WES-HOA-based ReScaleX-CNN for Plant Leaf Disease Classification

Spoorthi Pothaganahalli Akkalappa, Shivaputra, Meenakshi Laxman Rathod, Dinesha Puttaraje Gowda, and Smitha Sasi

Original scientific article

***Abstract*—Plant leaf disease classification involves identification and classification of different diseases based on indicators of plant leaves, which plays a crucial role in managing crop health. However, classifying plant leaf diseases is challenging due to wide variation in leaf shape, texture, and color, which leads to overlapping symptoms and inaccurate classification. In this research, the Wombat Escape Strategy-Hippopotamus Optimization Algorithm-based Recalibrated Multi-Scale Squeeze and Excitation Convolutional Neural Network (WES-HOA-based ReScaleX-CNN) is proposed to classify plant leaf disease accurately. In HOA, WES is incorporated to select the most appropriate features that enhance global searchability by guiding agents away from local optima. ReScaleX-CNN enhances the model's ability to concentrate on informative features by emphasizing significant spatial and channel-wise information. The multiscale approach captures disease patterns at various resolutions, leading to robust performance. Hence, the proposed method obtains a high accuracy of 99.92% on PlantVillage dataset in comparison with existing methods, such as DeepPlantNet.**

***Index terms*—Local Optima, PlantVillage, Premature Convergence, Recalibrated Multi-Scale Squeeze and Excitation Convolutional Neural Network, Wombat Escape Strategy-Hippopotamus Optimization Algorithm.**

I. Introduction

Recently, plant disease prevalence has significantly increased because of factors such as globalization, climate change, and international trade [1]. The application of image processing and computer vision technique is an alternative approach to enhance the process of determining diseases via leaf symptoms [2], [3].

Due to diseases caused by plants, more than 5% of agricultural production is wasted, which leads to major financial losses to farmers [4]-[7]. The agricultural industry uses expert knowledge and manual inspection to diagnose plant diseases using techniques such as microscopy, thermography, and biological diagnosis [8-10]. The application of Machine Learning (ML) and Deep Learning (DL) for detecting plant diseases is a rapidly evolving field that has demonstrated significant results. A major limitation of ML is the manual extraction of features that fail to capture complex and subtle patterns in plant leaves [11]-[13]. In contrast, DL automatically learns appropriate features, which makes it more effective for identifying complex and abstract visual cues [14], [15]. DL provides a significant advantage in integrating the feature extraction and classification methods into a unified model [16-18]. Kotwal et al. [19] developed a hybrid method depending on optimized automatic DL for Plant Leaf Disease Classification (PLDC). The weight of the U-Net was tuned using Hunter Prey Optimization (Hunt-PO) during segmentation whereas Artificial Driving-EfficientNet (AD-E Net) was employed to classify plant leaves. Shafik et al. [20] suggested the two Plant Disease Detection models (PDDNet) of Arithmetic Average Ensemble (AAE) and Lead Voting Ensemble (LVE).

Ullah et al. [21] presented an efficient DeepPlantNet to automatically classify plant diseases at different stages. To extract high-level characteristics from obtained images, DeepPlantNet utilized a convolution layer and Leaky Rectified Linear Unit (LReLU), which provided better performance. Zhang et al. [22] established a Capsule Network (CapsNet) and ResNet to classify plant leaf diseases accurately. The initial convolutional layer of ResNet was replaced by a kernel with a concatenation of small convolutional kernels that effectively extracted the features of plant leaf lesions. Shehu et al. [23] introduced three transfer learning methods, such as Vision Transformer (ViT)-ImageNet, ViT-Small, and ViT-Base, for predicting leaf diseases from images. Transfer learning had the ability to capture global contextual information via a self-attention mechanism, which provided better performance. Nobel et al. [24] presented an ensemble model in which DenseNetMini was integrated with a Gradient Product (GP) as an optimization method to classify plant leaf disease. However, the classification process was challenging because of the high intra-class variability in leaf color, shape, and texture, which led to minimized model performance. Hence, a feature-selection-based Wombat Escape Strategy-Hippopotamus Optimization Algorithm with a calibrated multi-scale squeeze and excitation Convolutional Neural Network (WES-HOA-based ReScaleX-CNN) is proposed for plant leaf disease. WES-HOA selects the most appropriate features by guiding the search process away

Manuscript received November 28, 2025; revised December 16, 2025. Date of publication August 26, 2026. Date of current version August 26, 2026. The associate editor prof. Ivo Stančić has been coordinating the review of this manuscript and approved it for publication.

S. P. Akkalappa, S. Shivaputra and M. L. Rathod are with the Department of Electronics and Communication Engineering, Dr. Ambedkar Institute of Technology, Bengaluru, Karnataka, India (e-mails: {spoorthiyadav.24, putrauvce}@gmail.com, meenakshilr.ec@drait.edu.in).

D. P. Gowda and S. Sasi are with the Department of Electronics and Communication Engineering, Dayananda Sagar College of Engineering, Bengaluru, Karnataka, India (e-mails: drdinesh-ece@dayanandasagar.edu, smitha-tce@dayanandasagar.edu)

Digital Object Identifier (DOI): 10.24138/jcomss-2025-0194

from the local optima, whereas ReScaleX-CNN captures both fine-grained and large-scale disease features, leading to accurate performance.

A primary contribution of this research is described as follows:

- WES-HOA is used to effectively select the most appropriate features from plant leaf images using an escape strategy that avoids local optima and enhances convergence speed. This results in enhanced classification performance and minimizes the computational complexity.
- ReScaleX-CNN enhances the feature representation by refining the spatial and channel responses. Its multi-scale design captures different leaf disease patterns over diverse resolutions through parallel pathways, leading to more reliable classification.
- Resizing standardizes the image dimensions by enabling consistent feature extraction, whereas min-max scales the values of pixel intensity to a uniform range that maximizes contrast and feature discrimination.

This research paper is organized as follows: Section II illustrates the proposed methodology, Section III presents the research method, Section IV analyzes the results and discussion, and the conclusion of this research paper is given in Section V.

II. Proposed Methodology

This research proposes a WES-HOA-based ReScaleX-CNN to select the most appropriate features and classify plant leaf diseases accurately. Initially, PlantVillage dataset is used for model performance evaluation. Resizing and min-max normalization are employed to resize the image dimensions and normalize the resized images. WES-HOA selects the most appropriate features, while ReScaleX-CNN classifies plant disease. Fig. 1. demonstrates the block diagram of the methodological process.

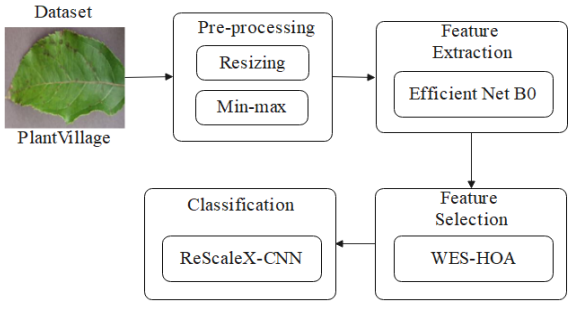


Fig. 1. Block diagram of the overall methodology process.

A. Datasets

This research employs the PlantVillage [25] and PlantDoc [26] datasets to evaluate the model performance for accurately classifying plant leaf diseases. PlantDoc is used for generalization analysis to evaluate unseen data. PlantVillage contains 54,303 images of healthy and diseased leaves from 14 crop species. The datasets are split into 80% training (43,443 images) and 20% testing (10,860 images), respectively. PlantDoc has 2,598 images gathered over 13 plant species and 17 disease classes with nearly 300 human-hours invested in annotating internet-scraped images. Fig. 2 depicts sample images of PlantVillage dataset, which is fed into the preprocessing stage.

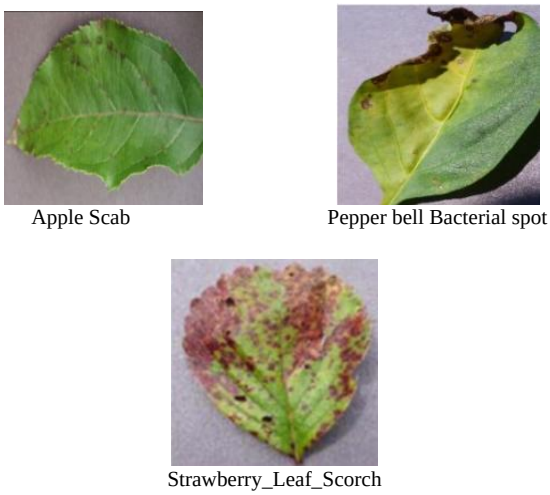


Fig. 2. Sample image from PlantVillage dataset.

B. Preprocessing

After obtaining the images, resizing is applied as 256×256 dimension, ensuring that all images have uniform dimensions. Then, min-max normalize all image pixel values to a [0, 1] for consistent input distribution in (1). The pixel intensity value is represented as x_i ; $min(x)$ and $max(x)$ indicates minimum and maximum values of pixel intensity and x_{norm} signifies normalized pixel intensity which are fed into feature extraction stage using EfficientNetB0.

$$x_{norm} = \frac{x_i - min(x)}{max(x) - min(x)} \tag{1}$$

C. Feature Extraction

The pre-processed images are extracted using EfficientNetB0 which preserves spatial consistency and captures fine-grained information effectively. EfficientNetB0 depends on inverted bottleneck residual blocks of MobileNetV2 with squeeze and excitation blocks, which learns more discriminative features. Table I lists the compound-scaling configurations for an EfficientNetB0. The extracted 1,280 features are passed as input into the feature-selection stage for further processing.

TABLE I
Configuration of Compound Scaling for EfficientNetB0

Method	Depth coefficient	Width coefficient	Dropout rate	Resolution
Efficient NetB0	1.0	1.0	0.2	224

III. Research Method

A. Wombat Escape Strategy-Hippopotamus Optimization Algorithm (WES-HOA)

After feature extraction, HOA is used to select the most appropriate features, which significantly contributes to

differentiating between healthy and diseased leaves. HOA [27] balance exploration of new areas and exploitation of known food sources, which makes HOA search for global features more effectively than Lion Optimization Algorithm (LOA) and Zebra Optimization Algorithm (ZOA).

Initialization: An initial solution of HOA is randomly generated using (2) where $x_{i,j}$ demonstrates position of i^{th} candidate solution; X denotes the population of candidate solutions; r represents a random value (0,1); and ub_j and lb_j are upper and lower bounds of j^{th} decision variables. Considering that N determines the hippopotamus population size and m represents number of decision variables, population matrix is formulated as (3).

$$x_{i,j} = lb_j + r.\left(ub_j - lb_j\right) \tag{2}$$

$$X = \begin{bmatrix} X_1 \\ \vdots \\ X_i \\ \vdots \\ X_N \end{bmatrix}_{N\times m} = \begin{bmatrix} x_{1,1} & \cdots & x_{1,j} & \cdots & x_{1,m} \\ \vdots & \ddots & \vdots & \iddots & \vdots \\ x_{i,1} & \cdots & x_{i,j} & \cdots & x_{i,m} \\ \vdots & \iddots & \vdots & \ddots & \vdots \\ x_{N,1} & \cdots & x_{N,j} & \cdots & x_{N,m} \end{bmatrix}_{N\times m} \tag{3}$$

Exploration (Location update): The male hippopotamus is driven away by a dominant male and competes with other male hippopotamuses to establish its advantage, and its position is represented in (4). Where x_{ij}^{M} determines male hippopotamus position, y_1 represents random range [0, 1], D refers to the position of the dominant hippopotamus, and I_1 and I_2 indicates random integer numbers selected from the population expressed in (5) and (6).

$$x_{ij}^{M} = x_{ij} + y_1\left(D - I_1 x_{ij}\right) \tag{4}$$

$$h = \begin{cases} I_2 \times r_1 + \left(\sim Q_1\right) \\ 2\times r_2 - 1 \\ \vec{r_3} \\ I_1 \times r_4 + \left(\sim Q_2\right) \\ r_5 \end{cases} \tag{5}$$

$$T = exp\left(-\frac{t}{\Gamma}\right) \tag{6}$$

$$x_{ij}^{FB} = \begin{cases} X_{ij} + h_1\left(D - I_2 MG_1\right), T > 0.6 \\ E, else \end{cases} \tag{7}$$

$$E = \begin{cases} x_{ij} + h_2\left(MG_i - D\right), r_6 > 0.5 \\ lb_j + r_7\left(ub_j - lb_j\right), else \end{cases} \tag{8}$$

where r_1 to r_6 represents a random vector, MG_1 denotes the best mean position of the hippopotamus population group, MG_i determines the best mean position of the hippo i^s group, Γ illustrates the exploration decay coefficient, T indicates a maximum number of iterations, E determines the intermediate-based position, h_1 refers to the control function that regulates the exploration stage, and h_2 describes the scaling function controlling local exploitation. Equations (7) and (8) refer to immature hippopotamuses and female positions in the population. Equations (9) and (10) represent the male hippopotamus' position in the population and the updated position of another vulnerable hippopotamus, where F_i denotes the objective function value and FB determines feedback-based updated position of the hippopotamus.

$$x_i = \begin{cases} x_i^{M}, F_i^{M} < F_i \\ x_i, else \end{cases} \tag{9}$$

$$x_i = \begin{cases} x_i^{FB}, F_i^{FB} < F_i \\ x_i, else \end{cases} \tag{10}$$

Exploitation (Hippo defense against predators): A vulnerable hippopotamus leaves the group because of various factors that make easy target for predation by larger organisms, using (11) to (12).

$$P_j = lb_j + r_8\left(ub_j - lb_j\right) \tag{11}$$

$$D = \left|P_j - x_{ij}\right| \tag{12}$$

here P_j denotes distance from predator to i^{th} group, whereas Hippo employs defensive behaviors depending on FP_j for protection from predators, and D indicates the distance among vulnerable hippopotamuses and current hippopotamus position, and FP_j represents feedback-based updated position of hippopotamus in j^{th} leader. If FP_j is greater, (13) indicates that the predator territory is located farther away. Where x_{ij}^{HR} represents the hippopotamus posture while facing predators and R_L denotes position changes of predator. A levy distribution is represented using (14), where ω and v are random. If FP_j is greater than F_i, which demonstrates the hippopotamus' position is replaced; else, hippopotamus returns to population using (15).

$$x_{ij}^{HR} = \begin{cases} R_L \oplus P_j + \left(\frac{f}{\delta - dcos\left(2\pi y\right)}\right).\left(\frac{1}{D}\right), FP_j < F_i \\ R_L \oplus P_j + \left(\frac{f}{\delta - dcos\left(2\pi y\right)}\right).\left(\frac{1}{2\times D + r_9}\right), FP_j \geq F_i \end{cases} \tag{13}$$

$$L\left(v\right) = 0.05 \times \frac{w \times \sigma_\omega}{\left|v\right|^{\frac{1}{v}}} \tag{14}$$

$$x_i = \begin{cases} x_i^{HR}, F_i^{HR} < F_i \\ x_i, F_i^{HR} \geq F_i \end{cases} \tag{15}$$

Development stage (Hippo escape from predators): The development stage is modeled to enhance local search exploitation. A random position is generated using (16)–(19) to simulate this behavior in vicinity of present hippopotamus location, where t indicates number of iterations. If a newly generated position results in a lower F_i, it specifies that hippopotamus determines a safer nearby location and moves to replace the previous position. In (17), x_{ij}^{HE} searches for a hippopotamus location to determine nearest safe location. The s_1 is selected from (18), which produces stronger local search abilities. Where r_{11} indicates a random vector [0, 1], and r_{12} and r_{13} determine the random variables that follow a normal distribution using (19). The Mean Square Error (MSE) is used as a fitness function due to its effective measure of difference between predicted and actual values in (20).

$$lb_j^{local} = \frac{lb_j}{t}, ub_j^{local} = \frac{ub_j}{t}, t = 1, 2, \ldots T \tag{16}$$

$$x_{ij}^{HE} = x_{ij} + r_{10}\left(lb_j^{local} + s_1\left(ub_j^{local} - lb_j^{local}\right)\right) \tag{17}$$

$$s_1 = \begin{cases} 2 \times r_{11} - 1 \\ r_{12} \\ r_{13} \end{cases} \tag{18}$$

$$x_i = \begin{cases} x_i^{HE}, F^{HE} < F_i \\ x_i, F_i^{HE} \geq F_i \end{cases} \tag{19}$$

$$Fitness\ function = \alpha.MSE(X) + \beta.\frac{|X|}{d} \tag{20}$$

where α represents weight coefficient controlling the significance of MSE in fitness calculation, β denotes weight coefficient controlling the significance of feature subset size penalty, and d determines total number of features.

A.1 Improved Strategy

To escape from hyenas and lions, hippopotamus jumps into water bodies but remains vulnerable to aquatic predators such as crocodiles. To address this issue, an escape strategy inspired by the wombat animal updates the position of search agents that enable them to navigate toward safer and predator-free regions more effectively. When a hippopotamus is in danger and attacked by a crocodile, it escapes by quickly moving toward a nearby safe location which protects from the threat. Modeling the hippopotamus movement as it escapes toward a safe location which increases exploitation ability and allows for more effective search space. Inspired by wombat's escape behaviour, new position is calculated for each member using (21). If newly generated position produces an improved objective function value, member's previous position is updated based on (22) when probability threshold is $p \leq 0.5$; otherwise, (17) is executed.

$$x_{ij}^{P2} = x_{ij} + \left(1 - 2r_{ij}\right).\frac{ub_j - lb_j}{t} \tag{21}$$

$$X_i = \begin{cases} X_i^{P2}, & F_i^{P2} \leq F_i \\ X_i, & else \end{cases} \tag{22}$$

where X_i^{P2} represents new position computed for i^{th} wombat during the escape stage, x_{ij}^{P2} indicates new position computed for i^{th} wombat during the escape stage in j^{th} position, r_{ij} represents a random number from interval [0, 1], and F_i^{P2} denotes fitness value of updated escape position. Fig. 3 illustrates a flowchart of the WES-HOA process. This process assists in removing redundant data, enhancing model accuracy, and minimizing computational complexity. After selecting 1,110 features, the classification process is performed using ReScaleX-CNN. By selecting most informative 1,110 features, learning process becomes more focused and ensures the classifier emphasizes disease-relevant patterns, which minimizes overfitting risk. Although selected features decrease from 1,280 to 1,110, representing only a 13% decrease, they offer several significant practical benefits for the learning process. Removing even a small number of irrelevant features minimizes noise in the learning process that makes the model concentrate on discriminative information, which enhances generalization.

B. Classification Using Recalibrated Multi-Scale Squeeze And Excitation Convolutional Neural Network (RESCALEX-CNN)

The selected features are passed through ReScaleX-CNN to classify the plant leaf effectively. CNN automatically learns feature representations directly from selected features and captures spatial and texture patterns to distinguish subtle disease variations. In CNN, multi-scale Squeeze and Excitation (SE) block is introduced to recalibrate channel-wise feature by capturing global and fine-grained information. The detailed design of ReScaleX-CNN are described below.

Convolution layer and Rectified Linear Unit (ReLU): Selected features are fed into convolutional layers containing 256 filters to extract rich and discriminative feature representation using (23). Where P denotes p^{th} layer, k_{ij} represents the convolutional kernel, N_j determines the bias, and M_j represents a set of input maps. ReLU introduces nonlinearity into the model while maintaining computational efficiency using (24).

$$M_j^p = f\left(\sum_{i \epsilon M_j} M_i^{p-1} \times k_{ij}^p + N_j^P\right) \tag{23}$$

$$ReLU(x) = max(0, x) \tag{24}$$

ReScaleX: Standard SE performs feature recalibration using single global descriptor from Global Average Pooling (GAP), which assumes that all feature maps are adjusted using a single-scale representation. However, plant leaf diseases exhibit variations in scale and morphology, ranging from small and isolated spots to large and medium-sized lesions. A global

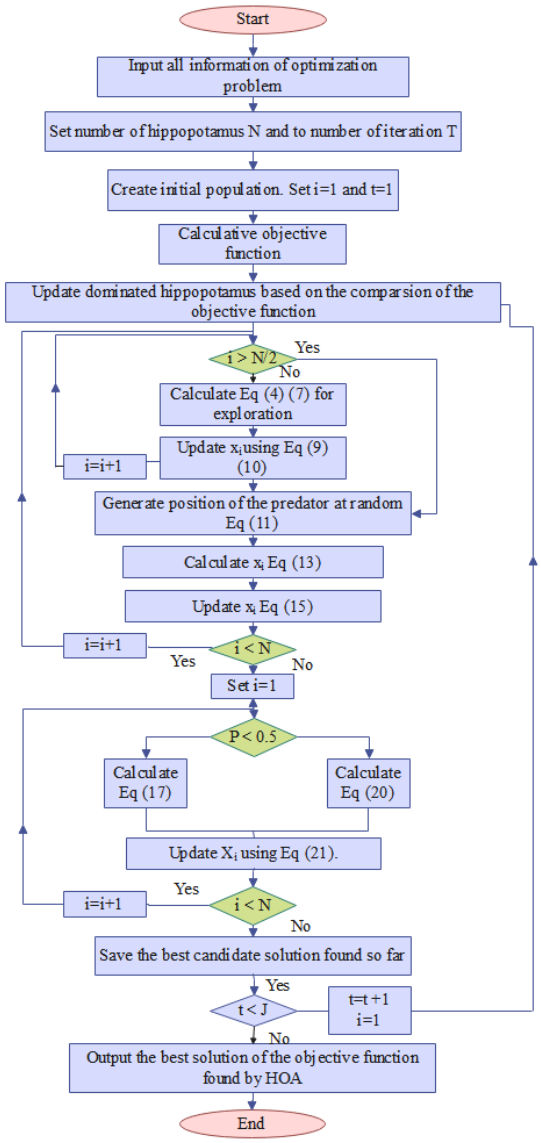


Fig. 3. Flowchart of WES-HOA process.

pooling fails to capture the rich variability in contextual and spatial patterns. To solve this, ReScaleX is used, which introduces a multi-scale excitation module containing three parallel SE pathways ($1 \times 1, 3 \times 3, 5 \times 5$) by capturing features at different spatial granularities. The $1 \times 1 (\alpha = 0.2)$ concentrates on fine-grained and local features such as tiny disease spots and dot-like symptoms, $3 \times 3 (\beta = 0.35)$ captures mid-range features such as medium-sized patches, blots, and spots, and 5×5 ($\gamma = 0.45$) emphasize large-scale structural features such as full-leaf blight and wide discolorations. GAP is performed on input feature maps in the squeeze step using (25), and the structure of ReScaleX is presented in Fig. 4.

$$z_c = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} X_c(i, j) \tag{25}$$

where X_c represents c^{th} input channel; H and W denotes height and width of feature map and z_c is GAP output for c^{th} channel. Each SE path utilizes a distinct reduction ratio m, $m_1 = 4, m_2 = 8, m_3 = 16$ resulting in varying degrees of feature compression across different scales. Each SE path outputs a vector of shapes, and these three outputs are weighted and fused dynamically, as formulated in (26). Where coefficients α, β, γ used in multi-scale SE are learnable parameters that are optimized within the network during backpropagation. The final SE vector is multiplied by feature maps to recalibrate feature importance using (27).

$$SE_{final} = \alpha.SE_{1\times1} + \beta.SE_{3\times3} + \gamma.SE_{5\times5} \tag{26}$$

$$X_c^{scaled} = SE_{final,c}.X_c \tag{27}$$

FC layer: After scaling, 2D spatial feature maps are flattened into a 1D vector and passed via FC layer. The layer-wise architecture of the proposed ReScaleX-CNN is demonstrated in Fig. 5. The hyperparameter values are chosen based on grid search [28] to ensure optimal performance which ensures reproducibility and transparency. The search space included the learning rate (0.01, 0.001, 0.005, and 0.0001), dropout rate (0.2 to 0.6), and batch size (16, 32, and 64). After tuning process, final configuration is determined with 0.001 learning rate to achieve an effective balance between convergence speed and training stability whereas dropout rate of 0.5 solve overfitting. Adam optimizer and a batch size of 32 increase stability, whereas ReLU introduces nonlinearity, that avoids vanishing gradient issue. The primary evaluation criteria for selecting final configuration are maximization of classification accuracy while minimizing MSE which act as fitness function to determine optimal model configuration. The stability of selected parameters is validated by 5-fold cross validation, providing balance between model variance and training noise ensuring consistent performance. The population size is explored in range of 10 to 30 where 24 is chosen due to providing balance between search space and computation cost. The maximum iterations are varied between 100 to 500 where 100 is selected since it achieves stable convergence. The lower and upper bounds are fixed with (0,1) to maintain normalized search space and consistent solution. A exploration decay factor is explored in range of 0.5 to 1.5 where 1.0 is selected and probability threshold is varied from 0.1 to 0.9 where 0.5 is chose as it ensures balanced trade-off between exploration and exploitation. The MSE weight coefficient is explored in range of 0.1 to 1.0 where 0.9 is selected to prioritize error minimization that leads to enhanced performance. The feature subset penalty is varied between 0.01 to 0.1 and 0.1 is chosen to effectively manage feature redundancy and maintain relevant feature selection.

IV. Results and Discussion

The proposed ReScaleX-CNN is simulated using MATLAB with 64 GB Random Access Memory (RAM), an Intel i5 processor, and Windows 10 operating system. Accuracy, F1-score, recall, and precision [22] are used as classification metrics to evaluate model performance.

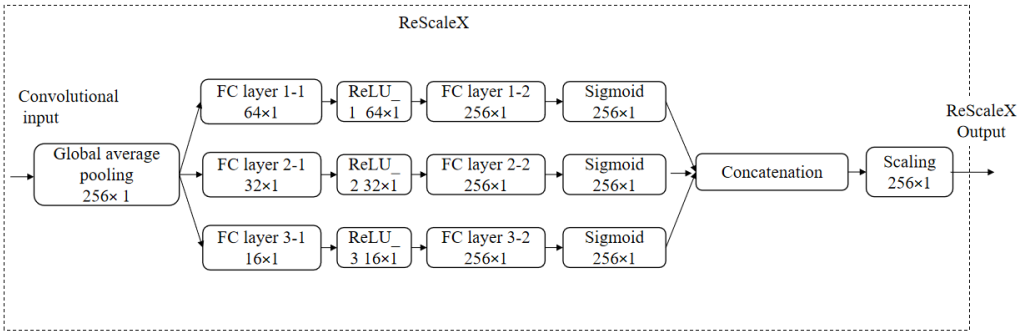


Fig. 4. Structure of ReScaleX.

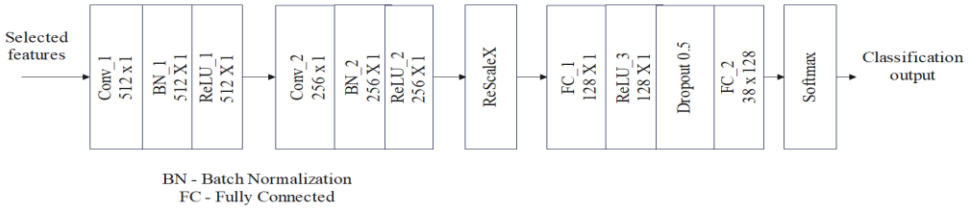


Fig. 5. Layer-wise architecture of the proposed ReScaleX-CNN.

A. Performance Analysis

Fig. 6. illustrates the performance analysis of various feature selection methods on the PlantVillage dataset. Compared to existing methods such as LOA, ZOA, Coati Optimization Algorithm (COA), and HOA, WES-HOA achieves a high accuracy of 99.92% on PlantVillage because of position-updating mechanism which avoid local optima. Table II presents performance evaluation of different classifiers without feature selection on PlantVillage. ReScaleX-CNN obtains a high accuracy compared to Deep Neural Networks (DNNs), ViT, Swin Transformer, Pyramid Vision Transformer (PVT), CNN, and CNN + SE. Without feature selection, model complexity is unnecessarily enhanced.

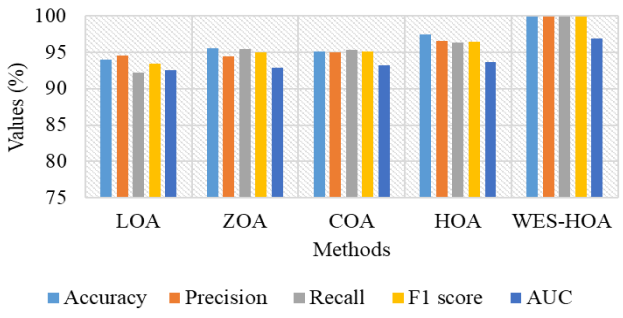


Fig. 6. Graphical representation of different feature selection methods on the PlantVillage dataset.

Table III presents performance analysis of different classifiers with feature selection. ReScaleX-CNN obtains a superior accuracy because redundant features are eliminated whereas SE block improve feature channels efficiently. Fig. 7. depicts performance evaluation of k-fold cross validation on PlantVillage. Using k=5 yields high accuracy because it offers optimal balance between variance and noise.

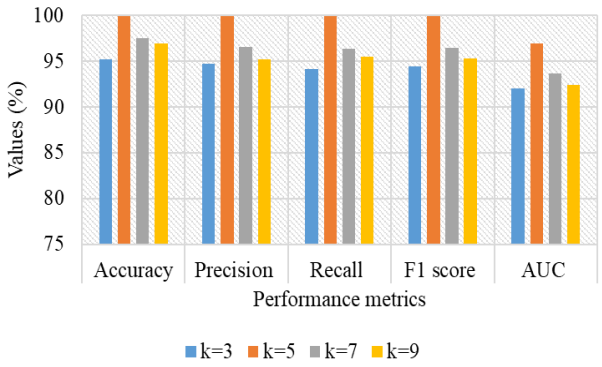


Fig. 7. Graphical representation of k-fold cross validation on the PlantVillage dataset.

Table IV provides a performance analysis of complexity and memory consumption. ReScaleX-CNN achieves a run time of 25.62 s. SE block emphasizes the significant channels and removes unwanted channels that minimize redundant computations. A paired t-test is employed to evaluate pairwise performance differences between proposed method and each baseline, whereas one-way ANOVA determines overall difference among all methods over 5 runs. Calculating CI is essential due to quantifying statistical reliability over 5 runs which shows robustness.

Table V indicates effectiveness of WES-HOA feature selection. By decreasing feature count from 1,280 to 1,110, WES-HOA minimizes training time and inference time, which enhances computational efficiency.

TABLE II
PERFORMANCE ANALYSIS OF DIFFERENT CLASSIFIERS WITHOUT FEATURE SELECTION ON THE PLANTVILLAGE DATASET

Methods	Datasets	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)	AUC (%)
DNN	PlantVillage	89.22	89.07	88.00	88.53	85.10
ViT		86.85	86.05	86.14	86.09	83.69
Swin transformer		87.59	86.59	89.36	87.95	86.74
PVT		88.26	88.17	90.15	89.14	88.36
CNN		90.44	90.69	90.87	90.78	90.74
CNN + SE		94.19	95.94	94.23	95.08	93.36
ReScaleX-CNN		96.62	96.22	97.28	96.75	93.15
DNN	PlantDoc	71.02	70.28	71.08	70.67	72.36
ViT		73.58	72.13	71.56	71.84	73.25
Swin transformer		74.02	73.15	72.31	72.72	74.06
PVT		75.28	73.89	72.69	73.28	74.87
CNN		76.85	74.25	73.25	73.74	75.14
CNN + SE		78.26	75.69	73.78	74.72	75.89
ReScaleX-CNN		80.21	75.96	74.08	75.00	75.95

TABLE III
PERFORMANCE ANALYSIS OF DIFFERENT CLASSIFIERS WITH FEATURE SELECTION

Methods	Datasets	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)	AUC (%)
DNN	PlantVillage	94.05	94.24	93.14	93.68	90.57
ViT		92.93	92.92	90.71	91.80	91.21
Swin transformer		93.87	93.14	92.48	92.80	91.85
PVT		95.89	94.26	94.28	94.26	92.30
CNN		96.16	95.61	96.57	96.09	92.88
CNN + SE		97.88	96.16	95.05	95.60	93.87
ReScaleX-CNN		99.92	99.91	99.9	99.90	96.92
DNN	PlantDoc	72.10	71.45	71.25	71.34	78.25
ViT		77.54	76.80	75.36	76.07	80.26
Swin transformer		78.12	77.85	77.05	77.44	80.69
PVT		79.36	78.03	77.89	77.95	81.05
CNN		81.35	80.92	78.05	79.45	82.36
CNN + SE		85.40	84.66	79.63	82.06	82.45
ReScaleX-CNN		89.92	89.15	83.02	85.97	83.06

TABLE IV
PERFORMANCE ANALYSIS OF COMPUTATIONAL COMPLEXITY AND MEMORY CONSUMPTION ON THE PLANTVILLAGE DATASET

Methods	Memory consumption (MB)	Run Time (s)	Parameters (M)	Flops (G)	p-value from paired t-test	ONE-WAY ANOVA	CI (%)
DNN	84.40	15.68	15.2	6.5	0.041	0.3658	92.63-95.47
ViT	39.41	14.66	58.0	15.5	0.039	0.4655	91.37-94.49
Swin transformer	121.58	16.58	55.4	14.8	0.038	0.4128	92.40-95.34
PVT	156.39	18.26	53.0	12.3	0.035	0.3569	94.63-97.15
CNN	61.17	18.68	25.6	8.1	0.036	0.2648	94.95-97.37
CNN + SE	303.06	28.65	72.1	20.5	0.032	0.0556	96.98-98.78
ReScaleX-CNN	293.28	25.62	65.5	17.2	0.024	0.0023	99.54-99.95

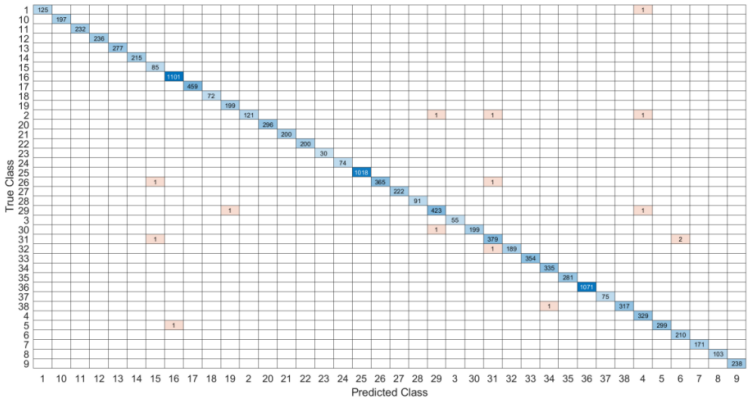


Fig. 8. Confusion matrix for ReScaleX-CNN on the PlantVillage dataset.

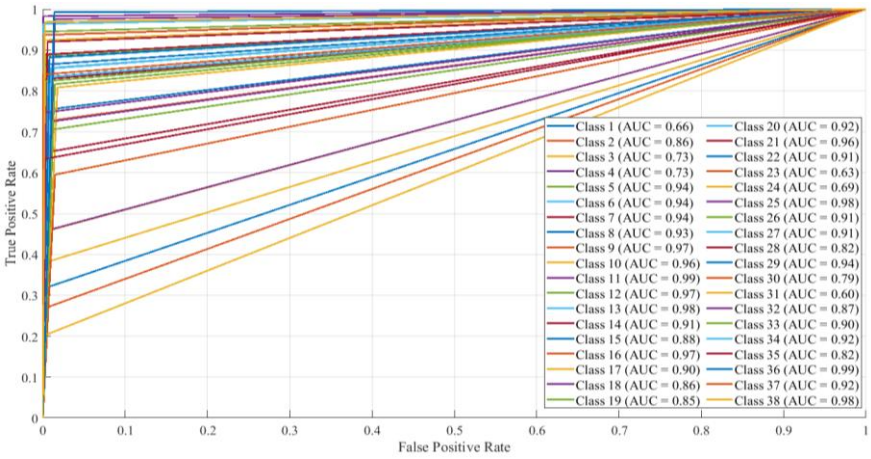


Fig. 9. Analysis of ROC curve for ReScaleX-CNN on PlantVillage dataset.

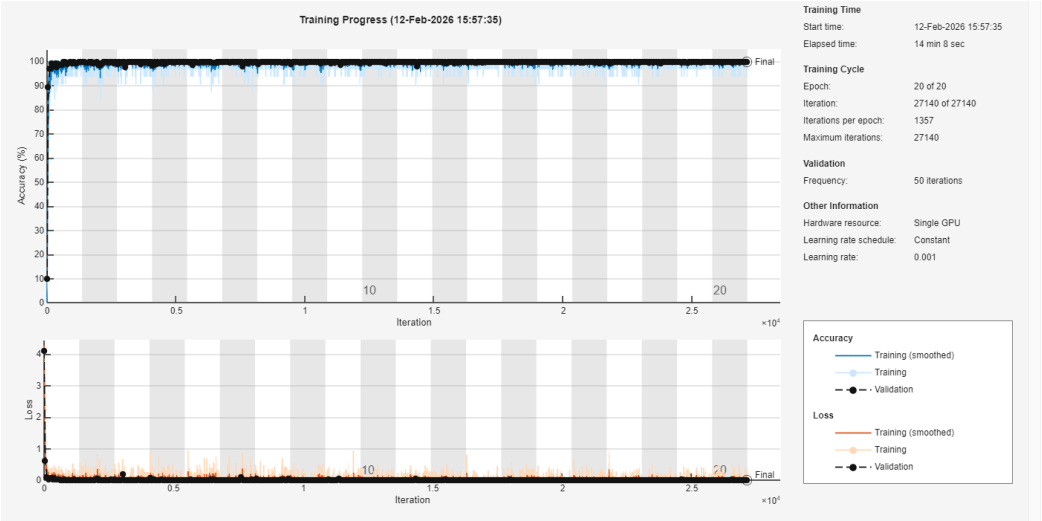


Fig. 10 Analysis of accuracy vs. iteration and loss vs. iteration of the proposed method on the PlantVillage dataset.

TABLE V
IMPACT OF WES-HOA FEATURE SELECTION ON PLANTVILLAGE DATASET

Methods	Feature count	Training time (s)	Inference time (s)
Without feature selection	1280	28.5	0.0024
With feature selection	1110	25.62	0.0019

Fig. 8. illustrates confusion matrix for ReScaleX-CNN on PlantVillage, which demonstrates strong and consistent classification performance for 38 different classes. The off-diagonal values show minor misclassification due to similarities among certain disease classes. Fig. 9. presents the Receiver Operating Curve (ROC) for ReScaleX-CNN on PlantVillage. Area Under Curve (AUC) values ranging from 0.85 to 1.00 illustrate strong discriminative ability. Fig. 10. represents the analysis of accuracy vs. iteration and loss vs. iteration of the proposed method on the PlantVillage dataset. Accuracy vs. iteration shows training and validation accuracy that increases rapidly and indicates effective learning. Loss vs. iteration decrease sharply and remain nearby to zero with less fluctuations. Table VI presents a analysis of ablation study based on different reduction ratios. Compared to other ratios, m=4, 8, and 16 achieves high accuracy due to effective balance of feature compression. Fig. 11. demonstrates analysis of convergence graph for WES-HOA. Since the fitness function employed is MSE, a lower value represents better performance in decreasing error during feature selection.

TABLE VI
PERFORMANCE EVALUATION OF ABLATION STUDY BASED ON DIFFERENT REDUCTION RATIOS ON PLANTVILLAGE

Reduction ratio m	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)	AUC (%)
4, 4, 8	97.69	96.96	96.19	96.58	94.10
8, 4, 8	95.56	94.33	95.51	94.92	92.10
4, 4, 16	96.46	95.44	96.45	95.94	93.30
16, 4, 16	97.49	95.72	96.93	96.32	94.87
4, 8, 16	99.92	99.91	99.90	99.90	96.92

Table VII provides ablation study analysis to evaluate the impact of individual components for proposed method on PlantVillage. The proposed method achieves high accuracy due to minimizing irrelevant information and rescaling feature maps, which enhance classification accuracy. Table VIII determines analysis of cross-dataset validation where model is trained on PlantVillage and tested on PlantDoc datasets. The PlantVillage dataset is used for training due to containing a large number of images, whereas a smaller number of images is used for testing. ReScaleX-CNN obtains better performance that enables more robust and discriminative feature learning. Table IX demonstrates analysis of multi-run performance variance, where seed 456 obtains high performance among all seeds due to random seed initialization influencing weight initialization. Table X demonstrates micro- and macro-averaging performance of proposed method. These results represent consistent high values showing model's robustness.

TABLE VII
ABLATION STUDY ANALYSIS TO EVALUATE THE IMPACT OF INDIVIDUAL COMPONENTS FOR THE PROPOSED METHOD ON THE PLANTVILLAGE DATASET

Methods	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)	AUC (%)
CNN	96.16	95.61	96.57	96.09	92.88
CNN+WES-HOA	97.80	97.42	97.93	96.52	92.88
SE Attention	97.30	96.80	97.61	96.24	93.36
SE+ReScaleX	98.61	98.32	98.46	97.92	93.15
WES-HOA+SE	98.25	97.82	98.15	97.37	93.87
WES-HOA+ReScaleX	99.21	99.02	99.15	99.24	96.12
WES-HOA+ReScaleX-CNN	99.92	99.91	99.90	99.90	96.92

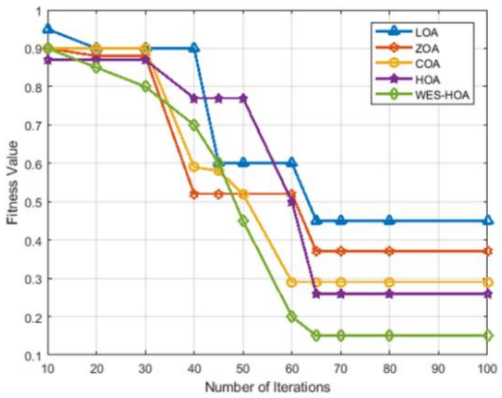


Fig. 11. Analysis of convergence graph on PlantVillage dataset.

A.1. Error Analysis

Type I (Healthy): On PlantVillage, 85% are correctly classified as healthy, whereas 15% are incorrectly predicted as diseased, which denotes a type I error (false alarm). This misclassification occurs because of minor color variation.

Type II (Diseased): For diseased leaves, 70% are identified correctly, whereas 30% are misclassified as healthy, which represents a type II error missed case. These missed predictions arise from subtle disease symptoms and low contrast between healthy and infected regions.

TABLE VIII
ANALYSIS OF CROSS-DATASET VALIDATION WITH TRAINED ON PLANTVILLAGE AND TESTED ON PLANTDOC DATASETS

Methods	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)	AUC (%)
DNN	64.22	63.85	65.23	64.53	71.32
ViT	69.15	68.30	67.12	67.70	75.12
Swin transformer	71.40	70.22	72.15	71.17	76.25
PVT	73.85	72.10	73.65	72.86	77.69
CNN	75.60	74.92	75.12	75.01	80.23
CNN + SE	79.12	78.44	75.89	77.14	81.02
ReScaleX-CNN	82.28	80.29	80.17	80.22	81.06

TABLE IX
ANALYSIS OF MULTI-RUN PERFORMANCE VARIANCE

Random seeds	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)	AUC (%)
42	99.87	99.89	99.88	99.88	95.90
123	99.90	99.89	99.89	99.89	94.94
456	99.92	99.91	99.90	99.90	96.23
789	99.90	99.88	99.89	99.88	95.88
2024	99.89	99.87	99.87	99.87	95.96

TABLE X
PERFORMANCE OF MICRO-AND-MACRO AVERAGING

Metrics	Averaging	Value (%)
Precision	Micro	99.92
	Macro	99.91
Recall	Micro	99.92
	Macro	99.90
F1-Score	Micro	99.92
	Macro	99.90
AUC	Micro	0.999
	Macro	0.997

B. *Comparative Analysis*

Table XI presents a comparative analysis of existing methods on PlantVillage. WES-HOA-based ReScaleX-CNN obtains a high accuracy of 99.92% on PlantVillage dataset by effectively choosing the most appropriate features and enhancing significant spatial channel information.

TABLE XI
COMPARATIVE ANALYSIS OF EXISTING METHODS ON PLANTVILLAGE DATASET

Methods	Accuracy (%)	F1-score (%)	Recall (%)	Precision (%)
PLDC [19]	99.91	99.84	99.81	99.87
PDDNet-LAE [20]	97.79	97.07	N/A	N/A
DeepPlantNet [21]	97.89	91.43	91.5	91.37
SE-SK-CapResNet [22]	98.58	98.58	98.58	98.58
ViT-Base [23]	99.17	N/A	N/A	N/A
Ensemble model [24]	99.65	N/A	N/A	N/A
Proposed WES-HOA-based ReScaleX-CNN	99.92	99.90	99.90	99.91

C. Discussion

The experimental analysis shows that proposed method obtains better performance due to effectively selecting most appropriate features for classification using WES-HOA. ReScaleX-CNN captures fine-grained disease patterns with enhanced sensitivity. Extensive analysis such as ablation study, k-fold validation, ROC curve, and confusion matrix exhibit superior performance which supports reliability of reported 99.92% accuracy. The integration of feature selection and multi-scale recalibration accurately distinguish complex leaf variation and minimize redundancy, leading to near-perfect accuracy of 99.92% on PlantVillage. Potential overfitting is minimized by WES-HOA by removing redundant features. The selected hyperparameters represented in Table 3 are determined through systematic tuning to achieve stable convergence and improved classification performance. The accuracy gap between PlantVillage dataset (99.92%) and PlantDoc dataset (82.28%) represents impact of domain shift when models trained under controlled laboratory conditions are applied to real-world agricultural environments. PlantVillage involves leaf images captured with uniform backgrounds and stable illumination whereas PlantDoc has field images characterized by intricate backgrounds, occlusions, varying lighting conditions, and multiple leaf instances. Although proposed method effectively reduces feature redundancy via feature selection and regularization mechanism like dropout, 18% performance reduction is observed in cross-dataset evaluation. This reduction represents that model trained on homogeneous backgrounds and stable lighting conditions of PlantVillage struggle with background clutter, varying illumination, and occlusion inherent in natural field environment. These factors create significant visual noise and intra-class variability that make disease recognition more challenging and representing limitations of training primarily on controlled datasets.

V. Conclusion

The WES-HOA is proposed for feature selection, and ReScaleX-CNN for accurate classification of plant leaf diseases. WES-HOA selects the most appropriate features by guiding search agents toward more promising solution regions, leading to lower complexity. ReScaleX-CNN includes a multi-scale squeeze and excitation block, which has recalibrated features with diverse spatial granularities by capturing both fine-grained and large-scale disease patterns. Experimental results indicates that proposed method achieved a superior accuracy of 99.92% compared with DeepPlantNet. An ablation study has confirmed that proposed method provides a highly efficient and reliable solution for plant leaf disease classification to support precise agriculture. In future, an advanced attention mechanism with DL will be considered to enhance model performance. To improve applicability of model in real-world agricultural environment, data augmentation methods like rotation, scaling, flipping, contrast adjustment, noise injection and background variation are used in future. These transformations learn from different visual patterns and environmental variations which improve robustness and cross-dataset generalization.

Notation Description

Symbol	Description
x_i	pixel intensity value of the image
min\(x\) and max\(x\)	minimum and maximum values of the pixel intensity
x_{norm}	normalized pixel intensity
$x_{i,j}$	position of i^{th} the candidate solution
X	population of candidate solution
r	random value (0,1)
ub_j and lb_j	upper and lower bounds of j^{th} decision variables
x_{ij}^{M}	male hippopotamus position
y_1	random range [0,1]
D	position of the dominant hippopotamus
I_1 and I_2	random integer numbers selected from the population
r_1 to r_6	random vector
MG_1	best mean position of hippopotamus population group
MG_i	best mean position of hippo i^s group
D	desired position
Γ	exploration decay coefficient
T	random threshold value uniformly distributed from [0,1]
h_1	control function that regulates the exploration stage
h_2	scaling function controlling local exploitation
F_i	objective function value
FB	feedback-based updated position of the hippopotamus
x_{ij}^{HR}	hippopotamus posture while facing predators
R_L	position changes of the predator
t	number of iterations
x_{ij}^{HE}	searches for a hippopotamus location to determine the nearest safe location
s_1	stronger local search abilities
α	weight coefficient controlling the significance of MSE in fitness calculation
β	weight coefficient controlling the significance of the feature subset size penalty
d	total number of features in dataset
X_i^{P2}	new position computed for the i^{th} wombat based on the escape stage
r_{ij}	random number from the interval [0,1]
F_i^{P2}	fitness value of escape updated position
P	p^{th} layer
k_{ij}	convolutional kernel
N_j	bias
M_j	set of input maps
X_c	c^{th} input channel
H and W	height and width of the input feature map
z_c	output of the GAP for c^{th} channel
γ	learnable parameters

Acknowledgement

All authors have equally contributed to the paper. All authors declare no conflict of interest.

References

[1] A. Tabbakh, and S.S. Barpanda, "A deep features extraction model based on the transfer learning model and vision transformer 'TLMViT' for plant disease classification," IEEE Access, vol. 11, pp. 45377-45392, May 2023.

[2] N. H. Shabrina, S. Indarti, R. Maharani, D.A. Kristiyanti, and N. Prastomo, "A novel dataset of potato leaf disease in uncontrolled environment," Data Brief, vol. 52, Art. no. 109955, Feb. 2024.

[3] E. Moupojou, A. Tagne, F. Retraint, A. Tadonkemwa, D. Wilfried, H. Tapamo, and M. Nkenlifack, "FieldPlant: A dataset of field plant images for plant disease detection and classification with deep learning," IEEE Access, vol. 11, pp. 35398-35410, Mar. 2023.

[4] K. Indira, and H. Mallika, "Classification of plant leaf disease using deep learning," J. Inst. Eng. India Ser. B, vol. 105, no. 3, pp. 609-620, Feb.

2024.
[5] E.A. Aldakheel, M. Zakariah, and A.H. Alabdalall, "Detection and identification of plant leaf diseases using YOLOv4," Front. Plant Sci., vol. 15, Art. no. 1355941, Apr. 2024.
[6] H. Alaeddine, and M. Jihene, "Plant leaf disease classification using Wide Residual Networks," Multimed. Tools Appl., vol. 82, no. 26, pp. 40953-40965, Apr. 2023.
[7] F. Naveed, A. Masih, J. Mahmood, M. Ahmed, A. Ali, A. Saddiqa, M.S.H. Abdulnabi, and E. Agbozo, "Sustainable AI for plant disease classification using ResNet18 in few shot learning," Array, vol. 26, Art. no. 100395, Jul. 2025.
[8] A. H. Ali, A. Youssef, M. Abdelal, and M.A. Raja, "An ensemble of deep learning architectures for accurate plant disease classification," Ecol. Inform., vol. 81, Art. no. 102618, Jul. 2024.
[9] R. S. R. Singh, and R.K. Sanodiya, "Zero shot transfer learning framework for plant leaf disease classification," IEEE Access, vol. 11, pp. 143861-143880, Dec. 2023.
[10] S. Srinivasan, S.M. Prabin, S.K. Mathivanan, H. Rajadurai, S. Kulandaivelu, and M.A. Shah, "Sugarcane leaf disease classification using deep neural network approach," BMC Plant Biol., vol. 25, no. 1, Art. no. 282, Mar. 2025.
[11] R. Polly, and E.A. Devi, "Semantic segmentation for plant leaf disease classification and damage detection: A deep learning approach," Smart Agric. Technol., vol. 9, Art. no. 100526, Dec. 2024.
[12] W. Macdonald, Y.A. Sari, and M. Pahlevani, "Grow light smart monitoring system leveraging lightweight deep learning for plant disease classification," Artif. Intell. Agric., vol. 12, pp. 44-56, Jun. 2024.
[13] M. Tiwari, H. Kumar, N. Prakash, S. Kumar, R. Neware, S. Tripathi, and R. Agarwal, "Tomato disease detection using vision transformer with residual L1 norm attention and deep neural networks," Int. J. Intell. Eng. Syst., vol. 17, no. 1, pp. 679-688, 2024.
[14] I. Bouacida, B. Farou, L. Djakhdjakha, H. Seridi, and M. Kurulay, "Innovative deep learning approach for cross crop plant disease detection: A generalized method for identifying unhealthy leaves," Inf. Process. Agric., vol. 12, no. 1, pp. 54-67, Mar. 2025.
[15] D. Raja, and M. Karthikeyan, "Red Deer Optimization with deep learning enabled agricultural plant disease detection and classification model," Int. J. Intell. Eng. Syst., vol. 16, no. 5, pp. 21-30, 2023.
[16] K. M. Hosny, W.M. El Hady, F.M. Samy, E. Vrochidou, and G.A. Papakostas, "Multi class classification of plant leaf diseases using feature fusion of deep convolutional neural network and local binary pattern," IEEE Access, vol. 11, pp. 62307-62317, Jun. 2023.
[17] X. Ma, W. Chen, and Y. Xu: ERCP Net, "A channel extension residual structure and adaptive channel attention mechanism for plant leaf disease classification network," Sci. Rep., vol.14, no.1, Art. no. 4221, Feb. 2024.
[18] A. R. Ala'a, "Potato leaves disease classification based on generalized Jones polynomials image features," MethodsX, vol. 14, Art. no. 103421, Jun. 2025.
[19] J.G. Kotwal, R. Kashyap, and P.M. Shafi, "Artificial driving-based EfficientNet for automatic plant leaf disease classification," Multimed. Tools Appl., vol. 83, no. 13, pp. 38209-38240, Oct. 2024.
[20] W. Shafik, A. Tufail, C. De Silva Liyanage, and R.A.A.H.M. Apong, "Using transfer learning-based plant disease classification and detection for sustainable agriculture," BMC Plant Biol., vol. 24, no. 1, Art. no. 136, Feb. 2024.
[21] N. Ullah, J.A. Khan, S. Almakdi, M.S. Alshehri, M. Al Qathrady, N. El Rashidy, S. El Sappagh, and F. Ali, "An effective approach for plant leaf diseases classification based on a novel DeepPlantNet deep learning model," Front. Plant Sci., vol. 14, Art. no. 1212747, Oct. 2023.
[22] X. Zhang, Y. Mao, Q. Yang, and X. Zhang, "A plant leaf disease image classification method integrating capsule network and residual network," IEEE Access, vol. 12, pp. 44573-44585, Mar. 2024.
[23] H. A. Shehu, A. Ackley, M. Marvellous, and O.E. Eteng, "Early detection of tomato leaf diseases using transformers and transfer learning," Eur. J. Agron., vol. 168, Art. no. 127625, Jul. 2025.
[24] S. N. Nobel, M. Afroj, M.M. Kabir, and M.F. Mridha, "Development of a cutting edge ensemble pipeline for rapid and accurate diagnosis of plant leaf diseases," Artif. Intell. Agric., vol. 14, pp. 56-72, Dec. 2024.
[25] PlantVillage, "PlantVillage dataset," Kaggle, 2025. [Online]. Available: https://www.kaggle.com. [Accessed: 24-Jun-2025].
[26] PlantDoc, "PlantDoc dataset," https://github.com/pratikkayal/PlantDoc-Dataset [Accessed: 22-Dec-2025].
[27] M.H. Amiri, M.N. Hashjin, M. Montazeri, S. Mirjalili, and N. Khodadadi, "Hippopotamus optimization algorithm: a novel nature inspired optimization algorithm," Sci. Rep., vol. 14, no. 1, Art. no. 5032, Feb. 2024.
[28] S. Sah, B. Surendiran, R. Dhanalakshmi, and M. Yamin, "Covid-19 cases prediction using the SARIMAX model by tuning hyperparameters through the grid search cross-validation approach," Expert Systems, vol. 40, no. 5, Art. no. e13086, 2023.



Spoorthi Pothaganahalli Akkalappa is currently working as an assistant professor in the Department of ECE, Dr. Ambedkar Institute of Technology, Bengaluru, Karnataka. She obtained her B.E. in Electronics & Communication Engineering from Dr. Ambedkar Institute of Technology, Bengaluru, in July 2013, and her M.Tech. degree in Digital Communication from M. S. Ramaiah Institute of Technology, Bengaluru, in July 2015, and is now pursuing a Ph.D. in the Department of ECE, Dr. Ambedkar Institute of Technology, Bengaluru, Karnataka.



Shivaputra obtained his B.E. in Electronics & Communication Engineering from Visvesvaraya College of Engineering (under Bangalore University) in 2002. He earned his M.Tech in VLSI Design & Embedded Systems from BMS College of Engineering (under Visvesvaraya Technological University) in 2009 and completed his Ph.D. at Jain University, Bangalore, in 2017. He serves as a reviewer for The Journal of Supercomputing (Springer) and several other international journals of high repute. He has published more than 25 research papers in respected international journals. Under his guidance, 2 Ph.D. candidates and 10 M.Tech students have been awarded their degrees. He also works as an editorial board member for both international and national journals.



Meenakshi L. Rathod is currently working as an assistant professor in the Department of ECE, Dr. Ambedkar Institute of Technology, Bengaluru, Karnataka. She obtained her B.E. in Electronics & Communication Engineering from P.D.A. College of Engineering, Gulbarga, in Feb. 2002; an M.Tech. degree in VLSI Design & Embedded System from UTL Technological Ltd., VTU extension center, under Visvesvaraya Technological University in 2009; and a Ph.D. under Visvesvaraya Technological University, Belagavi. She has published more than 20 research papers in the journals of international repute and has four patents filed. Her research interests include VLSI and communications.



Dinesha Puttaraje Gowda received his Ph.D. degree from the University of Mysore, India, in 2014. He is currently serving as a professor in the Department of Electronics and Communication Engineering at Dayananda Sagar College of Engineering, Bengaluru, India. His research interests include VLSI design and nanotechnology. Dinesha P has published multiple research articles in reputed international journals and conferences in the fields of digital system design with FPGA, antenna design, and nanotechnology. He is a member of numerous professional organizations, including the Institution of Electronics and Telecommunications Engineers (IETE-Fellow) and the Indian Society for Technical Education.



Smitha Sasi is working as an associate professor in the Department of Electronics and Telecommunication Engineering, Dayananda Sagar College of Engineering, has about 17 years of teaching experience. She received her B.Tech degree in Information Technology from Calicut University and her M.Tech degree in Digital Communication and Networking with rank from Visvesvaraya Technological University, Belagavi, Karnataka. She received a doctoral degree in cryptography from Visvesvaraya Technological University. She has published research papers in refereed international journals and research papers in the proceedings of various international conferences. She received funds from various funding agencies in the cryptography and security domain. She has authored various book chapters in the area of cryptography.