International Journal of Applied Mathematics
Volume 38 No. 11s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

MULTI VEGETABLE LEAF DISEASE CLASSIFICATION AND FERTILIZER RECOMMENDATION USING OPTIMIZED DEEP LEARNING MODEL WITH SHAPLEY ADDITIVE EXPLANATIONS

Ashwini KL Rao[1,] Rajeev Ranjan

Research scholars
School of Computer Science and Applications
REVA University
ashwiniklrao@gmail.com

ORCID iD: 0009-0003-7245-9638

Associate Professor
School of Computer Science and Applications
REVA University
rajeev.ranjan@reva.edu.in.

Abstract

Multi-vegetable leaf disease classification and fertilizer recommendation efforts on identifying and classifying diseases in vegetable leaves using advanced image processing and machine learning techniques. Moreover, several existing approaches faced major challenges such as low accuracy, complexity issues, scalability limitations, and efficiency problems. To overcome the aforementioned issues, this research proposes a very new Dense Scale-invariant feature transform-based Orientation-guided Convolutional Attention network with Quokka Swarm Optimization (DSOCA-QSO) and further adaptation of QSO towards multi-class disease classification. This approach achieves high accuracy with minimal training data, optimizing agricultural productivity through targeted fertilizer recommendations. The process in the pre-processing pipeline includes resizing, normalizing, denoised, and contrast-enhanced by the AEFGC which enhances the quality of images as well as noise-free images. The DSOCA network employs orientation-guided convolutional attention extraction for feature extraction, which allows the model to direct attention to important patterns for diagnosis purposes. In contrast, the QSO model fine-tunes the weight parameters of the model, which it uses to address the problems of computational complexity and sensitivity to parameter values. Additionally, SHAP is employed to enhance interpretability by explaining the model's predictions, ensuring transparency. Fertilizer recommendations are tailored to specific soil conditions, promoting plant health. The DSOCA-QSO approach achieves accuracy (99.9%), F1-score (99.35%), precision (99.7%), recall (99.99%), and sensitivity (98.99%). These results highlight its effectiveness in improving agricultural practices through precise disease diagnosis and optimized fertilizer recommendations.

Keywords: *Leaf Disease Classification, Quokka Swarm Optimization, Invariant Feature Transform, Deep Learning, Shapley Additive explanations*

1. Introduction

Smart farming, which combines the Internet of Things (IoT) with modern agricultural practices, represents a groundbreaking shift in traditional farming methods. By utilizing real-time data, farmers make more informed decisions, leading to improved efficiency and productivity in their operations. Smart farming powered by IoT technology plays a highly significant role in meeting increased global food production

requirements through precision agriculture , optimizing inputs such as water, fertilizers, and pesticides in light of reliable and pertinent information. [1, 2].

IoT has profound impacts on crop leaf disease classification as well as fertilizer recommendation - two critical components of precision agriculture. Traditional approaches have involved crop leaf disease observation in the field or basic algorithms fail to capture the complexity as well as the variability of agricultural environments. Similarly, the unique requirements of various crops, soil types, and environmental circumstances are not taken into consideration by standard fertilizer recommendations. These limitations lead to suboptimal crop health and productivity, affecting the sustainability of farming practices [3-6]. The application of advanced deep learning models over IoT-based systems for crop leaf disease classification and fertilizer recommendation addresses such issues. Many models fail to handle the high variability in data generated by IoT devices or require extensive computational resources, which is considered a significant drawback in resource-constrained environments. [7-10].

In summary, the objectives outlined focus on advancing IoT-based smart farming by addressing the limitations of current methods in job scheduling and agricultural decision-making. The invention of an ideal neural network-based task in conjunction with a complex deep learning model for the classification of multi-crop leaf diseases. Moreover, fertilizer recommendation has the potential to greatly improve the efficiency, accuracy, and sustainability of precision agriculture. These innovations improve resource utilization in cloud-based IoT systems and contribute to higher agricultural productivity and better management of natural resources.

In the field of multi-vegetable leaf disease classification, **Joseph DS et al. [11]** developed the Maize, Rice, Wheat (MRW-CNN) model in 2024 for detecting plant disease. By precisely recognizing the unique traits of both healthy and diseased leaves, the suggested MRW-CNN model was intended to diagnose diseases in the leaves of maize, rice, and wheat. The Convolutional Neural Network (CNN) was proposed by **Hassan SM et al. [12]** in 2022 for identifying plant diseases. For a variety of computer vision applications, such as pattern recognition and classification, CNN, a sort of neural network, proved especially successful. In 2023, **Ahmad A et al. [13]** presented VGG16 for VGG16 an older Deep Neural Network (DNN) architecture with fewer layers. It has significantly more trainable parameters, which increases the network's complexity, resulting in longer training times and higher memory requirements. The AlexNet was proposed by **Yang L et al. [14]** in 2022 for identifying plant diseases. The original AlexNet model consisted of five convolutional layers. Such changes

allowed recognition accuracy to remain unchanged while at the same time minimizing computational complexity. In 2023, **Guan H et al. [15]** developed a Dise-Efficient (DE) model for plant disease identification. The Dise-Efficient model helps to prevent them in agricultural production. Conducted A-series experiments to assess, and deploy factors to investigate the impact of transfer learning for training the model. Table 1 explains the description of various existing methods shown below,

International Journal of Applied Mathematics
Volume 38 No. 11s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Table 1: Description of various existing methods.

Ref.no	**Techniques used**	**Highlights**	**Limitations**	**Performance metrics**
[11]	MRW-CNN	The dataset includes real-life images with varying disease severities.	The presence of pre-processed images did not fully account for real-life environmental factors.	Accuracy-0.9728
[12]	CNN	The use of depth-wise reduces training time and enhances the effectiveness of plant disease classification.	The model was limited by dataset imbalance.	Accuray-99.39%
[13]	DNN	The research work achieved high generalization accuracy.	The study was limited by environmental variability and dependence on specific datasets.	Accuracy-81.60%
[14]	AlexNet	Timely identification of crop diseases' potential for deployment on portable devices.	The model was limited by disease similarity and environmental variability.	Accuracy-99.69%
[15]	DE	The model achieved high accuracy.	The model was limited to accuracy reduction with fewer layers.	Accuracy -64.40%
[16]	3D-DCNN-EOS	The proposed model achieved high accuracy and ensured efficient classification.	The proposed model encounters challenges including a limited dataset.	F1-score-97.8%, accuracy-98%, recall-97.9%,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

				precision- 97.8%,
[17]	PI-NMD	This research well with clean datasets.	The proposed model exhibits lower accuracy on noisy datasets.	F1-score- 76.52
[18]	3D MOT	Improved accuracy, robustness, and consistency in yield estimation.	Increased computational complexity and higher resource requirements.	-
[19]	ANN	Achieved better performance with free, accessible data.	Limited by sensor resolution and environmental variability	Accuracy - 99%
[20]	VGG-CNN	Delivered consistent performance across diverse datasets.	Optimization for real-time deployment was relatively larger than other lightweight models.	Accuracy- 99.16%

The 3D-Dense Convolutional Neural Network -Ebola Optimization Search (3D-DCNN-EOS) model was presented by **Ashwini C et al. [16]** in 2023 for detecting illness in corn leaf illness. The EOS algorithm enhances the 3D-CNN framework by optimizing its weights, thereby improving classification performance efficiently. The Plant Image-Based Nonlinear Motion Deblurring Method (PI-NMD) was developed by **Batchuluun G et al. [17]** in the year 2023 to classify plant images. Those deblurred images serve as input for the PI-Clas method, arguably the best-performing method against ones that do not consider the PI-NMD. However, remains fairly strong and persistent yet it does fade in time. In 2024, **Ahmedt-Aristizabal D et al. [18]** presented 3D Multi-Object Tracking (3D-MOT) approach for estimating Vineyard Yield. The videos from the three vineyard panels used to validate the bunch counting model were also utilized to assess yield estimation with the proposed system. In 2023, **Bolfe ÉL et al. [19]** presented Artificial Neural Networks (ANNs) for Analyzing Agricultural Intensification Patterns. This research work was the first to map agricultural intensification in Brazil's savanna regions using spectral indices and a variety of machine learning models with HLS time series data. The Visual Geometry Group-inception Convolutional Neural Network (VGG-CNN) was presented in the year 2023 by **Thakur PS et al. [20]**. For a range of plant species and illnesses associated with ITS, the primary goal was to minimize the model's size while preserving

classification effectiveness. The architecture of the model incorporates aspects of both Inception modules and VGG blocks.

1.1 Motivation

Precision agriculture still faces challenges, especially in the detection and control of vegetable leaf diseases

Volume 38 No. 11s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

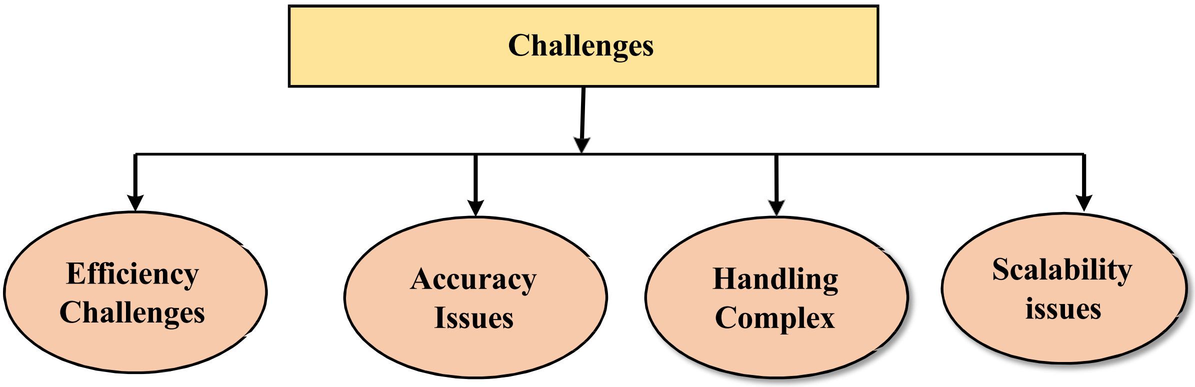
or insect pests, because the available conventional approaches become limited and suffer from a lack of real-time integration and information connectivity. Although significant research has focused on vegetable crops, they receive less attention despite their vital role in the global economy and nutrition. The vulnerability of vegetables to diseases, which impact both yield and quality, necessitates precise detection methods.

1.2 Problem statement

The classification and management of vegetables struggle with accuracy, efficiency, and scalability. Traditional image processing techniques demand substantial computational resources and produce errors when handling complex, high-dimensional data. This research introduces a novel solution, the DSOCA-QSO. Figure 1 shows the Considered existing challenges.

Figure 1: Considered existing challenges

The advanced combination of the latest deep learning techniques with optimization algorithms used will



boost the potential of disease classification accuracy, resource efficiency, and model generalization.

1.3 Contributions

➢ This research introduces the innovative DSOCA-QSO hybrid model to enhance multi-vegetable disease classification. DSOCA leverages orientation-guided convolutional attention to extract essential leaf image features, emphasizing key patterns for higher accuracy in disease identification.

➢ This research focuses on the preprocessing steps, involving resizing, normalization, denoising, and even contrast enhancement by using the Adaptive contrast Enhancement algorithm with Filtered Gamma Correction (AEFGC).

➢ This research employs SHAP to improve model interpretability by explaining predictions and ensuring transparency in multi-vegetable disease classification.

➢ The QSO algorithm fine-tunes the DSOCA network's weight parameters, overcoming computational complexity and sensitivity issues, and thereby enhancing optimization for complex problems. Furthermore, the system also advises on fertilizer utilization as a result of the diagnosis of diseases improving plant health and productivity.

The research is structured as follows: Section 2 offers a concise overview of the newly developed technique. Section 3 presents the results, highlighting the improved performance of the proposed methods. Section 4 discusses the findings and their implications, while Section 5 concludes the research.

Received: August 05, 2025

Volume 38 No. 11s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

2. Proposed methodology

The proposed method develops a multi-vegetable leaf disease classification method that effectively addresses challenges associated with data limitations and imbalanced class distributions. Figure 1 depicts a system that classifies plant diseases and provides fertilizer recommendations. An AEFGC that contains steps like resizing, denoising, normalization, and contrast enhancement. The images are then enhanced and analyzed using DSOCA for feature extraction and classification. Figure 2 shows the Structure of the proposed DSOCA-QSO.

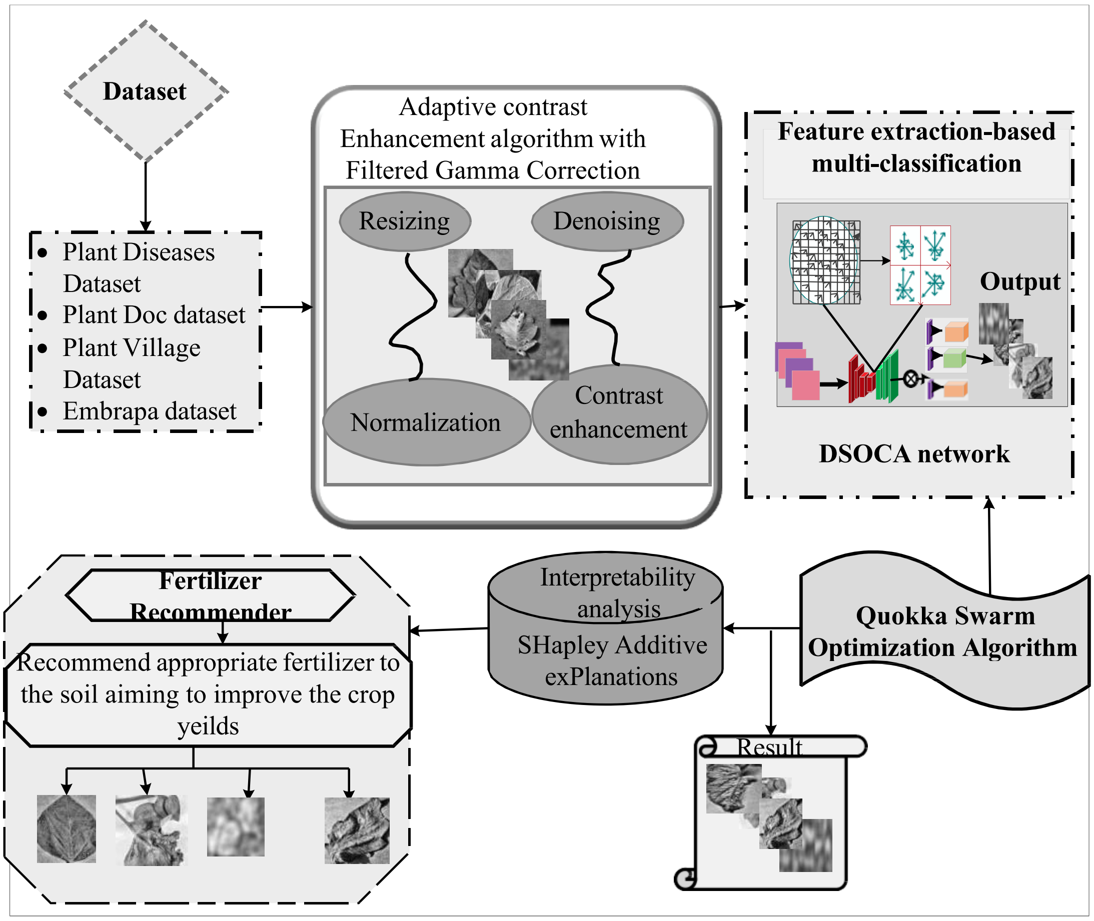


Figure 2: Structure of proposed DSOCA-QSO

To ensure interpretability, SHAP is applied. Finally, a fertilizer recommendation system is used to suggest fertilizers aimed at improving crop yields.

2.1 Data acquisition

The input images are gathered from four different datasets as the New Plant Diseases Dataset (Dataset 1) [27], the Plant Doc dataset (Dataset 2) [28], the Embrapa dataset (Dataset 3) [30],

and the Plant Village Dataset (Dataset 4) [29]. Dataset 1 has 37,286 training and 9,320 testing samples over 20 classes, including tomato and corn diseases. One specific example is tomato yellow leaf curl virus (1,961 train, 490 test). Dataset 2 is the smallest, with 1,381 training and 132 testing samples over 16 classes. Some examples include tomato septoria leaf spot (134 train, 10 test) and Corn leaf blight (177 train, 12 test). All these datasets provide a rich resource to train and evaluate plant disease classification models across multiple crops and diseases. Dataset 3 contained 17,195 training and 5,365 testing samples

International Journal of Applied Mathematics
Volume 38 No. 11s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

among 31 classes of plant diseases. The biggest class is Milho (Corn) - Mancha_Bipolaris (2,412 train, 754 test). Finally, the dataset 4 includes 16,502 training and 4,134 testing samples over 15 classes. Again, the most common is Tomato Yellow Leaf Curl Virus (2,566 train, 642 test). This dataset split into 80/20 ratio and the image size of this dataset is 256x256. These datasets are used to create and enhance models for detecting and classifying plant diseases accurately.

2.2 Image preprocessing

The input images are resized, normalized, and denoised, and their contrast is enhanced using the AEFGC to improve quality and prepare them for analysis. ACEA is very much supported to solve potential low-contrast problems. ACEA improves image contrast by redistributing pixel values according to the unique characteristics of each image's histogram. A straightforward yet powerful technique for enhancing visual contrast is ACEA [20]. The image's histogram is first determined, and then the intensity statistics of nearby pixels are used to equalize the final product. The images are resized to 750 × 1000, and the average runtime for processing a single image is 4.05 milliseconds. In particular, each pixel in the equation's normalized histogram is computed in equation (1).

$$R_{a,b}(j)=\frac{m_{a,b}(j)}{m}, \quad 0 \le j \le 255 \qquad (1)$$

where, $m_{a,b}(j)$ denotes the frequency of the pixel, m signifies overall pixels, j represents the pixels, a, b denotes the pixels, $R_{a,b}(j)$ represents the normalized histogram range and $0 \le j \le 255$ denotes pixel intensity range [21]. The associated cumulative distribution function is expressed in equation (2).

$$d_{a,b}(j)=\sum_{i=0}^{j} R_{a,b}(j) \qquad (2)$$

where, $d_{a,b}(j)$ denotes the cumulative distribution function, i denotes the pixel's intensity.

Based on this, equalization is achieved using the following transformation, which is given in equation (3).

$$c = 255\times d_{a,b}(j) \qquad (3)$$

International Journal of Applied Mathematics
Volume 38 No. 11s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

The output $g_{a,b}(j)$ is generated by applying (3) to map each pixel in the input with intensity

to a corresponding pixel with intensity. Therefore, CLAHE is utilized to constrain contrast amplification in smooth regions, thereby reducing noise amplification. A multi-scale sharpening technique that employs directed filtering, which improves texture and edge details. The capacity of directed filtering to increase low-exposure regions is limited, despite its usefulness for local contrast augmentation. To address these challenges, a gamma-corrected

version of the color-corrected image is developed as part of the fusion input. The standard gamma correction is defined by the following equation (4).

$$H = 255\left[\frac{J(a,b)}{255}\right]^{\lambda} \quad (4)$$

H represents the pixel output value, $J(a,b)$ denotes the pixel value, 255 denotes the maximum intensity value and λ signifies a parameter in gamma correction. Many image processing jobs frequently use gamma correction. This adjustment is effective for correcting images that are either overexposed, underexposed, or affected by both issues, and is tailored to the specific characteristics of each image to achieve the best results. After the image processing, the feature extraction-based multi-classification is used to extract the features like entropy, contrast, variance, homogeneity, and correlation.

2.3 Feature extraction-based multi-classification

In this phase, five texture features, entropy, contrast, variance, homogeneity, and correlation, are extracted. Entropy quantifies the complicated or unpredictable the texture of an image is, with higher values indicating greater variability or disorder, Contrast represents the difference in intensity between the darkest and brightest regions of an image, with higher values highlighting more distinct edges or transitions in texture, variance reflects the spread or dispersion of pixel intensity values. Table 2 shows the features and equation and Figure 3 shows the network structure of the DSOCA Technique.

Table 2: Features and Equation

Features	Equation
Entropy	$ENT = -\sum_{R=0}^{j-1}\sum_{P=0}^{i-1} f(R,P)\log_2 g(R,P)$
Contrast	$CONT = \sum_{y=0}^{n-1}\sum_{x=0}^{m-1}(y-x)^2 g(y,x)$

International Journal of Applied Mathematics

Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Variance	$\text{var}\,iance = \sum_{j}\sum_{j}(j-\eta)^2 R(j,i)$
Homogeneity	$HOM = \sum_{R=0}^{j-1}\sum_{P=0}^{I-1}\frac{1}{(R-P)^2} g(R,P)$
Correlation	$COR = \sum_{R=0}^{j-1}\sum_{p=0}^{i-1}(R,P)f(R,P) - N_{RNP}$

where, ENT mean the entropy, $f(R,P)$ represents the co-occurrence matrix range, R,P denotes the position of pixel intensity, $\log_2 g(R,P)$ denotes the logarithmic term, $CONT$ denotes the contrast, $g(y,x)$ denotes the matrix co-occurrence value, x,y denotes the co-occurrence matrix, $\text{var}\,iance$ denotes the variance, $R(j,i)$ represents the normalized co-occurrence matrix, η is the mean gray-level value, $(j-\eta)^2$ represents the squared deviation, HOM signifies the homogeneity, $\frac{1}{(R-P)^2}$ signifies the weighting factor, N_{RNP} denotes the mean intensity value, i, j denotes the pixel intensity ranges, η represents the average intensity value of the image, COR signifies the correlation, R,P represents the intensity pixel range, N_{RN} denotes the mean average value. The classification process for multi-vegetable leaf diseases is conducted through the implementation of an advanced framework known as the DSOCA network [22]. The SIFT algorithm detects and describes local features in images by identifying key points at extreme values in the spatial scale and extracting their position, scale, and rotational invariance.

International Journal of Applied Mathematics
Volume 38 No. 11s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

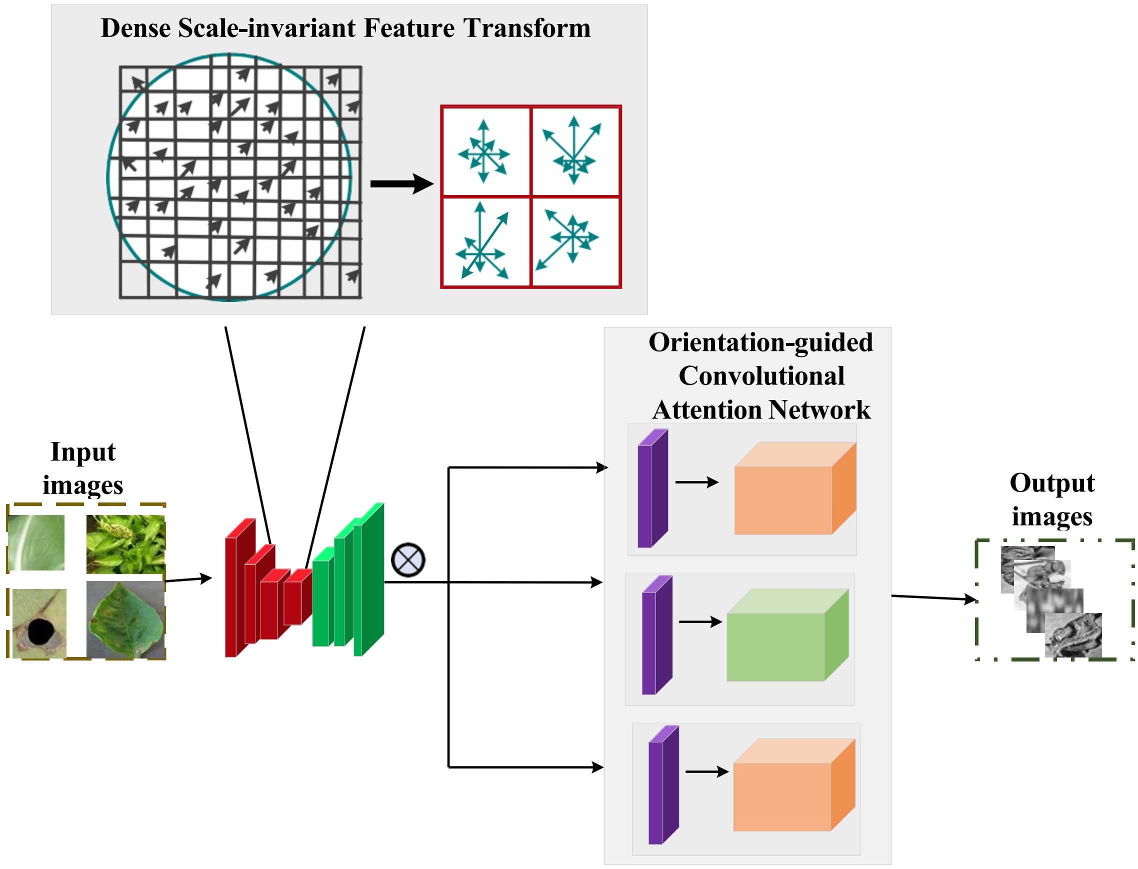


Figure 3: Network structure of DSOCA Technique

The scale-space theory in the SIFT algorithm [24] aims to represent the multi-scale features of images. The scale-space $H(a,b,\beta)$ An image at various scales is obtained by convolving the two-dimensional image with a Gaussian kernel function given in equation (5).

Received: August 05, 2025

International Journal of Applied Mathematics
Volume 38 No. 11s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

$$H(a,b,\beta)=\frac{1}{2\pi\omega^2}e^{\frac{-a^2+b^2}{2\omega^2}} \quad (5)$$

where, $H(a,b,\beta)$ denotes the gaussian kernel function, $2\pi\omega^2$ denotes the normalization factor and $e^{\frac{-a^2+b^2}{2\omega^2}}$ represents the exponential component. Moreover, a DSOCA is developed to emphasize regions or features [23] by leveraging learned patterns and spatial hierarchies shown in equation (6).

$$AC(a)=\lambda(MLP(Avgpool(y))+MLP(Maxpool(x)) \quad (6)$$

where, AC denotes the channel attention, a denotes the input feature map, $Avgpool$ represents the average pooling, $Maxpool$ denotes the maximum pooling and MLP represents the multilayer perceptron, which is shown in equation (7).

$$AS(a)=\lambda(h^{7\times7}([H^g_{avg};H^g_{\max}])) \quad (7)$$

where $AS(a)$ represents the Spatial attention maps, λ denotes the sigmoid function, $h^{7\times7}$ represents the 7×7 convolutional procedure, $H^g_{\max}$ denotes the Feature maps obtained from the max pooling operation and H^g_{avg} signifies the Feature maps obtained from the average

International Journal of Applied Mathematics
Volume 38 No. 11s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)
pooling operation. DSOCA network leverages the integration of features with different attributes of the same image in the same orientation, addressing the issue of ambiguous and weak attribute relationships. For a given orientation H, its graph is constructed by representing the features of D attributes as vertices given in equation (8).

$$h = (U, B) = (\hat{E}_j, M_h + M_C) \tag{8}$$

where, M_h represents the weight sources and M_C denotes the destination weights, U denotes the nodes and B represents the edges, h denotes the graph and $\hat{E}_j$ denotes the vertices. The residual structure is incorporated to avoid information loss given in equation (9).

$$\lambda_i^m = MLP((1+\in).\lambda_i^{m-1} + \tau_i) \tag{9}$$

where, λ_i^m denotes the feature value of the node i at layer, λ_i^{m-1} represents the embedding value of the node i at layer, $\in$ represents the learnable parameter. Additionally, the influence of attribute relations is amplified within each orientation, mitigating their inherent weakness. Thus, DSOCA is utilized to improve multi-vegetable leaf disease classification by employing advanced mechanisms. Subsequently, QSO is implemented for tuning the weight parameters in equation (8).

2.4 Quokka Swarm Optimization Algorithm

The QSO provides adaptability, rapid convergence, effective exploration, efficient exploitation, simplicity, scalability, and reliable performance across various optimization challenges. The QSO algorithm is employed because of its ability to effectively reduce loss function Furthermore, the QSO algorithm is developed to enhance the performance of DSOCA, especially in the application of multi vegetable leaf disease classification.

- **Initialization**

The ratio of the swarm's worldwide worst position to its global best position is known as the inertia parameter. Therefore, the following equation (10) defines the inertia weight parameter in this work

International Journal of Applied Mathematics
Volume 38 No. 11s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

$$h \quad \lambda_{Hworst} \tag{10}$$

$$m = m_{\max} - (m_{\max} - m_{\min}) \left(\frac{\lambda}{\lambda_{Hbest}} \right)$$

where, $m_{\max}$ and $m_{\min}$ denotes the maximum and minimum weight values, λ_{Hworst} and λ_{Hbest} denotes the worst and best performance in iteration and m^h represents the inertia weight defined in equation (8).

➢ **Fitness Function**

The Fitness Function (FF) evaluated in equation (11), which is used to tune the weight parameters derived from the DSOCA-QSO, ensuring optimal performance.

$$F(H)= \{tuning(M_h ,M_c)\} \tag{11}$$

where $F(H)$ denotes the FF of the problem. Moreover, FF improves efficiency by accelerating convergence to optimal parameter settings and prioritizing configurations with higher fitness scores to minimize weight parameters. Table 3 shows the pseudocode of the QSO algorithm.

Table 3: Pseudocode of QSO algorithm

Algorithm QSO pseudocode
Output: λ_{Hbest} - the better performance founded 1: Start 2: Initialize the swarm randomly 3: Initialize λ_{Hbest} with a minimum fitness value 4: Iterate through the swarm 5: $m^h = \lambda_{Hbest}$ (update the better particle position)

International Journal of Applied Mathematics
Volume 38 No. 11s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

```
        end if
    end for
  update the swarms to better performance
end while
End
```

The QSO algorithm offers adaptability, rapid convergence, and efficient exploration and exploitation capabilities. This approach effectively minimizes loss functions and improves DSO performance, especially in multi-vegetable leaf disease classification. The inertia weight parameter, described in equation (10), regulates swarm dynamics, while the Fitness Function (FF) in equation (11) drives the optimization process. By prioritizing high-fitness configurations, FF enhances tuning efficiency, accelerates convergence, and ensures optimal model performance through targeted loss minimization and metric improvement. After the multi-vegetable leaf disease classification, the explainability approach SHAP is used for enhancing the interpretability of each class.

2.5 Interpretability analysis:

SHAP method enhances transparency and interpretability in explaining the DSOCA's model classifications. SHAP [25] evaluates the relevance of model features. Every feasible combination of characteristics linked to leaf disease is taken into account while calculating the Shapley value. Red pixels represent a higher chance of accurately identifying a class, whereas blue pixels represent a lesser chance. Shapley values are represented as pixels. The SHAP value

is estimated using the equation (12).

$$\lambda_h = \sum_{N \subseteq M} \frac{|N|!(B-|N|-1)!}{B!}\left[g_y(N \cup H) - g_y(N)\right] \tag{12}$$

where, λ_h denotes the shapley value, N represents the feature subset, g_y denotes the marginal contribution, $\frac{|N|!(B-|N|-1)!}{B!}$ denotes the weighting features and $g_y(N \cup H)$ signifies model outcome. The estimated outcome represented by $g_y(N)$ is given in equation (13).

$$g_y(N) = R[g(y)\,y\,] \tag{13}$$

Received: August 05, 2025

Volume 38 No. 11s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

where, R denotes the model output. While SHAP enhances interpretability, it is computationally intensive for complex models, sensitive to feature correlations, and faces challenges in scaling with high-dimensional data. Once the explainability using the SHAP

approach has been completed, the fertilizer recommendation is designed to prompt the leaf disease detection.

2.6 Fertilizers recommendation

The proposed method facilitates the prompt detection of leaf diseases, thereby enabling timely intervention [26]. Additionally, it provides recommendations for appropriate fertilizers tailored to the soil, aiming to enhance the yield of various vegetables. Protecting crops from pests and applying the proper fertilizers to the soil are two essential components of farming that improve production. Fertilizers are vital for improving soil fertility and increasing crop productivity. Selecting the appropriate fertilizer depends on the type of crop, soil conditions, and specific nutrient deficiencies. Conducting a soil test is essential to determine nutrient levels and identify shortages in Nitrogen (N), Phosphorus (P), and Potassium (K), the primary macronutrients. For soils lacking nitrogen, urea or ammonium nitrate is effective. Phosphorus deficiencies are addressed with Di-Ammonium Phosphate (DAP) or rock phosphate, while muriate of potash (MOP) or Sulfate Of Potash (SOP) is suitable for potassium-deficient soils. Organic alternatives like compost and biofertilizers enhance soil health while reducing reliance on synthetic fertilizers. Integrated nutrient management, which combines organic and inorganic fertilizers, is a proven approach. Applying fertilizers at the correct growth stages and using location-specific technologies ensures efficient use, minimizes over-fertilization, and safeguards the environment. Sustainable practices are crucial for healthy crops and long-term soil fertility.

3. Simulation experiments

The introduced approach is executed using Python. The efficacy of the suggested approach is assessed using a thorough evaluation that incorporates several performance metrics. These measures include sensitivity, kurtosis, precision, variance, recall, computational complexity, F-measure, specificity, error rate, ROC analysis, entropy, mean, and accuracy.

Table 4: Parameter settings

Parameters	**Description**
Algorithm	Quokka Swarm Optimization Algorithm
Batch size	32
Iterations	100
Epochs	100
Dropout size	0.2, 0.5
Optimizer	QSO

International Journal of Applied Mathematics
Volume 38 No. 11s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Loss	Cross-Entropy, Margin Loss

Table 4 provides an overview of the parameters that are used to train the model. The algorithm used was the QSO algorithm. The batch size used is 32, allowing the model to work with the data in small sets efficiently. The model learning was undertaken for 100 iterations and 100 epochs to allow for adequate learning opportunities. Dropout values of 0.2 and 0.5 were used during training to mitigate the overfitting issue described above by turning off neurons during training at random. The QSO algorithm was used as an optimizer, while updating the model weights was accomplished with flexible weights. The loss functions are cross-entropy loss and margin loss, which help with prediction errors and classification performance.

3.1 Attained outcomes and analysis

A thorough presentation and analysis of the results from four distinct datasets on plant disease recognition are given in this section. It highlights the technique's efficiency, addresses performance issues, and offers key insights into its advantages.

Table 5: Training and testing values of the overall dataset

Datasets	Total samples	Training samples	Testing samples	Total classes
Dataset 1	46,606	37,286	9,320	20
Dataset 2	1,513	1,381	132	16
Dataset 3	22,560	17,195	5,365	31
Dataset 4	20,636	16,502	4,134	15

Table 5illustrates the training/testing distributions for each of the four datasets used for plant disease classification. Dataset 1 has the highest number of samples, requiring training and testing with 46,606 samples (37,286 samples were for training and 9,320 for testing), from 20 different classes. Dataset 2 had the fewest number of samples, with a total of 1,513 samples (1,381 samples were for training and 132 for testing), from 16 different classes. Dataset 3 contains 22,560 samples with 17,195 samples for training and 5,365 samples for testing, from 31 different classes. Dataset 4 had a total of 20,636 samples, with training samples being 16,502 and testing samples being 4,134, from 15 different classes. The plant disease datasets consist of varying amounts of classes and samples in each class allow for a thorough evaluation of the performance of each model of plant disease detection and classification.

3.1.1 Pre-processed images across all datasets

Figure 4 showcases a comparison between the original and pre-processed images across five different datasets (1 to 4).

International Journal of Applied Mathematics
Volume 38 No. 11s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

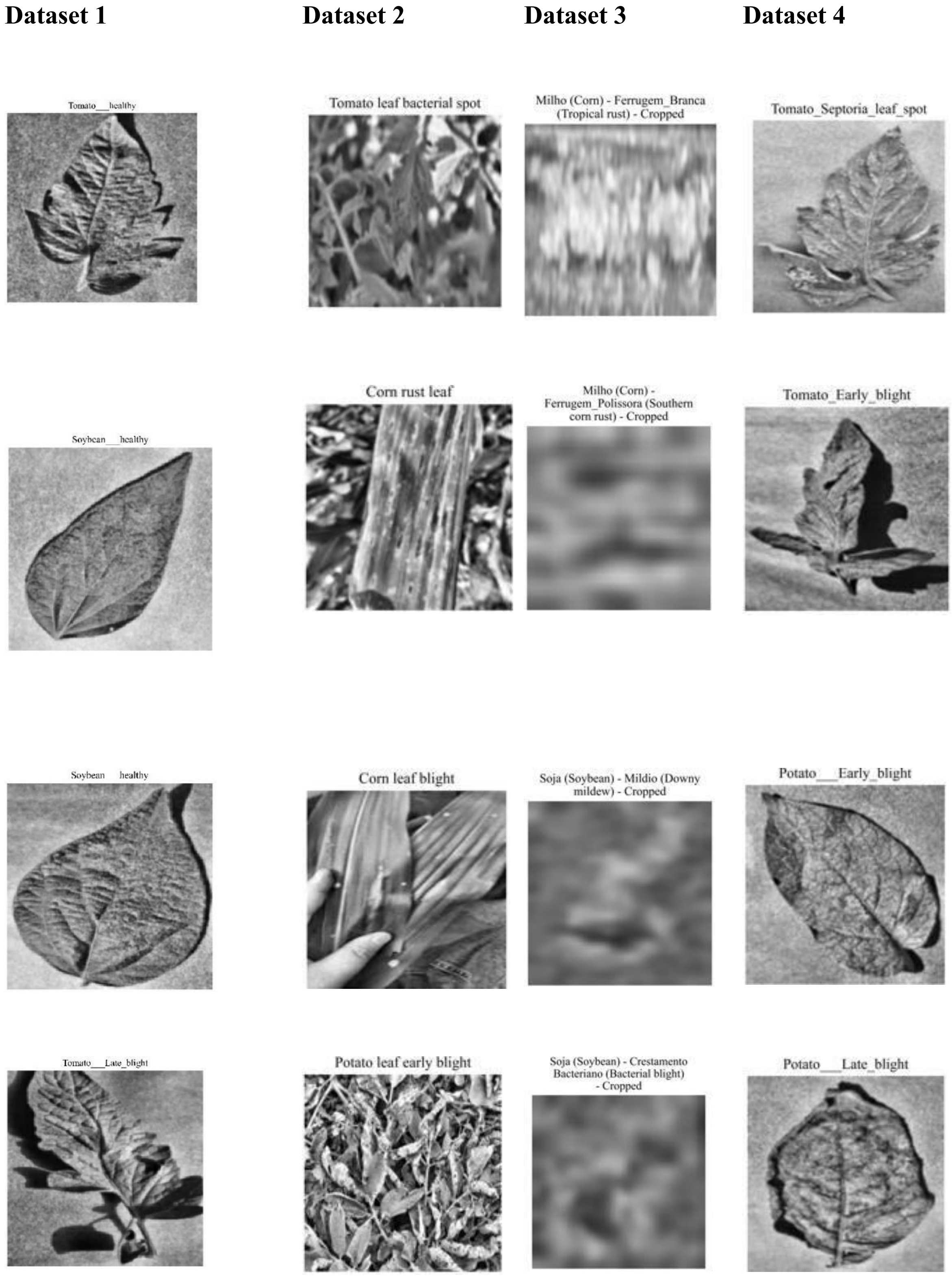


Figure 4: Pre-processed images in four datasets

Each column corresponds to a dataset (1 to 4) featuring diseases such as bacterial spot, rust, blight, and mildew in crops like tomatoes, corn, soybeans, and potatoes. The images differ in clarity, perspective, and annotation, highlighting dataset diversity for plant disease detection applications.

3.1.2 SHAP analysis of supervised image models across all datasets

International Journal of Applied Mathematics
Volume 38 No. 11s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Figure 5 presents the SHAP analysis of supervised image models applied to four plant disease datasets: Dataset 1, Dataset 2, Dataset 3 and Dataset 4.

Datasets	Original images	Outcome of SHAP analysis	Classification
Dataset 1			Corn
Dataset 2			Pepper leaf spot
Dataset 3			Pepper ball bacterial
Dataset 4			Cabbage

Figure 5: SHAP analysis of supervised image models across all datasets

Received: August 05, 2025

International Journal of Applied Mathematics
Volume 38 No. 11s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

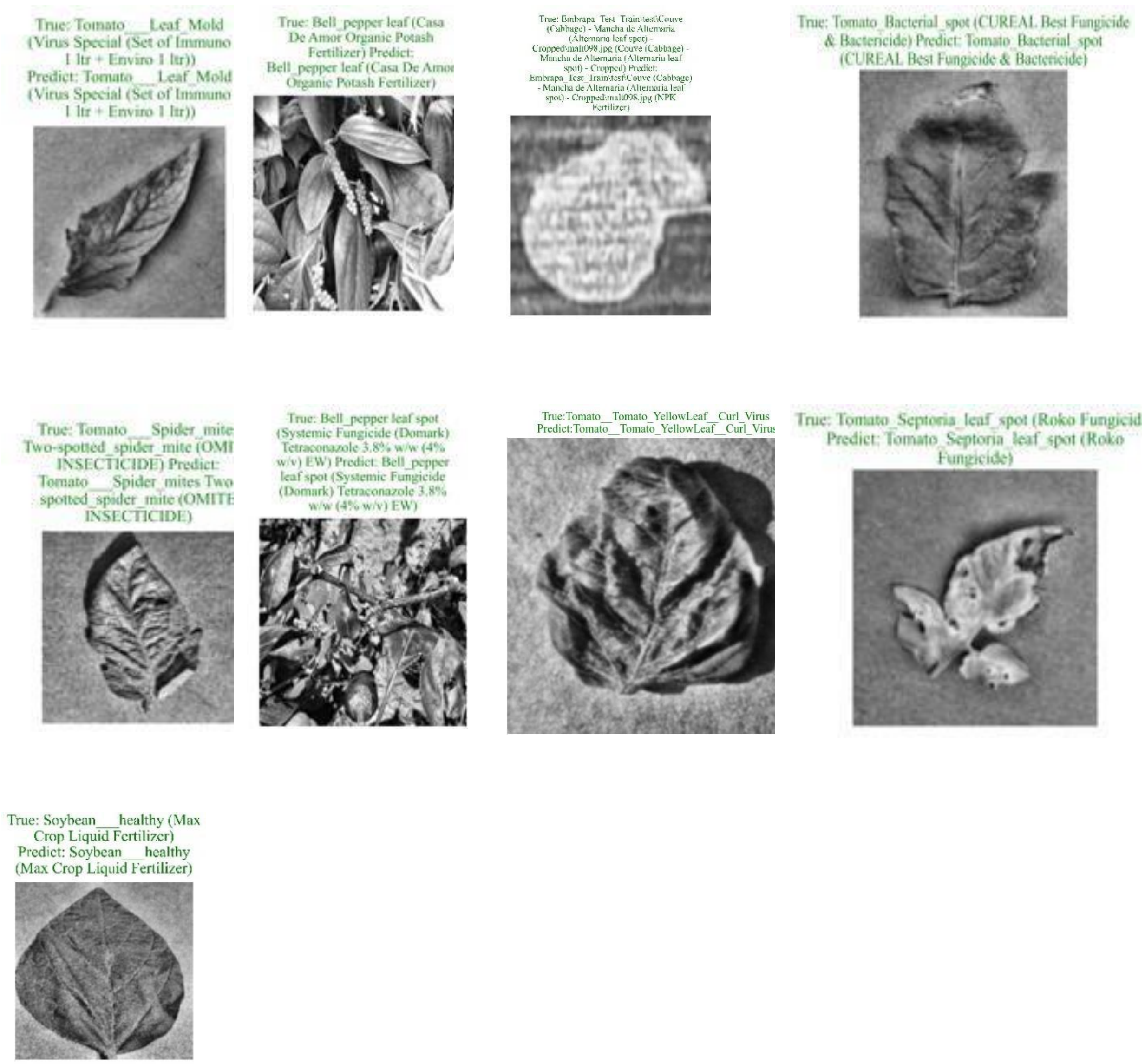
The SHAP analysis results from supervised image classification models across four datasets, all focused on different plant diseases. Dataset 1 focused on corn disease classification. Dataset 2 focused on pepper leaf spot. Dataset 3 focused on pepper bacterial spot. Dataset 4 looked at diseases related to cabbage. The SHAP analysis demonstrates the interpretability of the models by showing which image features contributed to classification, indicating each model differentiated between the disease-specific patterns within the four datasets. The visualizations

emphasize key regions (red/pink areas) that impact model decisions for detecting plant diseases.

3.1.3 Recommended fertilizers

Figure 6 presents an analysis of recommended fertilizers linked to specific plant diseases in 4 datasets. This image presents a grid illustrating various tomato plant diseases and pest issues alongside their corresponding treatments or preventive measures.

Dataset 1 **Dataset 2** **Dataset 3** **Dataset 4**



Received: August 05, 2025

International Journal of Applied Mathematics
Volume 38 No. 11s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

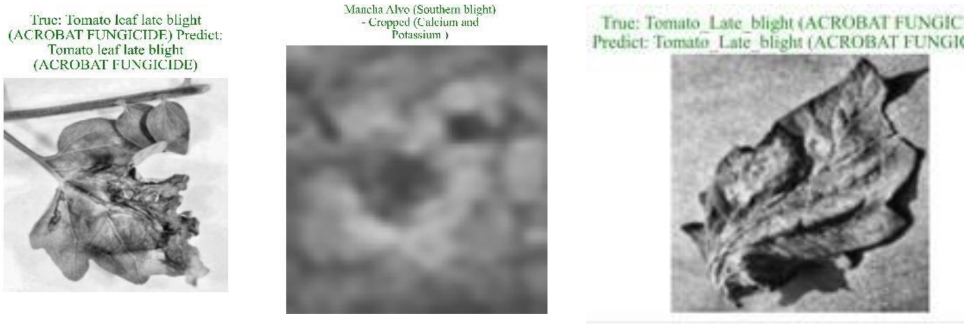


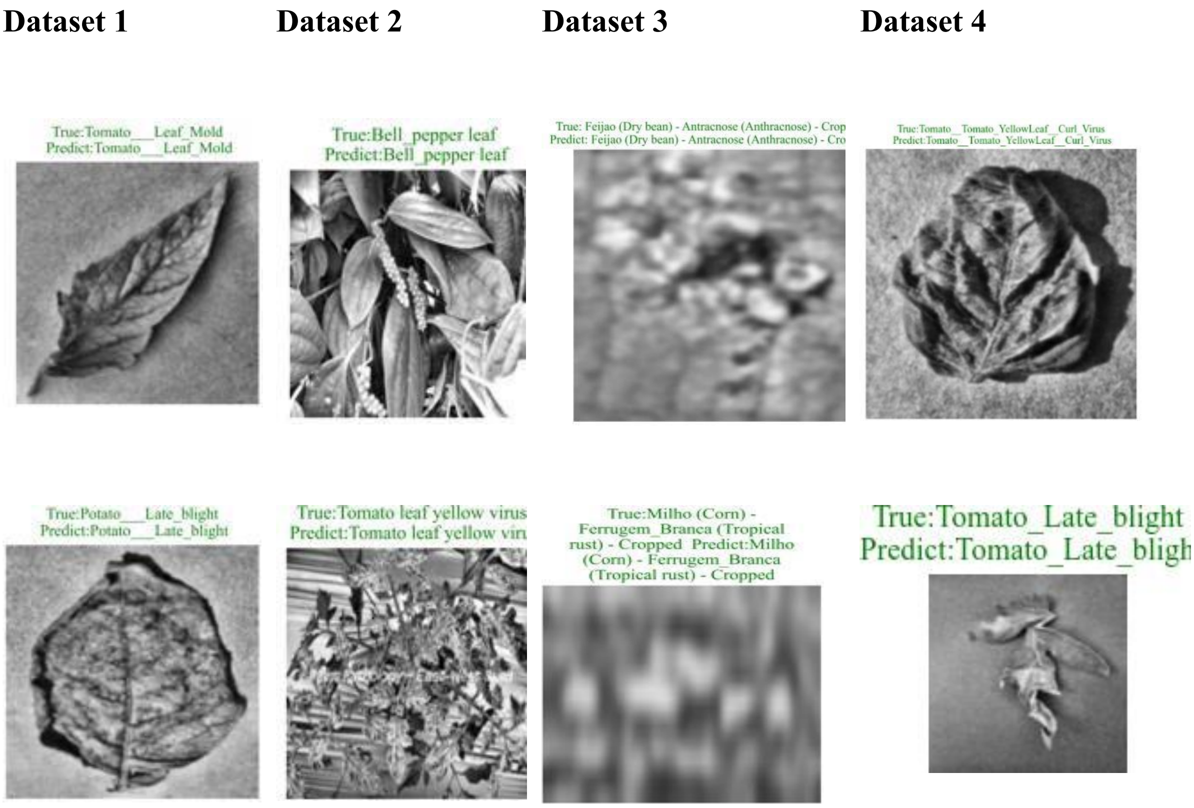
Figure 6: Analysis of recommended fertilizers across all datasets

Each cell specifies the plant type, the problem (e.g., fungal infections, pests), and recommended solutions such as fungicides or insecticides, serving as a practical guide for identification and management.

3.2 Predicted outcomes

Figure 7 presents the predicted outcomes of four datasets displaying images of plant leaves affected by various diseases and healthy conditions across crops like tomato, potato, corn, and soybean.

Dataset 1 **Dataset 2** **Dataset 3** **Dataset 4**



Received: August 05, 2025

International Journal of Applied Mathematics
Volume 38 No. 11s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

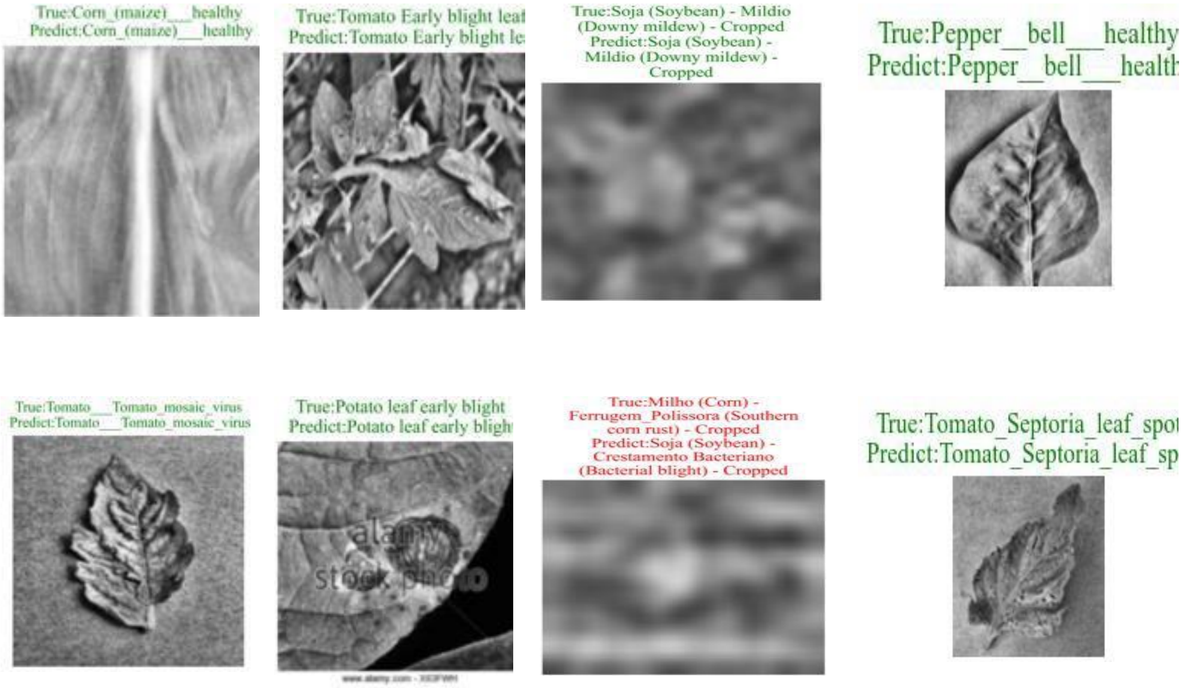


Figure 7: Predicted outcomes across four datasets

Each column represents a dataset, highlighting crop issues such as fungal infections, viruses, and blight. This visualizes model predictions versus true labels, facilitating the assessment of classification accuracy in plant health diagnostics.

3.3 Performance analysis

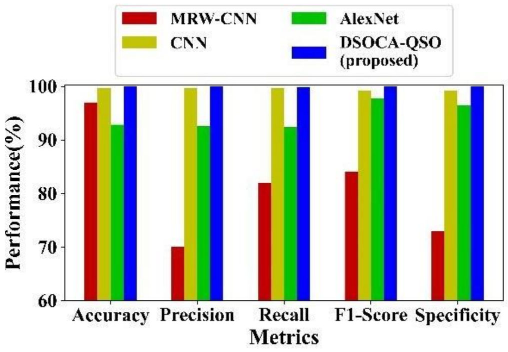
This analysis evaluates various metrics to accurately assess the performance and efficiency of the proposed method for plant disease detection across relevant datasets, highlighting its effectiveness in real-world applications.

3.3.1 Comparison of performance metrics

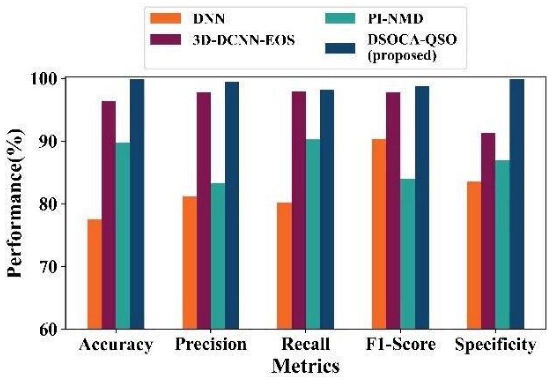
This section compares performance metrics (Accuracy, Precision, Recall, F1-Score, Specificity) across four datasets: Dataset 1, Dataset 2, Dataset 3 and Dataset 4. Models such as CNN, DNN, AlexNet, VGG-CNN, and the proposed DSOC-A-QSO are evaluated.

Received: August 05, 2025

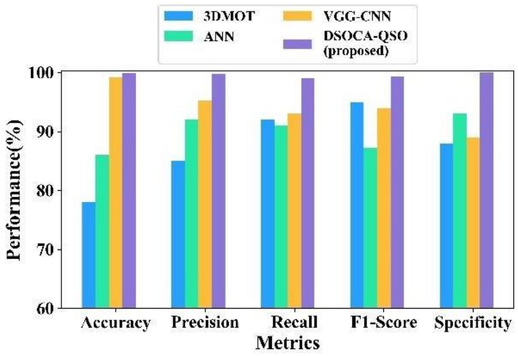
International Journal of Applied Mathematics
Volume 38 No. 11s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)



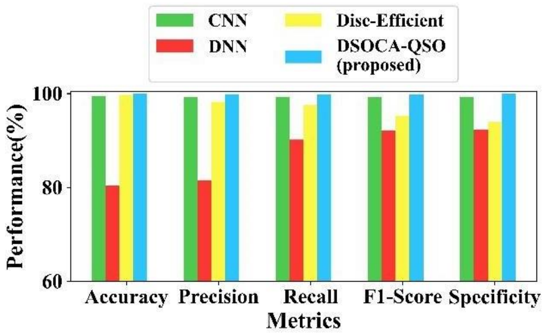
(a)



(b)



(c)



(d)

Figure 8: Comparison of performance metrics in (a) Dataset 1,(b) Dataset 2,(c) Dataset 3 and (d) Dataset 4

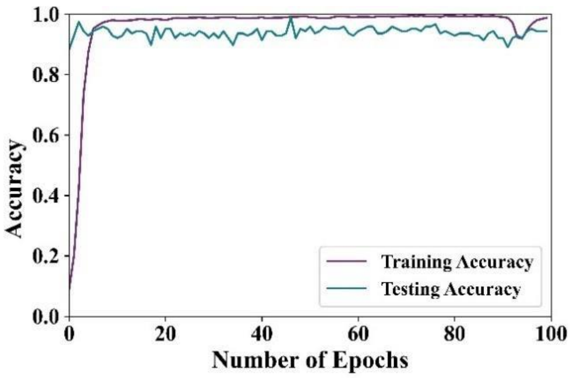
Figure 8 compares DSOCA-QSO with various models across Accuracy, Precision, Recall, F1-Score, and Specificity. Figure 8 (a), DSOCA-QSO achieves 99.9% accuracy, surpassing MRW-CNN, CNN, and AlexNet (75-95%). Figure 8 (b), it outperforms DNN, 3D-DCNN-FOS, and PI-VMD, exceeding 99%, while others range between 70-90%. Figure 8 (c), DSOCA-QSO maintains peak performance above 99.5%, compared to VGG-CNN, ANN, and 3DMOT (80-95%). Figure 8 (d), it surpasses CNN, DNN, and Disc-Efficient, reaching 99.8%,

International Journal of Applied Mathematics
Volume 38 No. 11s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

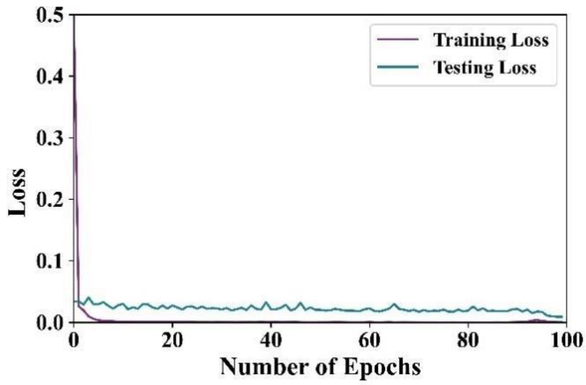
while others range from 70-96%. These results emphasize DSOCA-QSO's superior efficiency in disease classification and fertilizer recommendation.

3.3.2 Comparison of Accuracy and Loss

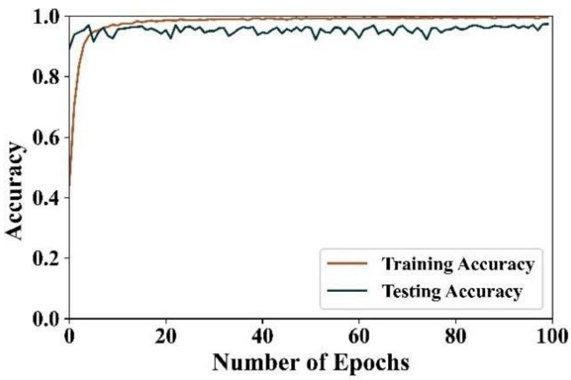
Figure 9 presents the performance evaluation of multiple models in terms of training and testing accuracy and loss over 100 epochs. Subplots 9 (a), (c), (f), and (h) illustrate accuracy trends, showing that both training and testing accuracies converge between 0.95 and 1.0, indicating excellent model performance and strong generalization. Subplots 9(b), (d), (e), and (g) depict the loss trends, with both training and testing losses decreasing steadily from initial values around 0.5–0.3 to below 0.1, reflecting effective learning and reduced error rates.



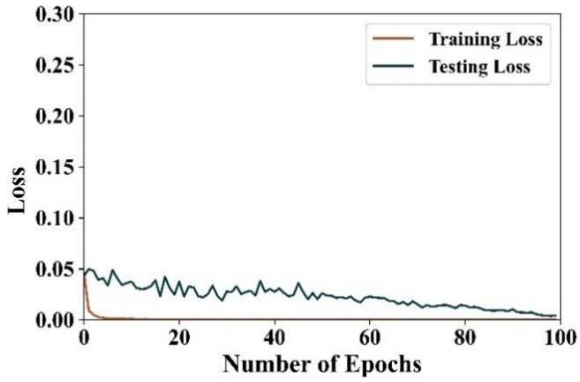
(a)



(b)



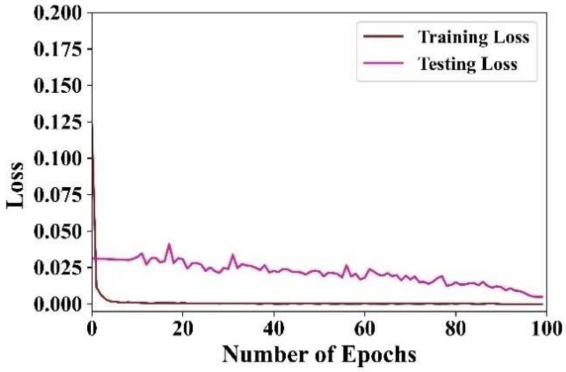
(c)



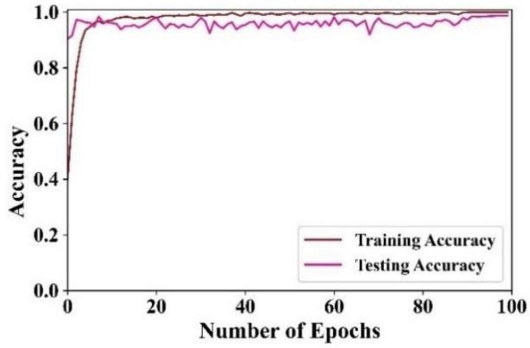
(d)

Received: August 05, 2025

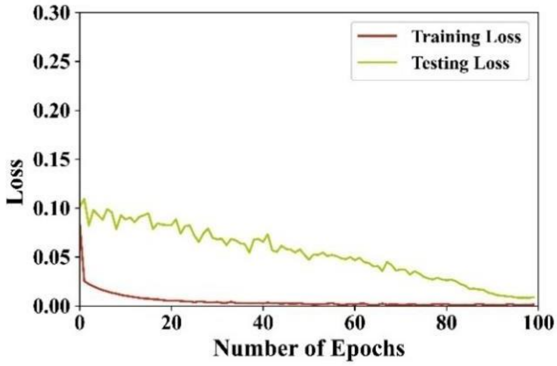
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)



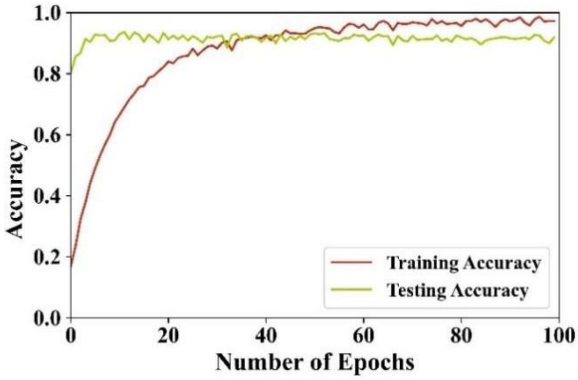
(e)



(f)



(g)



(h)

Figure 9: Comparison of Accuracy and loss in (a-b) Dataset 1 (c-d) Dataset 2 (e-f) Dataset 3 Dataset 3 (g-h) Dataset 4

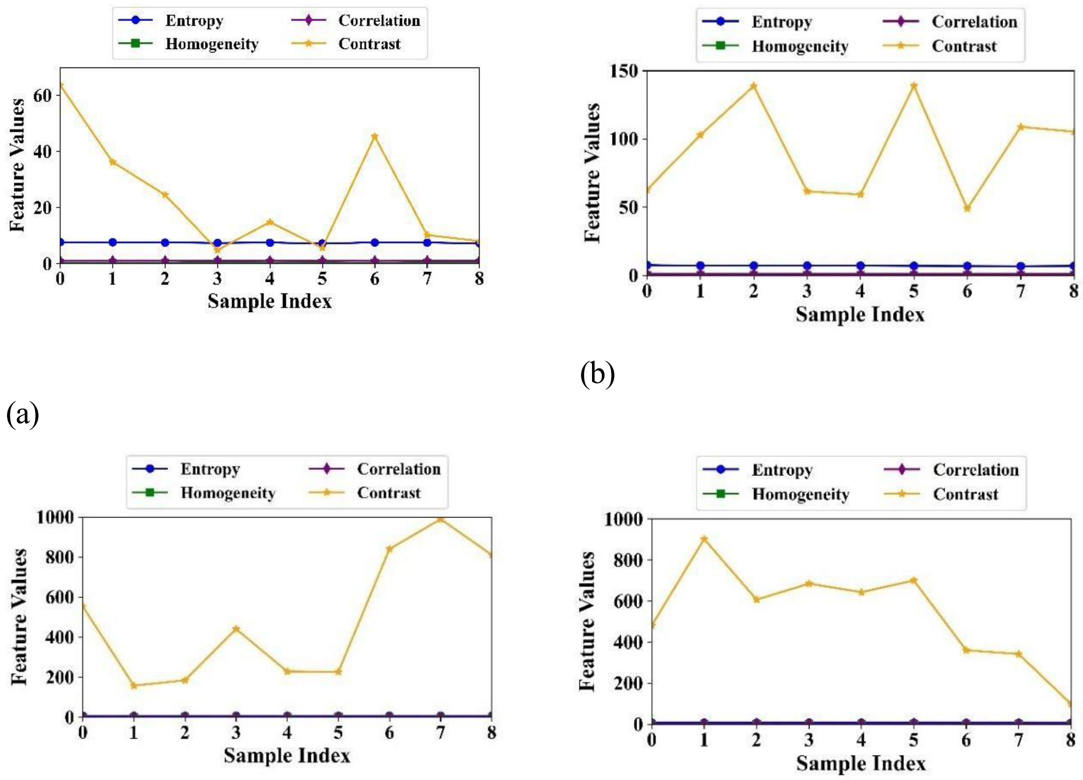
Notably, in subplot 9(e), the loss remains exceptionally low (below 0.05) early in training, suggesting a highly optimized model or dataset-specific learning behavior. Overall, these visualizations demonstrate efficient training processes, with the models achieving near-perfect accuracy and minimal losses by the end of 100 epochs.

3.3.3 Feature values across four datasets

Figure 10, four-line graphs show trends of four features, entropy, correlation, homogeneity, and contrast, with sample indices ranging from 0-8. In Figure 10 (a), entropy begins at around 60 for sample index 0 and rapidly declines, but Correlation, Homogeneity, and Contrast barely range from close to zero throughout. In Figure 10(b), contrast is the most robust feature, peaking at around 150 for sample index 1, while the other features stay close to zero throughout.

Received: August 05, 2025

International Journal of Applied Mathematics
Volume 38 No. 11s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)



(a) (b)

(c) (d)

Figure 10: Feature values across four datasets

Figure 10(c) illustrates the contrast peaks at around 1000 for sample index 0 and gradually declines, with the remaining features showing minimal variation near 0. Figure (d) signifies the feature values for entropy, homogeneity, correlation, and contrast across sample indices 0 to 8. The contrast feature values are quite large, with values starting around 500 at index 0, peaking at about 900 at index 1, and tapering off gradually to around 100 by index 8. The entropy, homogeneity, and correlation values collected from the samples remain very close to zero at all sample indices, indicating there is little variation and little contribution to the classification and scoring processes. These details demonstrate that contrast is the overall feature with the highest discriminative nature of the dataset.

3.3.4 Variance across four datasets

Figure 11 illustrates the variance across four datasets, like Dataset 1, Dataset 2, Dataset 3 and Dataset 4 for sample indices ranging from 0 to 9. The variance spans from 1000 to 6000.

International Journal of Applied Mathematics
Volume 38 No. 11s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

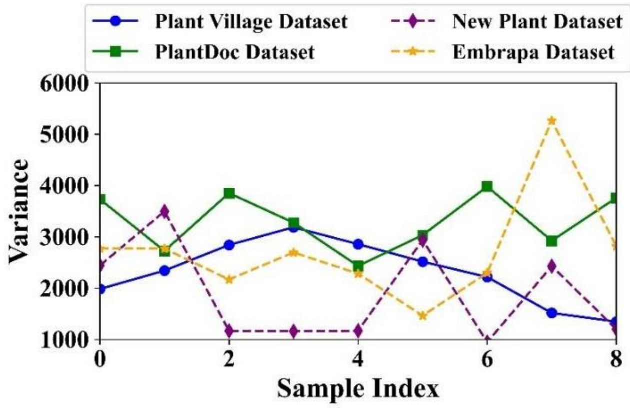


Figure 11: Variance across four datasets

Notably, the Dataset 1 reaches its highest point, near 6000, at index 6, while Datasets 2 and 4 maintain a relatively steady range between 2000 and 3000.

3.4.4 Confusion matrix of all datasets

Figure 12 contains four confusion matrices (a, b, c, and d) comparing different classification models or datasets for plant disease identification. Each matrix represents true versus predicted labels, with the diagonal indicating correct classifications.

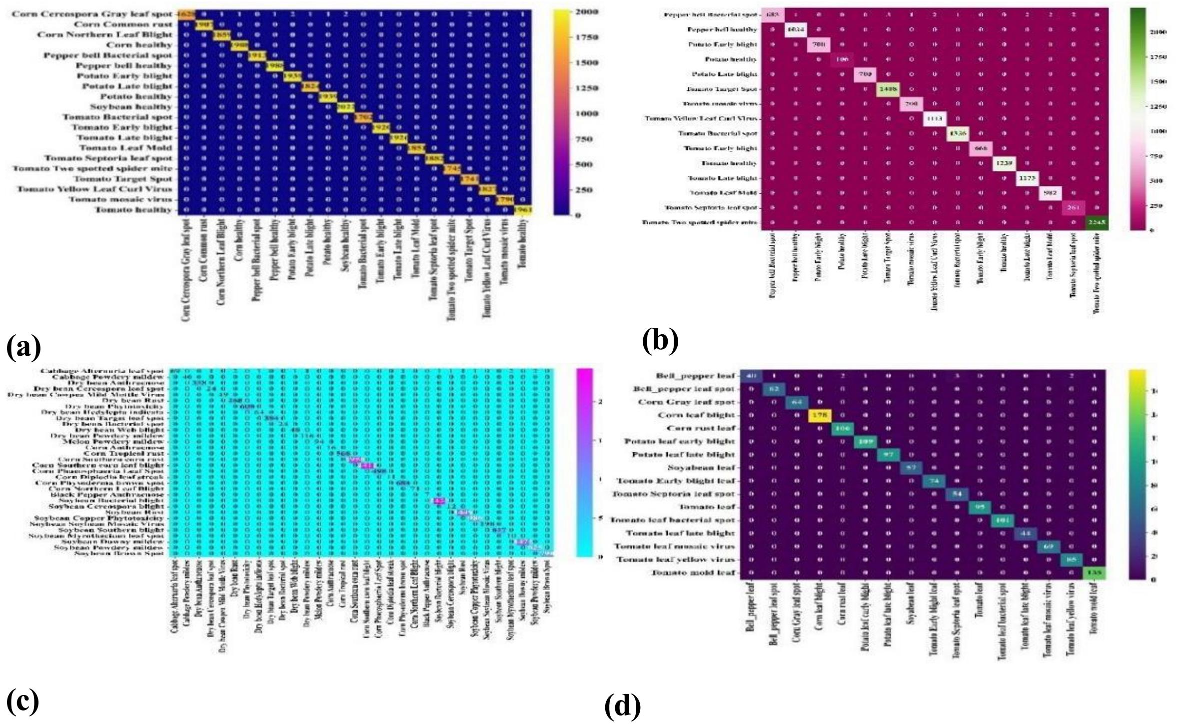


Figure 12: Confusion matrix of (a) Dataset 1(b) Dataset 2 (c) Dataset 3 (d) Dataset 4 The color intensity varies with classification accuracy, highlighting performance differences among classes. These matrices are essential for evaluating the effectiveness of machine learning models in accurately diagnosing plant diseases across diverse datasets.

International Journal of Applied Mathematics

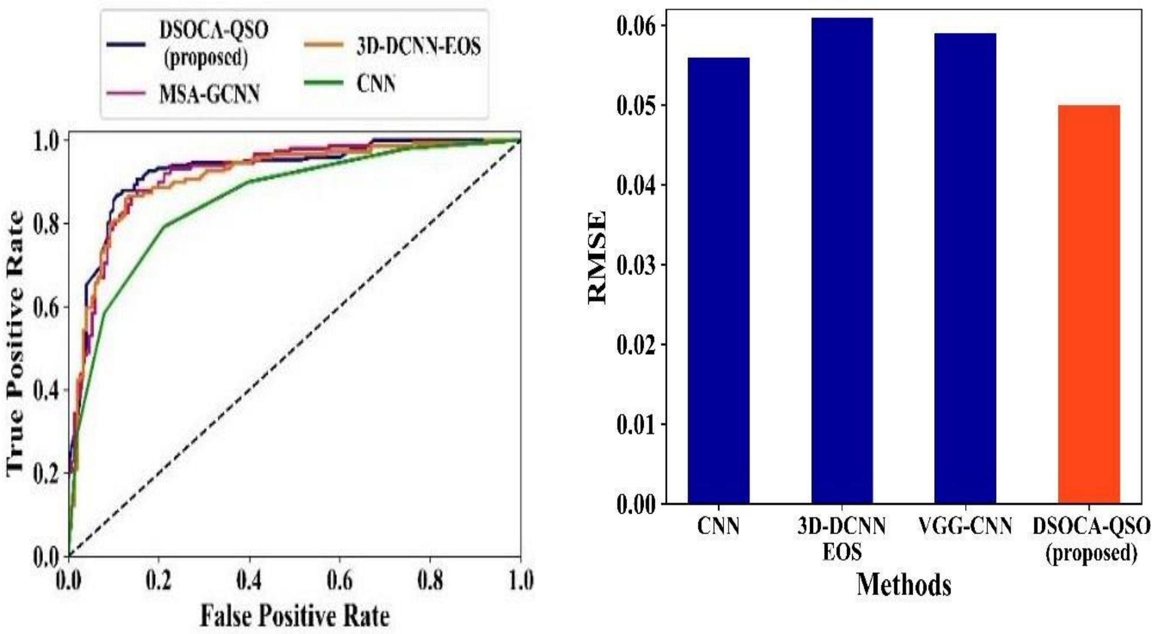
Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

3.4.5 Comparison of true positive value and RMSE:

Figure 13(a) displays the Receiver Operating Characteristic (ROC) curves for four methods: DSQCA-QSO (proposed), 3D-DCNN-EOS, MSA-GCNN, and CNN. The DSQCA-QSO

method achieves a consistently higher true positive rate across various false positive rates, outperforming the other approaches.



(b)

(a)

Figure 13: Comparison of (a) False positive rate and (b)RMSE

Figure 13(b) presents RMSE comparisons, highlighting DSQCA-QSO's performance (0.048), which surpasses CNN (0.052) and VGG-CNN (0.054) while remaining comparable to 3D- DCNN-EOS (0.045). Table 6 evaluates various methods across datasets based on sensitivity, RMSE, computational time, and kurtosis.

International Journal of Applied Mathematics
Volume 38 No. 11s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Table 6: Comparison of overall datasets

Dataset	Methods	Sensitivity (%)	RMSE	Computational time (sec)	kurtosis
Dataset 1	MRW-CNN [11]	93.26	0.21	8	0.56
	CNN [12]	92.56	0.22	9	0.44
	AlexNet [14]	94	0.23	10	0.61
	DSOCA-QSO (proposed)	**99.99**	**0.20**	**7.97**	**0.66**
Dataset 2	DNN [13]	91	95.91	4	0.61
	3D-DCNN-EOS [16]	93	96.58	6	0.70
	PI-NMD [17]	90.50	99.19	3.21	0.65
	DSOCA-QSO (proposed)	**98.20**	**95.90**	**2.24**	**0.72**
Datasets 3	3DMOT [18]	95.21	0.764	4.34	0.75
	ANN [19]	97.29	0.876	5.12	0.82
	VGG-CNN [20]	96.29	0.892	6.26	0.83
	DSOCA-QSO (proposed)	**98.99**	**0.604**	**3.19**	**0.84**
Dataset 4	CNN [12]	99.6	0.29	5.46	0.103
	DNN [13]	92.89	0.30	4.45	0.200
	Dise-Efficient [15]	93.56	0.31	3.26	0.109
	DSOCA-QSO (proposed)	**99.7%**	**0.28**	**3.48**	**0.201**

DSQCA-QSO consistently achieves the highest sensitivity, competitive RMSE, and efficient computational time, highlighting its superior accuracy and performance. Additionally, its stable kurtosis values demonstrate its robustness and reliability across all datasets.

3.4.6 Fitness performance

Figure 14 illustrates the fitness performance of four optimization techniques namely Ebola optimization

International Journal of Applied Mathematics
Volume 38 No. 11s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)
search (EOS) [16], Spider Monkey optimization (SMO) [20], Adam optimizer [15], and the proposed QSO over 100 iterations. The EOS method exhibits a steep initial decrease in fitness, dropping from approximately 50 to below 10 within the first 20 iterations before stabilizing. In contrast, SMO, Adam optimizer, and QSO achieve significantly faster convergence, reducing fitness values to below 2 within the first 10 iterations.

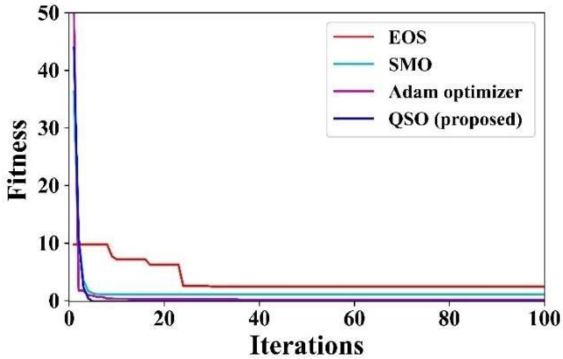


Figure 14: Fitness performance

Notably, the proposed QSO consistently maintains the lowest fitness value throughout all iterations, highlighting its superior performance. These results indicate that QSO achieves more effective optimization and faster convergence compared to the other methods.

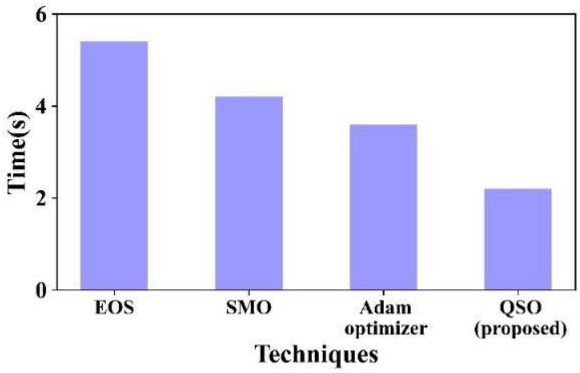


Figure 15: Comparison of time

Figure 15 shows the time taken by four optimization techniques, like EOS, SMO, Adam optimizer, and QSO, to solve a problem. QSO is the fastest, requiring approximately 2 seconds, whereas EOS is the slowest, taking around 5.5 seconds.

3.4.7 Ablation Study Analysis

An ablation study is a method for assessing the different parts or modules of a model that affect its overall performance. The performance metrics show that the Full DSOCA-QSO model outperforms its variants, achieving 99.9% accuracy, 99.7% precision, and 99.99% recall,

demonstrating exceptional robustness. Table 7 shows the impact of key components in DSOCA-QSO.

Table 7: Impact of key components in DSOCA-QSO

International Journal of Applied Mathematics
Volume 38 No. 11s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Components	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)	Sensitivity (%)
DSOCA-QSO without Attention	95	89	87	85	92.56
DSOCA-QSO without orientation	96	95	91	90	93.47
DSOCA-QSO without convolutional	98.65	98	96	95	97.59
Full DSOCA-QSO model	**99.9**	**99.7**	**99.99**	**99.35**	**98.99**

Removing convolutional layers reduces accuracy to 98.65%, while excluding orientation further lowers it to 96%. The DSOCA-QSO without attention yields the lowest performance, with an accuracy of 95%. These findings emphasize the critical role of attention, orientation, and convolution in enhancing model performance.

4. Discussion:

There have been notable improvements in multi-vegetable leaf disease classification through the DSOCA-QSO framework by integrating advanced preprocessing, feature extraction, and optimization techniques. The AEFGC enhances image quality for precise analysis, while the DSOCA network leverages orientation-guided attention to extract key features for disease identification effectively. QSO optimizes parameter tuning, tackling computational complexity and improving performance in high-dimensional scenarios. The SHAP interprets model output to build confidence in the decisions it makes. According to it, the fertilizer suggestion module strengthens the framework as it promotes plant health and increases yields. The performance measures add to the amount of accuracy and reliability of the system.

4.1 Limitation

The research faces several real-time challenges such as high computational demands, latency in large-scale applications, and reliance on high-resolution images. Variations in lighting and occlusions impact accuracy, limiting its effectiveness in dynamic agricultural environments.

International Journal of Applied Mathematics
Volume 38 No. 11s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

5. Conclusion

This research presents the DSOCA-QSO model, an innovative hybrid approach that combines the DSOCA with the QSO for efficient multi-vegetable leaf disease classification. The model has been greatly advanced in the accuracy of disease detection and the energy resource is reduced by deep learning and optimization integration. This approach evaluated the image preprocessing steps like resizing, denormalization, denoising, and contrast enhancement, bringing an image to its optimum quality for better classification performance. The DSOCA method extracts the significant features from the leaf images, while the orientation-guided convolutional attention focuses on the relevant patterns. Adding the QSO has improved the model further and fine-tuned the weight parameters that help it achieve higher efficiency on challenging tasks. By offering details about the model's predictions, SHAP enhances interpretability. The performance is evaluated using metrics such as recall (99.99%), accuracy (99.9%), precision (99.7%), and ROC analysis. Moreover, the automatic fertilization recommendations are provided following diagnosis-based disease detection, plant health and ultimately plant productivity. This is the most sophisticated solution compared to the traditional disease management practices in agriculture through the DSOCA-QSO model. Future research focuses on minimizing computational costs, enhancing real-time efficiency, and increasing resilience to environmental variations. Implementing edge computing and adaptive learning improves scalability, making it more suitable for practical agricultural applications.

References

1. Joseph DS, Pawar PM, Chakradeo K. Real-time plant disease dataset development and detection of plant disease using deep learning. IEEE Access. 2024 Jan 25.
2. Hassan SM, Maji AK. Plant disease identification using a novel convolutional neural network. IEEE Access. 2022 Jan 7; 10:5390-401.
3. Ahmad A, El Gamal A, Saraswat D. Toward generalization of deep learning-based plant disease identification under controlled and field conditions. IEEE Access. 2023 Jan 26; 11:9042-57.
4. Yang L, Yu X, Zhang S, Zhang H, Xu S, Long H, Zhu Y. Stacking-based and improved convolutional neural network: a new approach in rice leaf disease identification. Frontiers in Plant Science. 2023 Jun 6; 14:1165940.
5. Guan H, Fu C, Zhang G, Li K, Wang P, Zhu Z. A lightweight model for efficient identification of plant diseases and pests based on deep learning. Frontiers in Plant Science. 2023 Jul 14; 14:1227011.
6. Ashwini C, Sellam V. EOS-3D-DCNN: Ebola optimization search-based 3D-dense convolutional neural network for corn leaf disease prediction. Neural Computing and Applications. 2023 May;35(15):11125-39.
7. Batchuluun G, Hong JS, Wahid A, Park KR. Plant image classification with nonlinear motion deblurring based on deep learning. Mathematics. 2023 Sep 21;11(18):4011.
8. Ahm edt-Aristizabal D, Smith D, Khokher MR, Li X, Smith AL, Petersson L, Rolland V, Edwards EJ. An In-Field Dynamic Vision-Based Analysis for Vineyard Yield Estimation. IEEE Access. 2024 Jul 19.
9. Bolfe ÉL, Parreiras TC, Silva LA, Sano EE, Bettiol GM, Victoria DD, Sanches ID, Vicente LE. Mapping agricultural intensification in the Brazilian savanna: A machine learning approach using

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

harmonized data from Landsat Sentinel-2. ISPRS International Journal of Geo-Information. 2023 Jul 2;12(7):263.

10. Thakur PS, Sheorey T, Ojha A. VGG-ICNN: A Lightweight CNN model for crop disease identification. Multimedia Tools and Applications. 2023 Jan;82(1):497-520.
11. Yuan Z, Zeng J, Wei Z, Jin L, Zhao S, Liu X, Zhang Y, Zhou G. CLAHE-based low-light image enhancement for robust object detection in overhead power transmission system. IEEE Transactions on Power Delivery. 2023 Apr 21;38(3):2240-3.
12. Zhang D, He Z, Zhang X, Wang Z, Ge W, Shi T, Lin Y. Underwater image enhancement via multi-scale fusion and adaptive color-gamma correction in low-light conditions. Engineering Applications of Artificial Intelligence. 2023 Nov 1;126:106972..
13. Lu WQ, Hu HM, Yu J, Zhou Y, Wang H, Li B. Orientation-aware pedestrian attribute recognition based on graph convolution network. IEEE Transactions on Multimedia. 2023 Mar 22;26:28-40.
14. Bhujel A, Kim NE, Arulmozhi E, Basak JK, Kim HT. A lightweight Attention-based convolutional neural network for tomato leaf disease classification. Agriculture. 2022 Feb 4;12(2):228.
15. Bo J, Lei L, Sheng Z, Shan-shan Y, Shu-guang Z, Yao H, Xiao-yu L. An automatic detection for solar active regions based on scale-invariant feature transform and clustering by fast search and find of density peaks. Chinese Astronomy and Astrophysics. 2022 Jul 1;46(3):264-76.
16. Thorat T, Patel BK, Kashyap SK. Intelligent insecticide and fertilizer recommendation system based on TPF-CNN for smart farming. Smart Agricultural Technology. 2023 Feb 1; 3:100114.
17. https://www.kaggle.com/datasets/vipoooool/new-plant-diseases-dataset
18. https://www.kaggle.com/datasets/abdulhasibuddin/plant-doc-dataset
19. https://www.kaggle.com/datasets/emmarex/plantdisease
20. https://www.kaggle.com/datasets/poornimasinghthakur/embrapa