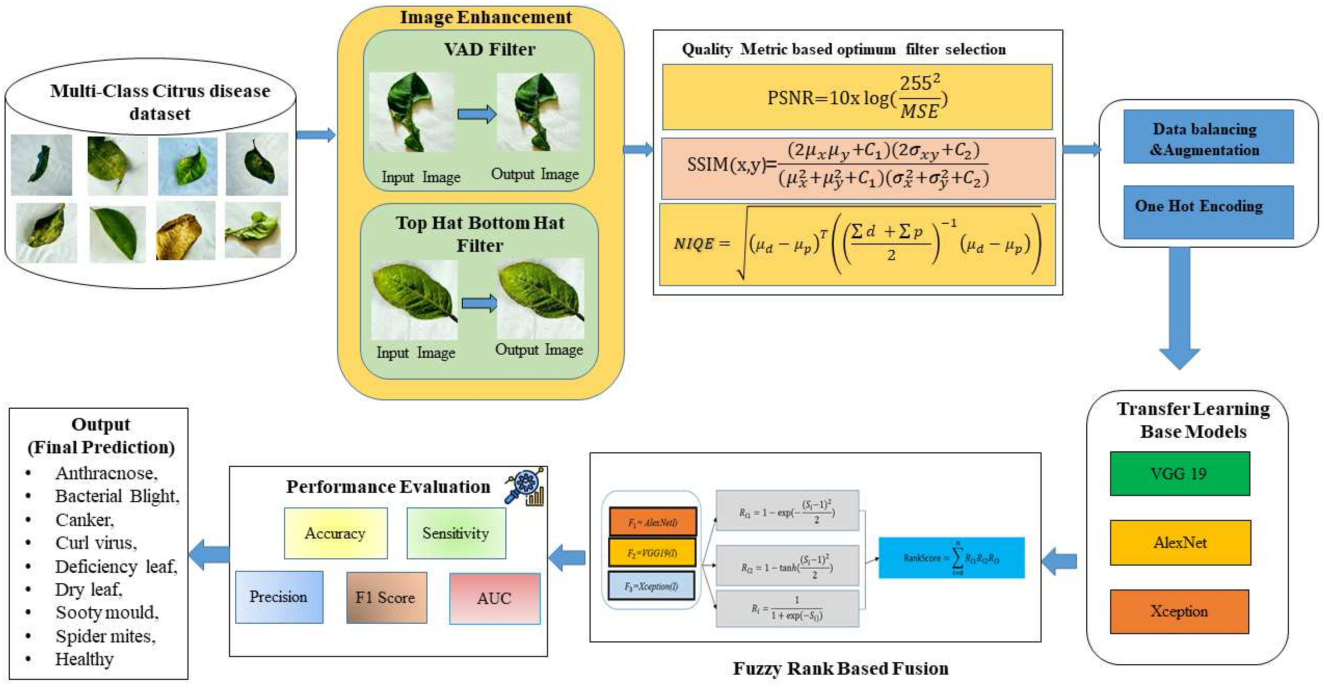

Figure 8. Flow diagram for the Proposed Methodology

The final predicted class $\hat{y}$ for the test sample S_x is the class with the smallest aggregated fuzzy rank:

$$\hat{y} = \arg \min_{c \in \{1,2,\ldots,C\}} \left(R_c^{\text{fused}}\right) \tag{12}$$

The class $\hat{y}$ with the maximum aggregated fuzzy rank is considered the final prediction of the proposed methodology. This decision rule prioritizes the class that receives consistently high confidence and low uncertainty scores across the ensemble of base models thereby improving the classification accuracy.

5 Experimental Results

5.1 Experimental set-up:

The experiments conducted in this study utilize the Python framework with Google Colab as the integrated development environment. The implementation system is configured with 8 GB of RAM and utilizes with hardware accelerator v2-8 TPU. The study aims to classify citrus disease images available through publically available dataset as Anthracnose, Bacterial Blight, Canker, curl virus, Deficiency leaf, Dry leaf, Sooty mould, Spider mites, Healthy employing consistent hyper parameters for the training of three utilized transfer learning models: AlexNet, VGG19 and Xception. Subsequent to input pre-processing through various filters, the images of different categories of diseases are ultimately scaled to 128 x 128 and subsequently utilized in transfer learning models for feature extraction. To split the data into train, test and validation set 5-fold cross validation is used.

5.2 Performance Indicators:

To access the performance of proposed NL-FuRBE method, the performance metrics considered includes- Accuracy, F1-Score, Recall and Precision. Through these metrics we are going to present the tabular results and access the performance of proposed methodology w.r.t existing methods. These metrics access the model's performance quantitatively for the given dataset.

$$\text{Accuracy} = \frac{\sum(T_{pos}, T_{neg})}{\sum(T_{pos}, T_{neg}, F_{pos}, F_{neg})}$$

$$\text{Precision (Specificity)} = \frac{\sum T_{pos}}{\sum(T_{pos}, F_{pos})}$$

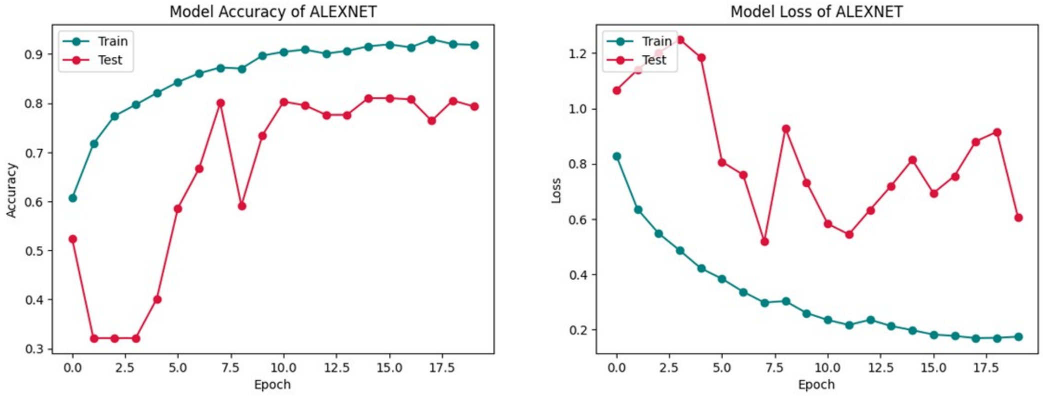

Figure 9. Accuracy and Loss for Alexnet with Conventional Ensemble approach

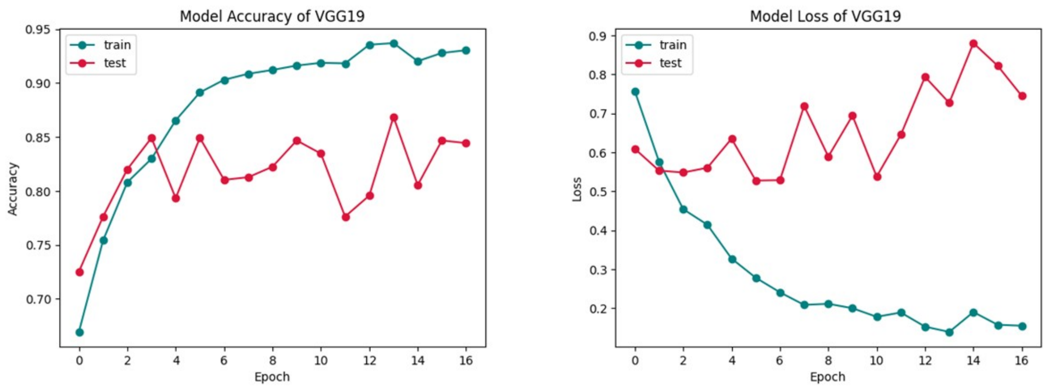

Figure 10. Accuracy and Loss for VGG-19 with Conventional Ensemble approach

$$\text{Recall (Sensitivity)} = \frac{\sum T_{pos}}{\sum(T_{pos}, F_{neg})}$$

$$\text{F1-Score} = 2 \times \frac{(\text{Precision} \times \text{Recall})}{\text{Precision} + \text{Recall}}$$

where $T_{pos}, T_{neg}, F_{pos}, F_{neg}$ stands for True Positives, True Negatives, False Positives, False Negatives respectively.

6 Results and Discussions

This section presents the results obtained after implementation, considering the mentioned experimental setup. We have tested the two methodologies—standard fuzzy-rank-based ensemble and the proposed model explained in the previous section. A comparative analysis considering both methods is also presented.

6.1 Results using Base Learners with fuzzy ensemble model

The standard fuzzy rank-based ensemble is trained using the available dataset and the results are shown and presented in this section. Initially 3 base learners, AlexNet, VGG19 and Xception models were trained and a total of 20 epochs were used. Figure 9 shows the accuracy and loss value for individual AlexNet with the ensemble approach. Accuracy, F1 score, recall, and precision achieved by the AlexNet classifier with the fuzzy ensemble model are 79.5%, 82.4%, 81.4%, and 83.5%, respectively. Figure 10 shows the accuracy and loss value for individual AlexNet with the ensemble approach. Accuracy, F1 score, recall, and precision achieved by the AlexNet classifier with the fuzzy ensemble model are 85.7%, 90.0%, 88.8.4%, and 91.2%, respectively. Figure 11 shows the accuracy and loss value for individual Xception with the ensemble approach. Accuracy, F1 score, recall, and precision achieved by the AlexNet classifier with the fuzzy ensemble model are 88.3%, 91.5%, 91.3%, and 91.7%, respectively.

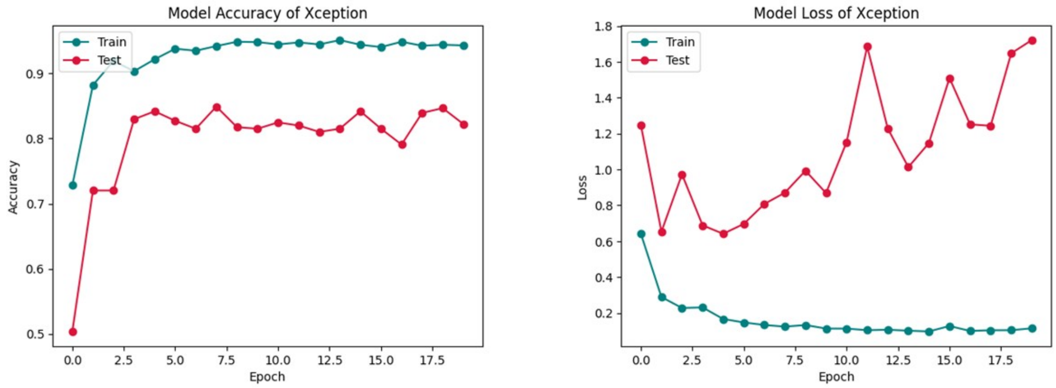

Figure 11. Accuracy and Loss for Xception with Conventional Ensemble approach

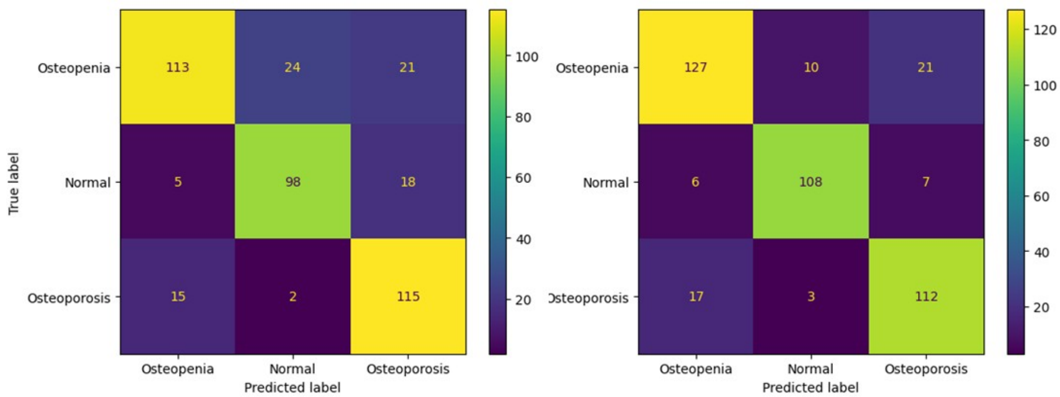

Figure 12. Confusion Matrix for Alexnet and VGG-19 with Conventional Ensemble approach

The confusion matrix illustrates the performance of a Multi-class classification model for different disease variants available in the dataset. Figure 12, Figure 13, show the confusion matrix for the base learners and ensemble model with conventional approach.

6.2 Results for the Proposed Model

The proposed fuzzy rank-based ensemble is trained using the Enhanced dataset (refer algorithm 1 and 3) and the results are shown and presented in this section. Initially, 3 base learners, AlexNet, VGG19 and Xception models were trained, and a total of 20 epochs were used. The proposed approach outperforms the conventional approach, depicting the advantage of applying Image quality improvement before providing the dataset for the classifier training. The image quality was assessed using standard metrics and the results shown support the relevance of applying image quality enhancement to the original dataset. Table 6 shows the model performance for the lemon disease dataset corresponding to the fifth fold.

The whole enhanced dataset obtained after applying enhancement is split into five parts so that a 5-fold cross-validation can be done. The first four folds of the dataset were used for training in the first fold. The last fold tested all the base learners and the suggested fuzzy-rank-based ensemble model. After 20 epochs, Figure 14 shows the accuracy and loss of the AlexNet classifier with the proposed methodology. Accuracy, F1 score, recall, and precision achieved by the AlexNet classifier with the fuzzy ensemble model are 85.2%, 86.3%, 85.8% and 86.9% respectively. Figure 15 shows the accuracy curve, and loss curve of the VGG-19 classifier with the fuzzy ensemble model. Accuracy, F1 score, recall, and precision achieved by the VGG-19 classifier with the fuzzy ensemble model are 90.3%, 91.7%, 91.6%, and 91.7%, respectively. For the Xception classifier, Figure 16 shows the accuracy and loss curve with the fuzzy ensemble model. Accuracy, F1 score, recall, and precision achieved by the AlexNet classifier with the fuzzy ensemble model are 96.5%, 96.4%, 96.4%, and 96.5% , respectively.

Table 5 compare the performance of four classifiers—AlexNet, VGG, Xception, and an ensemble model—under two approaches: fuzzy rank-based methodology and the proposed methodology. Four key evaluation metrics are reported: Accuracy, F1 Score, Recall, and Precision. Figure 17, Figure 18, show the confusion matrix for the individual learners and ensemble model with proposed approach.

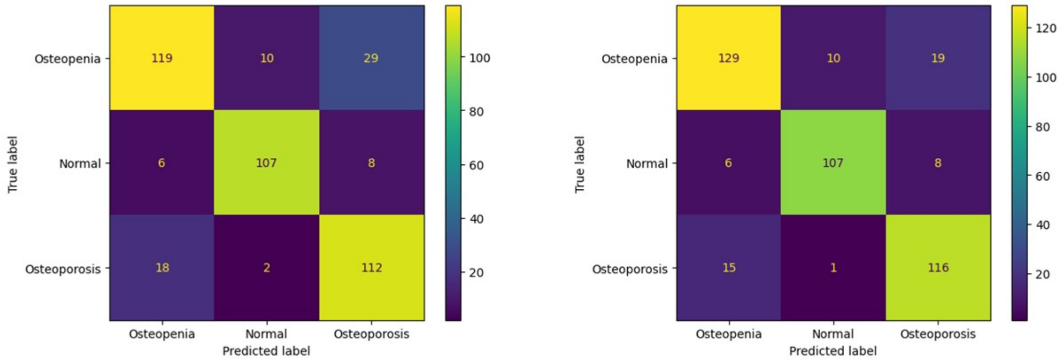

Figure 13. Confusion Matrix for Xception and Ensemble with Conventional Ensemble approach

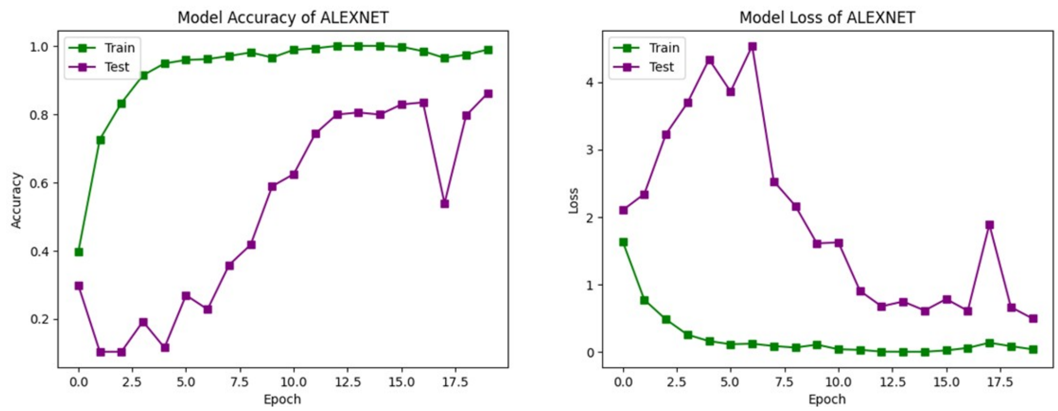

Figure 14. Accuracy and Loss for Alexnet with Proposed approach

The results evidently demonstrate that the Ensemble model outperforms the individual models across all evaluation metrics. AlexNet shows the lowest performance across all metrics, with an accuracy of 85.2% and F1 Score of 86.3%, suggesting its limitations in capturing the complexity of the disease patterns. VGG markedly surpassed AlexNet, attaining an accuracy of 90.3% and an F1 score of 91.7%, because to its deeper architecture that facilitates enhanced feature extraction. Xception somewhat surpassed VGG, achieving an accuracy of 92.2% and a recall of 92.3%, demonstrating robust sensitivity in accurately detecting positive instances. Table 5 presents metric wise comparison for conventional and proposed methodology considering individual base learners and ensembled version.

Table 5. Performance Comparison Between Conventional and Proposed Methodologies

Model	Conventional Fuzzy Rank based Methodology				Proposed Methodology			
	Accuracy	F1 Score	Recall	Precision	Accuracy	F1 Score	Recall	Precision
AlexNet	79.52	82.42	81.41	83.52	85.26	86.31	85.84	86.91
VGG	85.75	90.00	88.81	91.22	90.34	91.79	91.60	91.71
Xception	88.36	91.54	91.32	91.75	92.25	91.82	92.30	91.41
Ensemble	91.46	92.23	92.26	92.36	96.51	96.43	96.49	96.55

Table 6 shows the metric values obtained during 5th fold for different categories of diseases for individual as well as ensemble model. The Ensemble model, integrating the advantages of separate models, attained superior performance across all metrics: Accuracy (96.5%), F1 Score (96.4%), Recall (96.4%), and Precision (96.5%). This illustrates that the ensemble method successfully alleviates the shortcomings of individual models, resulting in enhanced and more generalizable performance. AlexNet showed a marked improvement, with accuracy increasing from 79.5% to 85.2%, F1 score from 82.4% to 86.3%, and similar gains in recall (81.4% to 85.8%) and precision (83.5% to 86.9%). This improvement reflects the benefit of the proposed

Table 6. Class-wise Performance Metrics for different classes- Fold 5

Fold no.	Model	Class	Precision	Recall	F1-score	Accuracy
5	AlexNet	Spider Mites	0.86	0.68	0.79	85.20
		Sooty Mould	1.00	0.81	0.83	
		Healthy Leaf	0.79	0.97	0.89	
		Curl Virus	0.93	0.86	0.84	
		Anthracnose	0.51	0.88	0.85	
		Citrus Canker	0.95	0.85	0.85	
		Bacterial Blight	0.91	0.84	0.88	
		Deficiency Leaf	0.86	0.83	0.84	
		Dry Leaf	1.00	1.00	1.00	
	VGG-19	Spider Mites	1.00	0.75	0.98	90.30
		Sooty Mould	1.00	0.76	0.84	
		Healthy Leaf	0.70	1.00	0.84	
		Curl Virus	1.00	0.87	0.87	
		Anthracnose	0.65	0.89	0.78	
		Citrus Canker	1.00	0.97	0.99	
		Bacterial Blight	0.96	1.00	0.98	
		Deficiency Leaf	0.96	1.00	0.98	
		Dry Leaf	0.98	1.00	1.00	
	Xception	Spider Mites	1.00	0.76	0.86	92.20
		Sooty Mould	0.80	1.00	0.82	
		Healthy Leaf	1.00	0.97	0.97	
		Curl Virus	1.00	0.85	0.92	
		Anthracnose	0.82	1.00	0.91	
		Citrus Canker	0.82	1.00	1.00	
		Bacterial Blight	0.97	0.86	0.99	
		Deficiency Leaf	0.85	1.00	0.80	
		Dry Leaf	0.96	0.86	1.00	
	Ensemble	Spider Mites	1.00	0.91	0.99	96.50
		Sooty Mould	1.00	0.95	0.97	
		Healthy Leaf	0.96	1.00	0.96	
		Curl Virus	1.00	0.85	0.92	
		Anthracnose	0.77	0.98	0.86	
		Citrus Canker	1.00	0.98	1.00	
		Bacterial Blight	1.00	1.00	1.00	
		Deficiency Leaf	0.96	1.00	0.98	
		Dry Leaf	1.00	1.00	1.00	

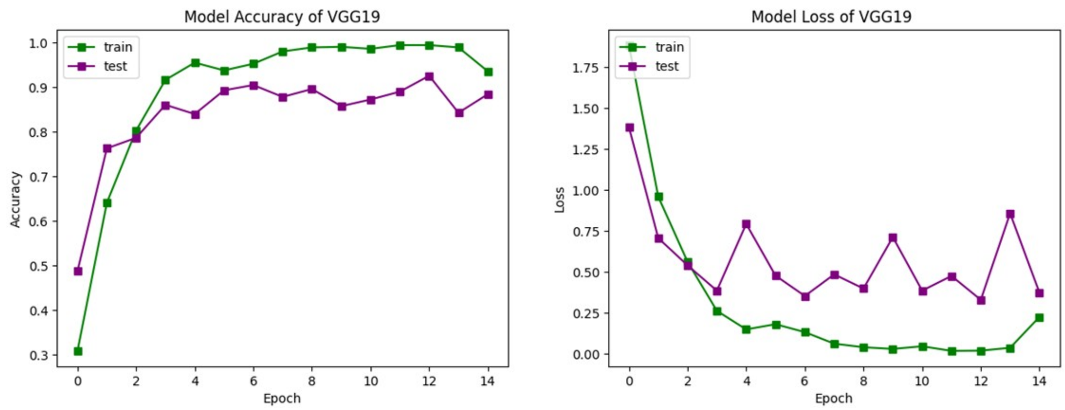

Figure 15. Accuracy and Loss for VGG-19 with Proposed approach

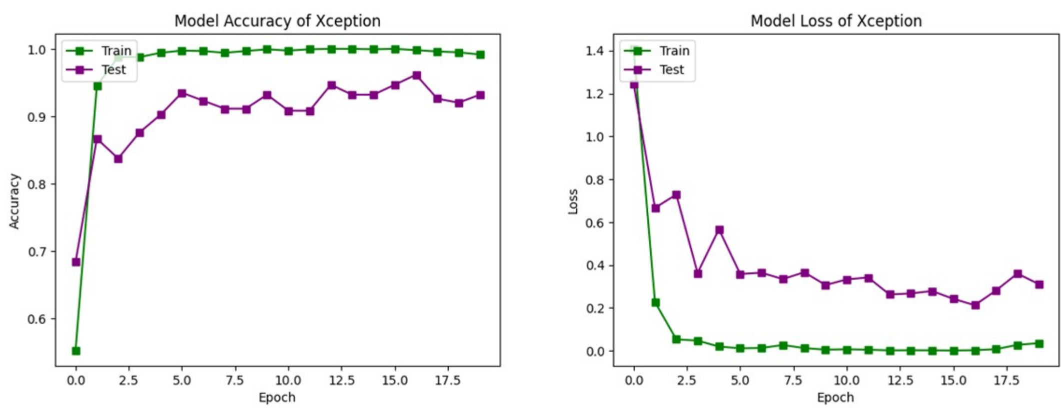

Figure 16. Accuracy and Loss for Xception with Proposed approach

model enhancements even for earlier-generation architectures. VGG displayed notable gains as well. Accuracy rose from 85.8% to 90.3%, and F1 score improved from 90.0% to 91.7%. The model also saw increases in recall (88.8% to 91.6%) and precision (91.2% to 91.7%), highlighting the improved generalization capability of the proposed approach. Xception, already a high-performing model under the fuzzy rank-based methodology (accuracy of 88.4%, F1 of 91.5%), further benefited from the proposed framework. Accuracy reached 92.2%, and recall rose to 92.3%, while maintaining strong precision (91.4%) and F1 score (91.8%). The ensemble model outperformed all individual architectures in both the approaches. Under the conventional method, it achieved an accuracy of 91.5% and F1 score of 92.2%, whereas under the proposed methodology, it reached 96.5% accuracy, 96.4% F1 score, and identical recall and precision values (96.4% and 96.5%, respectively). The application of image enhancing techniques before model training markedly enhanced the overall performance of the proposed method. The preprocessing stage improved image contrast, minimized noise, and emphasized essential characteristics, hence providing the model with superior inputs that facilitated more efficient feature extraction and classification.The enhanced images promoted superior convergence during training, increased accuracy, and higher resilience in prediction results, consequently affirming the significance of image enhancement as an essential phase in the model development process. The results demonstrate that whereas individual models such as Xception exhibit robust performance, ensembling is an effective method for enhancing reliability and accuracy in practical applications, especially where high precision and recall are essential.

6.3 Comparison of Proposed methodology with state-of-the-art techniques

The Table 7 presents a comparative analysis of Proposed methodology with existing methods. The proposed method, which uses an enhanced fuzzy rank-based ensemble on a dedicated 2025 lemon disease dataset, achieved a strong accuracy of 96.51%.

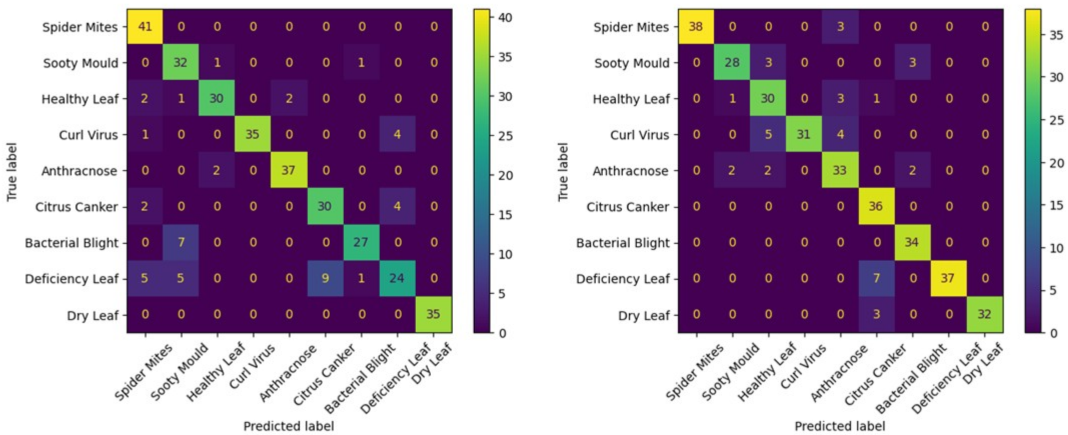

Figure 17. Confusion Matrix for Alexnet and VGG-19 with Proposed approach

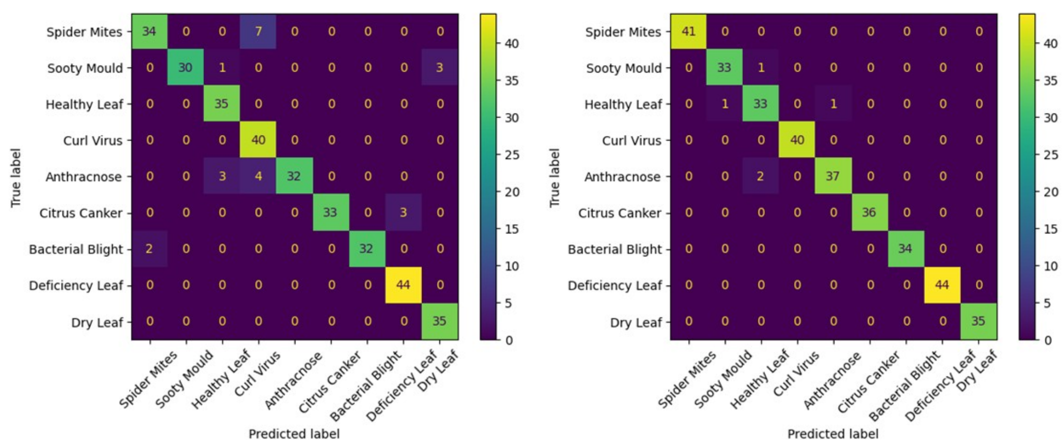

Figure 18. Confusion Matrix for Xception and Ensemble with Proposed approach

Table 7. Comparison of Methods for Citrus and Lemon Disease Detection

Studies	**Methods**	**Dataset**	**Performance**
Butt et.al.[44]	DenseNet, AlexNet	citrus fruits and leaves dataset (2019)	Accuracy = 99.6%
Rahman et.al.[29]	2 stage deep CNN	citrus fruits and leaves dataset (2019)	Accuracy = 94.37%
Dong et.al.[30]	Faster RCNN	in-field leaf dataset	Average Precision = 84.13%
Janarthan et.al.[26]	Deep metric learning-based framework	citrus fruits and leaves dataset (2019)	Accuracy = 95.04%
Mehmood et.al.[46]	Ensemble method	fruit imagery datasets	Accuracy = 99%
Proposed	Enhanced Fuzzy rank based ensemble	Lemon disease dataset (2025)	Accuracy = 96.51%

The fuzzy rank mechanism effectively prioritizes and integrates classification outputs, enhancing decision reliability under uncertainty, which is common in early-stage plant disease symptoms. The proposed approach's competitive performance demonstrates that a fuzzy logic-based ensemble model, when combined with recent, domain-specific datasets, can outperform traditional deep learning models in specific agricultural contexts. This highlights the importance of tailored model architectures and up-to-date datasets in improving precision agriculture applications.This demonstrates its effectiveness and suitability for lemon-specific disease detection, outperforming many existing models on modern datasets and showing promise for future real-time applications.

Conclusion and future work

A novel fuzzy-rank-based ensemble learning approach with enhanced image quality consideration for the classification of citrus leaf diseases using deep learning is presented in this paper. The classification task for the citrus diseases involved categorizing input leaf images into distinct disease classes using proposed ensemble model. The image enhancement stage prior to classification and individual base learners (AlexNet, VGG, Xception) predcition combined through fuzzy ranking mechanism resulted into improvement in accuracy value as compared to conventional approach. A five-fold cross-validation protocol was employed to ensure model generalization and reduce bias. Specifically, the ensemble model achieved a peak accuracy of 96.5%, F1 Score of 96.4%, and recall and precision values of 96.4% and 96.5%, respectively—showing marked improvement over the conventional methodology (accuracy: 91.5%, F1 Score: 92.2%) for the latest publically available dataset. The proposed model performed optimally for the majority of disease categories except Anthracnose where model performance can be improved further.These results demonstrate the effectiveness of the proposed ensemble strategy in enhancing the performance of individual networks by intelligently combining their predictions using a fuzzy ranking mechanism. In future the work can be extended to mjultimodal image datasets with more balanced class wise image distribution to dtecet the diseases at its early onset.

Acknowledgment

The authors would like to acknowledge the support of Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2025R435), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Funding Statement

This research was funded by Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2025R435), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Consent for publication

All authors have read and approved the manuscript

Ethics statement

This study exclusively utilizes publicly accessible Lemon disease dataset. The utilized datasets exclude any private information or identifiable data, thereby negating the requirement for ethical approval. The generative AI and AI-assisted technologies have not been utilized in the writing process.That generative AI and AI-assisted technologies have not been used to create or alter images.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Author contributions

Bobbinpreet Kaur contributed to the conception and design. Shashi Kant Gupta and Midhunchakkaravarthy Janarthan contributed to the collection and assembly of data. Deema Mohammed Alsekait contributed to the development of methodology. Diaa Salama AbdElminaam contributed to the data analysis and interpretation. All authors read and approved the final manuscript.

Data Availability Statement

The dataset used during the current study are available in the Mendeley Data repository under the title "Comprehensive Lemon Leaf Disease Dataset for Advanced Detection and Sustainable Agriculture" (DOI: 10.17632/44nrn4593f.1), https://data.mendeley.com/datasets/44nrn4593f/1

References

1. Khan, M.A., Akram, T., Sharif, M., et al. (2020). "Detection and classification of citrus diseases in agriculture based on optimized weighted segmentation and feature selection." Computers and Electronics in Agriculture, 170, 105354.
2. Mukhametzyanov, Rafail R., Snezhanna V. Brusenko, Akhmed M. Khezhev, Eldar M. Kelemetov, and Sofia S. Kirillova. "Changing the Global Production and Trade of Citrus Fruits." In Sustainable Development of the Agrarian Economy Based on Digital Technologies and Smart Innovations, pp. 19-24. Cham: Springer Nature Switzerland, 2024.
3. Naqvi, S. A. M. H. "Diagnosis and management of pre and post-harvest diseases of citrus fruit." In Diseases of Fruits and Vegetables Volume I: Diagnosis and Management, pp. 339-359. Dordrecht: Springer Netherlands, 2004.
4. Saleem, M.H., Potgieter, J., Arif, K.M. (2020). "Plant disease detection and classification by deep learning." Plants, 9(11), 1302.
5. Kurup, S.M., Amaraja, V.S., Kumar, N.R. (2021). "Ensemble deep learning approach for citrus leaf disease classification." Journal of Plant Diseases and Protection, 128(4), 929-943.
6. Zhang, Y., Song, C., Zhang, D. (2021). "Deep learning-based object detection improvement for tomato disease." IEEE Access, 9, 24248-24257.
7. Osorio, K., Puerto, A., Pedraza, C., et al. (2021). "A multi-core deep learning framework for early diagnosis of citrus greening disease using progressive learning techniques." Applied Soft Computing, 98, 106859.
8. Wagh, P., Sane, S., Kekre, H.B. (2021). "Detection of citrus greening disease using hybrid CNN-LSTM architecture with visual explanations." Expert Systems with Applications, 167, 114346.
9. Tian, K., Li, J., Pan, J., et al. (2022). "Fast citrus disease identification via hierarchical vision Transformer." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 15, 5187-5197.
10. Sharma, P., Hans, R., Grover, J. (2022). "Vision transformer-based deep learning model for citrus disease detection." Neural Computing and Applications, 34, 16933-16948.
11. https://www.fao.org/
12. https://www.ibge.gov.br/en
13. https://www.iari.res.in/en/index.php
14. Csillik, Ovidiu, John Cherbini, Robert Johnson, Andy Lyons, and Maggi Kelly. "Identification of citrus trees from unmanned aerial vehicle imagery using convolutional neural networks." Drones 2, no. 4 (2018): 39.
15. Chane, Camille Simon, Alamin Mansouri, Franck S. Marzani, and Frank Boochs. "Integration of 3D and multispectral data for cultural heritage applications: Survey and perspectives." Image and Vision Computing 31, no. 1 (2013): 91-102.
16. Mohanty, Sharada P., David P. Hughes, and Marcel Salathé. "Using deep learning for image-based plant disease detection." Frontiers in plant science 7 (2016): 215232.
17. Wagle, Shivali Amit. "A Deep Learning-Based Approach in Classification and Validation of Tomato Leaf Disease." Traitement du signal 38, no. 3 (2021).
18. Saha, Rupali, and Shubhangi Neware. "Orange fruit disease classification using deep learning approach." International Journal 9, no. 2 (2020).
19. Sladojevic, Srdjan, Marko Arsenovic, Andras Anderla, Dubravko Culibrk, and Darko Stefanovic. "Deep neural networks based recognition of plant diseases by leaf image classification." Computational intelligence and neuroscience 2016, no. 1 (2016): 3289801.
20. Agarwal, Mohit, Suneet Kr Gupta, and Kanad Kishore Biswas. "Development of Efficient CNN model for Tomato crop disease identification." Sustainable Computing: Informatics and Systems 28 (2020): 100407.
21. Rangarajan, Aravind Krishnaswamy, Raja Purushothaman, and Aniirudh Ramesh. "Tomato crop disease classification using pre-trained deep learning algorithm." Procedia computer science 133 (2018): 1040-1047.
22. Xing, Shuli, Marely Lee, and Keun-kwang Lee. "Citrus pests and diseases recognition model using weakly dense connected convolution network." Sensors 19, no. 14 (2019): 3195.
23. Wang, Zhiqiang, Guofeng Yuan, Huili Pu, Shuangshuang Shan, Zhengke Zhang, Hongmiao Song, and Xiangbin Xu. "1-Methylcyclopropene suppressed the growth of Penicillium digitatum and inhibited the green mould in citrus fruit." Journal of Phytopathology 169, no. 2 (2021): 83-90.

24. Gangwar, Ankita, Vijaypal Singh Dhaka, Geeta Rani, Shrey Khandelwal, Ester Zumpano, and Eugenio Vocaturo. "Time and Space Efficient Multi-Model Convolution Vision Transformer for Tomato Disease Detection from Leaf Images with Varied Backgrounds." Computers, Materials & Continua 79, no. 1 (2024).
25. Barman, Utpal, Parismita Sarma, Mirzanur Rahman, Vaskar Deka, Swati Lahkar, Vaishali Sharma, and Manob Jyoti Saikia. "Vit-SmartAgri: vision transformer and smartphone-based plant disease detection for smart agriculture." Agronomy 14, no. 2 (2024): 327.
26. Janarthan, Sivasubramaniam, Selvarajah Thuseethan, Sutharshan Rajasegarar, Qiang Lyu, Yongqiang Zheng, and John Yearwood. "Deep metric learning based citrus disease classification with sparse data." IEEE Access 8 (2020): 162588-162600.
27. Dhaka, Vijaypal Singh, Nidhi Kundu, Geeta Rani, Ester Zumpano, and Eugenio Vocaturo. "Role of internet of things and deep learning techniques in plant disease detection and classification: A focused review." Sensors 23, no. 18 (2023): 7877.
28. Zhou, Xin, Chunjiang Zhao, Jun Sun, Yan Cao, Kunshan Yao, and Min Xu. "A deep learning method for predicting lead content in oilseed rape leaves using fluorescence hyperspectral imaging." Food chemistry 409 (2023): 135251.
29. Syed-Ab-Rahman, Sharifah Farhana, Mohammad Hesam Hesamian, and Mukesh Prasad. "Citrus disease detection and classification using end-to-end anchor-based deep learning model." Applied Intelligence 52, no. 1 (2022): 927-938.
30. Dong, Ruihao, Aya Shiraiwa, and Takefumi Hayashi. "A Simple Diagnostic Method for Citrus Greening Disease With Deep Learning." Electronics and Communications in Japan 108, no. 1 (2025): e12472.
31. Siam, AKM Fazlul Kobir, Prayma Bishshash, Md Asraful Sharker Nirob, Sajib Bin Mamun, Md Assaduzzaman, and Sheak Rashed Haider Noori. "A comprehensive image dataset for the identification of lemon leaf diseases and computer vision applications." Data in Brief 58 (2025): 111244.
32. Perona, Pietro, and Jitendra Malik. "Scale-space and edge detection using anisotropic diffusion." IEEE Transactions on pattern analysis and machine intelligence 12, no. 7 (1990): 629-639.
33. Whitaker, Ross, and Guido Gerig. "Vector-valued diffusion." In Geometry-driven diffusion in computer vision, pp. 93-134. Dordrecht: Springer Netherlands, 1994.
34. Peter, Pascal, and Joachim Weickert. "Colour image compression with anisotropic diffusion." In 2014 IEEE International Conference on Image Processing (ICIP), pp. 4822-4826. IEEE, 2014.
35. Roussos, Anastasios, and Petros Maragos. "Vector-valued image interpolation by an anisotropic diffusion-projection PDE." In International Conference on Scale Space and Variational Methods in Computer Vision, pp. 104-115. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007.
36. Fernández, José-Jesús, and Sam Li. "An improved algorithm for anisotropic nonlinear diffusion for denoising cryo-tomograms." Journal of structural biology 144, no. 1-2 (2003): 152-161.
37. Bala, A. Anilet, P. Aruna Priya, and Vivek Maik. "Hybrid technique for fundus image enhancement using modified morphological filter and denoising net." The Journal of Supercomputing 80, no. 9 (2024): 13317-13340.
38. Le Van, Toan, and Liet Van Dang. "Enhancement of Mammographic Images Based on Wavelet Denoise and Morphological Contrast Enhancement."
39. Maidana, María Belén Fleitas, José Luis Vázquez Noguera, Diego P. Pinto-Roa, and Julio César Mello-Román. "Noise removal and contrast enhancement in fundus images via morphological operations." In 2022 17th Iberian Conference on Information Systems and Technologies (CISTI), pp. 1-7. IEEE, 2022.
40. Samajdar, Tina, and Md Iqbal Quraishi. "Analysis and evaluation of image quality metrics." In Information Systems Design and Intelligent Applications: Proceedings of Second International Conference INDIA 2015, Volume 2, pp. 369-378. Springer India, 2015.
41. Hore, Alain, and Djemel Ziou. "Image quality metrics: PSNR vs. SSIM." In 2010 20th international conference on pattern recognition, pp. 2366-2369. IEEE, 2010.
42. Mittal, Anish, Rajiv Soundararajan, and Alan C. Bovik. "Making a “completely blind” image quality analyzer." IEEE Signal processing letters 20, no. 3 (2012): 209-212.
43. Udayaraju, Pamula, P. Jeyanthi, and B. V. D. S. Sekhar. "A hybrid multilayered classification model with VGG-19 net for retinal diseases using optical coherence tomography images." Soft Computing 27, no. 17 (2023): 12559-12570.
44. Butt, Nouman, Muhammad Munwar Iqbal, Shabana Ramzan, Ali Raza, Laith Abualigah, Norma Latif Fitriyani, Yeonghyeon Gu, and Muhammad Syafrudin. "Citrus diseases detection using innovative deep learning approach and Hybrid Meta-Heuristic." PloS one 20, no. 1 (2025): e0316081.

45. Deb, Sagar Deep, and Rajib Kumar Jha. "Breast UltraSound Image classification using fuzzy-rank-based ensemble network." Biomedical Signal Processing and Control 85 (2023): 104871.

46. Mehmood, Abid, Muneer Ahmad, and Qazi Mudassar Ilyas. "On precision agriculture: enhanced automated fruit disease identification and classification using a new ensemble classification method." Agriculture 13, no. 2 (2023): 500.