



OPEN Plant disease detection using a hybrid dilated CNN with attention mechanisms and optimized mask RCNN segmentation

Kalicharan Sahu^{1✉}, Shivam Tiwari², Manoj Kumar Singh³, Jankisharan Pahareeya⁴, Harish Kumar Shakya^{5✉}, Gyanendra Kumar⁶ & Shitharth Selvarajan⁷

In accordance with human life, agriculture has main role in it, and in addition to that most people are involved in some kind of agricultural activity either in a direct or indirect manner. Moreover, the agricultural sectors acquired a major role in supplying better quality food and thus made the greatest attribution to the growth of populations and economics. But, the disease over the crop has influenced the growth of the corresponding species and thus requires an earlier diagnosis of plant disease by utilizing the most adequate and automatic detection approach for improving the quality of the production of food as well as to reduce the loss in economic. But, there are no techniques in the conventional system for identifying the disease in diverse crops in the agricultural environment. In modern times, deep learning approaches have acquired tremendous enhancement in the identification of image categorization as well as the object detection system. For precise detection of plant disease, an improved classification model is developed. Initially, from the standard publicly available database, the images of the plants are aggregated. The gathered images are segmented using Dilated, Adaptive, and Attention-based Mask Recurrent Convolutional Neural Networks (DAA-MRCNN). Then, it is fed into a hybrid classification phase, where the new model namely Dilated, Adaptive, and Attention-based Multiscale DenseNet termed as (DAA-MDeNet) for classification. The classifier performance is improved by optimizing the parameter in Mask RCNN and Multiscale DenseNet using the hybrid optimization algorithm named African Vulture and Lemur Optimizer (AVLO). When compared with the other model, a superior performance is shown in the proposed model.

Keywords Plant disease detection, Dilation and attention mechanism, Mask recurrent convolutional neural network, Multiscale DenseNet, African vulture and lemur optimizer

In general, the precision agricultural research on identifying diseases over plant leaves by means of image analysis has been regarded as a critical area¹. Some of the conventional techniques for determining the severity of the plant disease have been depending on the examination of the plant tissues². Moreover, the management as well as the cultivation of the expert system has been widely utilized become of the widespread adoption of digital techniques and thus it has maximized the capacity of production³. Moreover, the disease and pest retrieving as well as the description characteristics over the expert's system basically based expert's exposit, and thus resulted in lower efficiency and maximized costs⁴. In addition to that, the advancement over artificial intelligence as well as image processing techniques has attained the opportunity for extending the research over the agricultural sector⁵. Further, the deep learning techniques are defined as a kind of machine learning approach and thus, it has been regarded as active research and also successfully deployed to diverse fields⁶. Thus, it has also been employed in various sectors like communications and networking, the automotive industry,

¹Department of Computer Science and Engineering, Sharda University, Greater Noida, Uttar Pradesh 201310, India. ²Department of Computer Science & Engineering, Sharda University, Greater Noida, Uttar Pradesh 201310, India. ³School of Computer and Systems Sciences, Jawaharlal Nehru University (JNU), New Delhi 110067, India. ⁴Rustamji Institute of Technology, Tekanpur, BSF Academy, Tekanpur, Gwalior, Madhya Pradesh 475005, India. ⁵Department of Artificial Intelligence and Machine Learning, Manipal University Jaipur, Jaipur, Rajasthan 303007, India. ⁶Department of IoT and Intelligent Systems, Manipal University Jaipur, Jaipur, Rajasthan 303007, India. ⁷School of Built Environment, Engineering and Computing, Leeds Beckett University, Leeds LS6 3HF, UK. ✉email: kalicharan8u@gmail.com; harish.shakya@jaipur.manipal.edu

the automotive industry, business, agriculture, and so on has utilized both image categorization and object detection techniques⁷. In considering sector, conventional techniques for identifying the disease in the plant have acquired experts to perform the process of visual inspection and also the more in-depth detection in labs as time-consuming process⁸.

More commonly, the plant disease is diagnosed depending on the symptoms that occur through diverse regions of the plants like as in pulp, stem, and leaf^{9,10}. Moreover, the expert's knowledge has been considered necessary for diagnosing the disease accurately in the plant leaf¹¹. In addition to that, some of the systems have been extended to the correct detection of plant disease through recommending possible remedies and then it has served as advertisements through diverse agricultural product vendors¹². Thus, the crop disease diagnosis depending on the visual symptoms may vary at diverse parts, where the farmer takes a picture of the plant part and then, the mobile disease identification approaches that have the potential to detect as well as to label the disease¹³. It has aided to prevent crop losses by reducing the phases in the usual detection approaches¹⁴. But, it is very crucial to attain the rural areas, where the crop has been generated through smallholder farmers, along with the utilization of some improved technologies; it is possible to attain the diagnosis at an expert level¹⁵.

Here, diverse AI techniques identify as well as categorize the disease over the plant that has been implemented. In general, various approaches like CNN, and so on are used¹⁶. Thus, this approach has been utilized along with various pre-processing approaches for maximizing the feature extraction process. Consequently, the Deep Learning (DL) algorithm model is used for extracting the features through the images and then used the features for performing regression or classification processes based on the requirements. In addition to that, it is depending on the similarity in classifying the data and it has been further considered as the unlabeled objects that are categorized by utilizing the neighbouring labelled objects. But, some of the conventional techniques have certain restrictions like overlapping and over-fitting issues of data¹⁷. In general, DL techniques have the potential to retrieve the features in a specified manner and then fused the features in order to offer better-classified outcomes¹⁸. On considering the earlier phases of the research depending on the process of detection, it has intended to remain the categorization tasks also. In addition to that, computers are utilized to detect the disease by means of DL approaches. Further, a fine-tuned deep learning model is implemented for identification process. In order to further improve the identification process for diseased plants, a new model has been implemented.

Certain attributions that are highlighted in the given model are shown as follows.

- To build the new precise networking model for detecting the disease over the plant that has aided for the earlier identification and prevention to secure the plant in a significant manner.
- To perform the segmentation process by utilizing the novel architectural model termed DAA-MRCNN, where the AVLO model is used for truning the parameters along with the objective function for maximizing the dice and Jaccard co-efficient.
- To perform the classification process, the DAA-MDeNet is utilized to attain the final classified outcomes, where the AVLO model is used for truning the parameters along with the objective function for maximizing the rate of accuracy.
- To develop the new algorithm that is used for optimization process for both the segmentation and classification phase is termed AVLO, which is designed with the new formulation to overcome the limitations in conventional models.
- To experimentally validate the entire performance of the given model using several measures to show the betterment over others.

Other sections that are followed in the newly designed model are given as follows. Tier II offers Literature survey, illustration of detection process: adaptive segmentation and classification model in Tier III, DAA-MRCNN for segmentation in Tier IV, identifying the plant disease using multi-scale DenseNet with AVLO in Tier V, results and conclusion are in Tier VI and VII.

Literature survey Related works

In 2023, Moupojou et al.¹⁹ have suggested new DL-dependent techniques that were helped the farmers to identify the crop-related disease in order to neglect the yield delay. This technique has been trained over the publically available datasets and was composed of laboratory images attained under laboratory conditions but faced diverse limitations. Thus, a FieldPlant was recommended in this model that was aggregated through the plantations. In addition to that, the manual annotation of every image over the individual leaves was carried out to assure the quality. Finally, the determination was made for the object detection model and then identified the classification tasks.

In 2022, Saleem et al.²⁰ have presented the model for recognizing the disease in the plant by utilizing the newly aggregated datasets. After detecting the most adequate deep learning techniques, the data augmentation approaches have been validated. Consequently, the impact over the resizers along with the interpolators, and batch normalization has also been carried out. At the last the overall performance enhancement was carried out through empirical observation. Moreover, the robustness of the given model was determined through k-fold cross-validations.

In 2022, Patil et al.²¹ have utilized the model for detecting the disease in the plants due to its spectacular accomplishment. Further, the standard approach has been utilized for neglecting the irrelevant background in the input images by multi-scale features selection. This model implemented the detection of disease in the cardamom plant by utilizing the EfficientNetV2 techniques. In addition to that, a comprehensive set of

determinations has been made to ascertain the ability of the given model and then assimilate it with other techniques like CNN and the Efficient Net model.

In 2023, Hosny et al.²² implemented the new lightweight deep model for attaining better depictions of maximized-level features. Further, the deep features have been fused together along with the conventional standard features in order to attain the local texture information over the plant. Then, this model has been trained as well as then tested over the datasets.

In 2022, Amin et al.²³ have proposed end-to-end DL techniques for detecting the unhealthy as well as the healthy leaves into consideration. This system has utilized two pre-trained CNNs and some other standard techniques for retrieving the deep features through the plant images. In addition to that, the data augmentation approaches have been utilized to add variations to the images to train the model and also ensure the model learns more complex cases. Thus, this model has acquired fewer parameters than the conventional techniques and attained less proceeding power.

In 2023, Vishnoi et al.²⁴ have used the CNN model that included a lesser amount of layers, which may lead to the reduced burden computation process. In addition to that, various augmentation processes like flipping, zoom, scaling, shear, and shift have been deployed for producing the samples and thus maximized the training sets without attaining more images. Even though various conventional techniques were used in the detection process, this model has offered an accurate performance with computational and lower storage. Moreover, the rigorous validation outcomes have revealed that the given model was fit.

In 2021, Zhao et al.²⁵ have provided a detection model by utilizing the Double GAN for offering high-resolution images by detecting the leaf with disease and it was carried out in two different phases. To the standard techniques, the healthy leaf images were utilized as the input for attaining the pre-trained model was performed in the initial phase and consequently, the unhealthy leaves were utilized to pre-train the model. Further, the other standard techniques were utilized to attain the corresponding images to extend the unbalanced datasets. The recognition issues have offered better outcomes when assimilated with original datasets.

In 2021, Ahmad et al.²⁶ have designed techniques for systematically categorizing the symptoms of plant disease by utilizing the CNN model. While coupled with the given techniques, it has assured the industrial applications by minimizing the training times. In addition to that, the transfer learning techniques were used for training even the small datasets that have the potential to transfer the pre-trained weights over the larger datasets. But the negative transfer learning was a common issue in transfer learning. Thus, step-wise transfer learning techniques were recommended that aided the fast convergence.

In 2025, Hassan et al.²⁷ have developed a transfer learning-based deep learning model for the classification of breast cancer in women. Multiple deep learning models are combined in this model to get robust results in the classification process. In 2024, Hassan et al.²⁸ have presented a deep learning model for early detection of black fungus in medical images. This model detects the black fungus more effectively at an earlier stage to prevent the mortality of humans.

In 2025, Hassan et al.²⁹ have suggested a Real-Time Adaptation Framework for dysarthria detection under the resource-constrained scenario. Across the diverse dataset, the enhanced WaveNet is used for capturing the long-term dependencies in the audio signal. The reliability of the clinical application is improved through this model. In 2024, Chouhan et al.³⁰ have proposed an artificial intelligence-based approach for improving the agricultural sector.

In 2019, Chouhan et al.³¹ have developed a Fuzzy Competitive Learning based Counter Propagation Network (FCPN) for the purpose of segmenting the natural scene images. This model has high parallel learning capability and also handles the uncertainty in the segmentation process. In 2024, Sharma et al.³² have proposed Generative Adversarial Networks (GANs) and Vision Transformers (ViTs) for providing a solution to the farming situation. Here, the robust and high-quality dataset is used for constructing a robust model for the agricultural task.

In 2025, Chouhan et al.³³ have developed an artificial intelligence model for soil health and crop health monitoring. The timely and accurate plant disease diagnosis system is offered through this model. In 2020, Mahmood et al.³⁴ have presented a deep learning model for breast cancer detection on multimodalities. In 2025, Rehman et al.³⁵ have proposed the Swin-ViT model for robust kidney carcinoma prognosis. In 2023, Ali et al.³⁶ have presented TESR (Two-stage approach for Enhancement and super-resolution) for improving the resolution of the images in an artificial manner. In 2024, Mahmood et al.³⁷ have developed a deep-learning algorithm for improving the detection and classification of breast cancer. In 2024, Mahmood et al.³⁸ have presented a Depth Double Deep Learning Method of Linear Attention Network (D3LM-LAN) for detecting cognitive impairment at an earlier stage. In³⁹ Multi-Modal Feature Fusion Network for Histopathology (MFF-HistoNet) is proposed for improving the accuracy of breast cancer detection by addressing the multigrading challenges. In⁴⁰, squeeze-and-excitation and dilated dense convolution are proposed for analyzing the intricate brain tissues more accurately.

Problem specifications

In general, there is usually advancement as well as certain limitations over the classical model of disease detection in leaves that are tabulated in Table 1. Deep learning models¹⁹ technique detects the disease in the individual leaves. It has also assured the quality of the entire detection process. But the modelling of a global ensemble, along with the segmentation process, is limited in this model. The deep learning²⁰ method has been embedded into the robotic system for deploying over the disease control approach. It has also been utilized for developing a cost-effective protection system. But, more in-depth validation needs to be explored to strengthen the process. The CNN²¹ process can neglect the complex background from the images. The model needs to be extended for detecting nutritional deficiency. The CNN²² model offered better determination with accurate outcomes for detection. It required less number of parameters. But the application of practical crop disease detection is restricted. CNN²³ has acquired small parameters for retrieving the features as well as integrated the sets of features that offer more robustness to the given model. Detecting the disease by means of a digital imaging

Author [citation]	Methodology	Features	Challenges
Moupojou et al. ¹⁹	Deep learning models	It has the capability to identify the impact on the individual leaves It has also assured the quality of the entire detection process	But the modelling of a global ensemble, along with the segmentation process, is limited in this model
Saleem et al. ²⁰	Deep learning	This method has been embedded into the robotic system for deploying over the disease control approach It has also been utilized for developing a cost-effective protection system	But, more in-depth validation needs to be explored to strengthen the process
Patil et al. ²¹	CNN	The complex background of the images is removed using this model	The model needs to be extended for detecting nutritional deficiency
Hosny et al. ²²	CNN	This model has the ability to offer better determination with accurate outcomes for detection It required less number of parameters	But the application of practical crop disease detection is restricted
Amin et al. ²³	CNN	It has acquired small parameters for retrieving the features as well as integrated the sets of features that offer more robustness to the given model	Detecting the disease by means of a digital imaging process needs to be developed
Vishnoi et al. ²⁴	CNN	To carry out the identification process with the aid of the leaf image It is more consistent as well as reliable	Better image variability is required
Zhao et al. ²⁵	DoubleGAN	It has been effectively utilized in the field of image generation It has also detected unhealthy leaves easily	But, the high-resolution images along with less number of samples are limited
Ahmad et al. ²⁶	Ensemble	This method has played an essential role in enhancing the entire performance It has offered better feasibility	But it needs to be improved for practical applications

Table 1. Certain advancements and their limitations in the plant disease detection model.

process needs to be developed. CNN²⁴ detect the disease over the crops with the aid of the leaf images. It is more consistent as well as reliable. Better image variability is required. Double GAN²⁵ has been effectively utilized in the field of image generation. It has also detected unhealthy leaves easily. But, the high-resolution images, along with a smaller number of samples, are limited. The ensemble²⁶ method has played a significant role in enhancing the entire performance of the detection system. It has offered better feasibility. But it needs to be improved in the practical applications.

Illustration of plant disease detection: adaptive segmentation and classification model

Proposed system of plant disease detection

In general, there is an ancient interaction between future crops as well as the disease that has caused the never-ending competition, and identification for controlling the pest. On considering crop protection, it is a difficult process for validating the resistance of the plants to employ pesticides in a cost-effective manner. But the determination of symptoms and disease severity is essential for addressing. Plant disease may affect the growth of the species, and thus, there is a need for earlier detection. For detecting the disease, various mobile-dependent techniques have been employed. Certain machine learning and deep learning approaches are taken into consideration for diagnosis. For tackling the losses, diverse techniques have been designed to diagnose the diseases. In addition to that, precision agriculture has utilized recent technologies for optimizing the decision-making process. Some of the standard techniques are utilized to offer optimal decisions that lead to the reduction of costs. But, this area still needs to be improved, specifically over the decision-support systems that aid in more useful recommendations. It has also provided more accurate predictions. Moreover, deep learning approaches are utilized for resolving complex issues in a reasonably short amount of time. But, there are limitations as well in the conventional approaches to detecting the disease in the plant, and thus, the new technique is implemented, and it has been represented in the architecture in Fig. 1.

In this research work designed an efficient plant disease classification model. From the standard publicly available database, the images are acquired in the first step. Then, the attained images have been given to the phase of segmentation, that have aided the plant disease detection process to the next level, which has the potential to offer a more accurate and time-saving process. For performing the segmentation process, the newly developed DAA-MRCNN is used. Then, the images are given to the final phase of classification for classifying the diseased plant. In this phase, the newly designed DAA-MDeNet effectively performs the classification process. In both models, the integration of attention impacts precision and reduces false positives in complex field conditions during the segmentation and classification of plant diseases. During the segmentation operation, the pixel-level lesion areas are mainly focused on by the DAA-MRCNN because of the incorporation of the attention module. This may be used for suppressing the irrelevant region of the images and also the lighting artefacts. But in the classification of the plant disease, the lesion texture, edge patterns are focused on by the DAA-MDeNet model via the incorporation of an attention module. Thus, the diagnostic precision and the interpretability of the model is greatly improved through the incorporation of an attention mechanism in the proposed models. Further, by implementing the new AVLO algorithm for optimizing the parameters in both the DAA-MRCNN and DAA-MDeNet models for improve the performance. The determination of the detection process is validated in the final phase.

Plant disease dataset

The images related to performing the detection process are aggregated in this phase, and by using the relevant dataset, the details are given as:

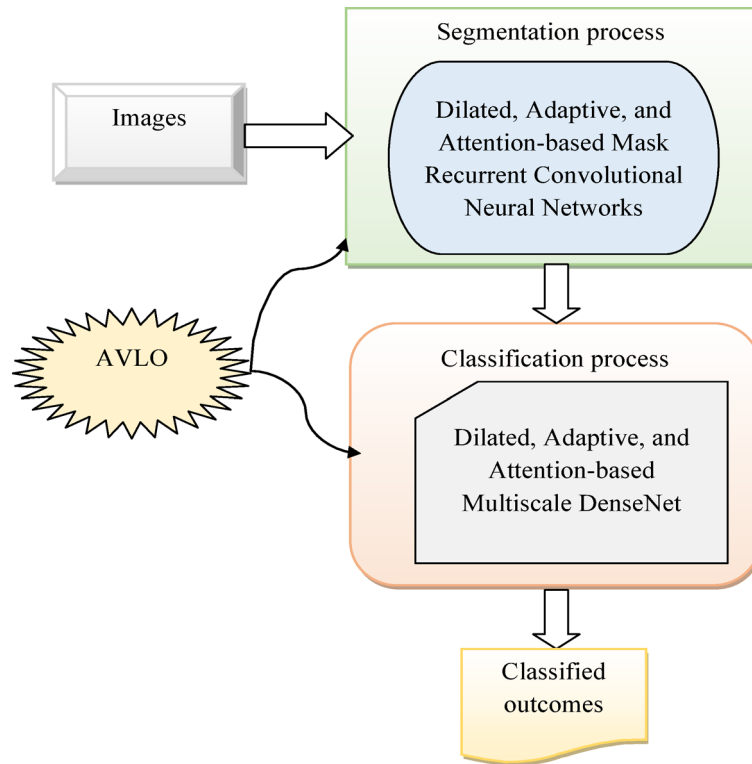


Fig. 1. Architectural modelling of the given detection system for plant disease.

Dataset: This dataset is named as PlantifyDr. It includes a total of 12,500 images in it from 10 different plant types, where the 10 different types are considered as 10 individual datasets. (1) Apple, (2) Cherry, (3) Citrus, (4) Corn, (5) Grape, (6) Peach, (7) Pepper, (8) Potato, (9) Strawberry, and (10) Tomato. It contains a total of 37 as plant diseases. It was collected through “<https://www.kaggle.com/datasets/lavaman151/plantifydr-dataset>”: “Access Date: 2023-08-09”.

Thus, the images are significantly aggregated, and it has been termed as PD_{zz} , $zz = 1, 2, \dots, ZZ$ which denotes the images that are aggregated during the process. Then, the sample images with the disease name are given in Fig. 2.

Novel heuristic algorithm: AVLO

In the newly designed novel hybrid heuristic-based detection process for detecting the disease over the plant, a new model, AVLO, has been implemented depending on its position to tackle some of the limitations of the conventional AVOA and LO models. Because of the complex parameter spaces, achieving peak performance by tuning the parameters of the deep learning model is a difficult task. In this work, the AVLO is developed that simulate the natural activities of the lemur for effectively searching the optimal hyperparameter space. The searching process using the AVLO improves the convergence and accuracy of the model used for both segmentation and the classification operation. The AVLO-based parameter tuning improves the reliability of the detection, assisting the fine-tuning process in both segmentation and classification. The proposed AVLO is developed by hybridizing the LO and AVOA algorithms since they are more effective in the exploitation and exploration phases. The hybridization of both algorithms improves the convergence and also maintains the diversity of the population. Even in the large and non-convex parameter space, the dynamic balance between the exploitation and the exploration is effectively managed by the combination of both algorithms. As compared to the conventional optimization algorithms, faster convergence, better generalization are attained by leveraging the explorative strength of the AVOA along with the exploitative strength of the LO, so it provides high quality solution for the complex optimization issues.

AVOA model only requires the minimal computational complexity and then it maximizes flexibility. It also has the capability to solve the continuous issues that arise in the optimization process. But, it has time-related issues to process the entire phases. Consequently, the LO model has provided essential competition over other models and managed both the parameter and optimal control-related issues. Here, the binary version is limited and degrades its performance.

Thus, the new formulation based on population is implemented; it is expressed in Eq. (1).

$$ps = \text{mean}(ps1, ps2) + \frac{bs}{wf} \quad (1)$$

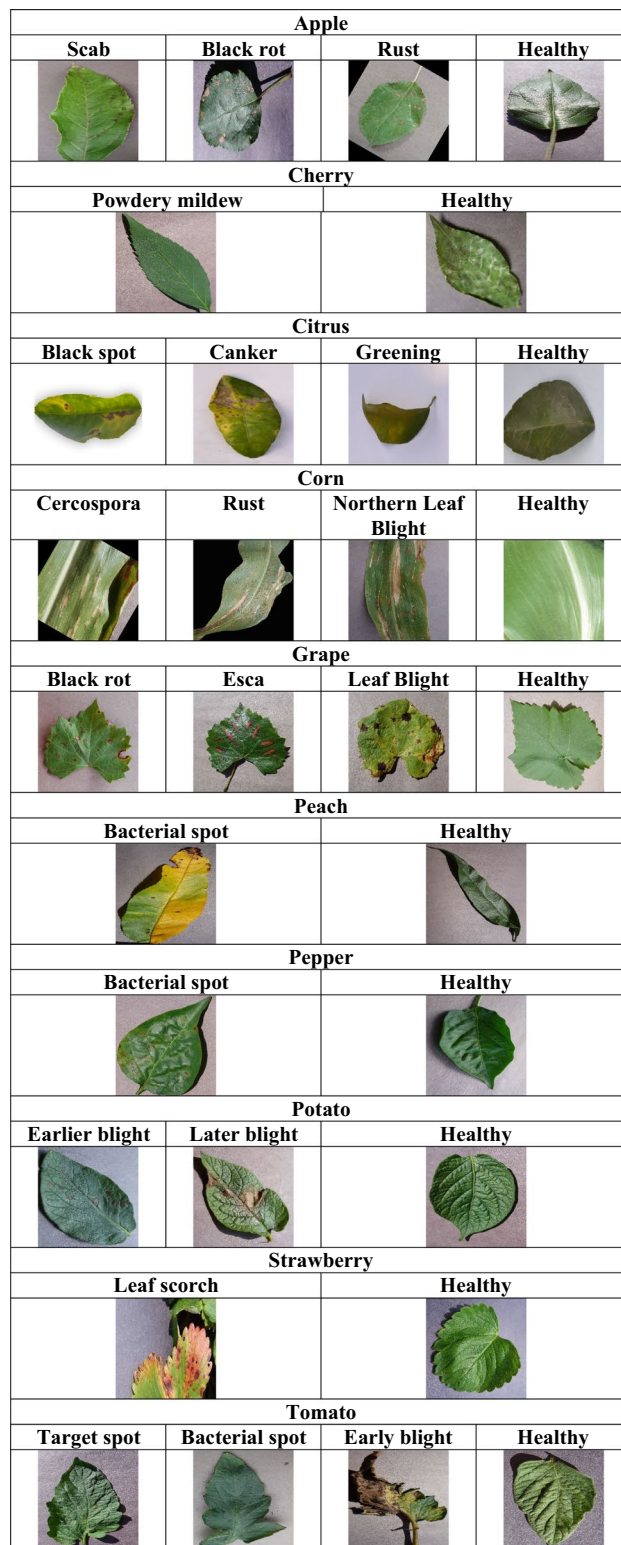


Fig. 2. Some of the sample images aggregated using the given dataset.

Here, ps denotes the new position formulation that is derived for update, $mean$ indicates the mean value, $ps1$ and $ps2$ represents updates on the positions that take place using AVOA and LO accordingly.

AVOA⁴¹: Vultures are divided into two significant groups of hunting birds. In addition to that, the vultures have aided in protecting the stinging as well as infecting carcasses.

Phase one: The fitness as well as the best solution has been selected as the best vultures for all the solutions and other solutions over the second group using Eq. (2). To validate the iteration over the fitness function and

also the population over the vulture is also determined. In addition to that, the probability of choosing the better solution has been derived.

$$A(z) = \begin{cases} bsvu_1 & \text{if } a_b = B_1 \\ bsvu_2 & \text{if } a_b = B_2 \end{cases} \quad (2)$$

Here, a better solution is required to validate the probability of the chosen vultures, where B_1 and B_2 are also measured.

Phase two: It is regarded as the rate of being satiated, and then depicted the behaviour of the model is derived in Eq. (3). The mathematical modelling included in this behaviour is expressed in Eq. (4).

$$c = d \times \left(\sin^* \left(\frac{\pi}{2} \times \frac{it_f}{it_{tn}} \right) + \cos \left(\frac{\pi}{2} \times \frac{it_f}{it_{tn}} \right) - 1 \right) \quad (3)$$

$$vs = (2 \times rnd_1 + 1) \times e \times \left(1 - \frac{it_f}{it_{tn}} \right) + c \quad (4)$$

Here, the term rnd_1 depicts the random value among $[0, 1]$, e and d is given as a random number among $[-1, 1]$ and $[-2, 2]$, it_f which is the current iteration. vs is the vultures and it_{tn} indicated as the total number of iterations.

Exploration: the vulture usually acquires various random areas for analysis, in which it is based on two different strategies, and then the parameters are denoted as C_1 that is used to elect the strategy. On considering the number that is maximized over the parameter C_1 , where Eq. (5) is used. Subsequently, when rnd_{C_1} is minimized the parameters C_1 , Eq. (6) is used.

$$C(z+1) = \begin{cases} eq.(6) & \text{if } C_1 \geq rnd_{C_1} \\ eq.(8) & \text{if } C_1 < rnd_{C_1} \end{cases} \quad (5)$$

$$C(z+1) = A(z) - D(z) \times sv \quad (6)$$

$$D(z) = |E \times A(z) - C(z)| \quad (7)$$

The best cultures in the surrounding area are the vulture's random search for prey, which $C(z+1)$ is regarded as the vulture location vectors. In addition to that, the rate of the vulture is sv and it is derived in Eq. (8). Further, the term E represents the coefficient vector that has enhanced the random motion. Moreover, $C(z)$ indicates the position of the current vector of the vulture.

$$C(z+1) = A(z) - sv + rnd_2 \times ((u_{bo} - l_{bo}) \times rnd_3 + l_{bo}) \quad (8)$$

Here, l_{bo} and u_{bo} are given as lower and upper bounds, and rnd_3 the coefficient of random nature is enhanced. Then, the best vultures are selected using Eq. (6), and it is termed as $A(z)$ and also the vulture satiation rate attained using Eq. (7).

Exploitation: Then, the effectiveness of the AVOA has been defined in the phase. This phase has been carried out when the value $|sv|$ becomes less than 1. Moreover, the parameters C_2 and C_3 are used to select the strategy.

Further, the rotating flight strategy is made while the random number is minimized within the C_2 parameters. It is expressed in Eq. (9).

$$C(z+1) = \begin{cases} eq.(6) & \text{if } C_2 \geq rnd_{C_2} \\ eq.(8) & \text{if } C_2 < rnd_{C_2} \end{cases} \quad (9)$$

LO⁴²: The lemurs are defined as the classification of primates in general, in which it includes all primates. In addition to that, the searching process over the LO algorithm model is classified into two diverse steps: population-based algorithm model, as dance-hup behaviour is involved in exploration, and the leap-up behaviour is included in the exploitation.

Then, the set of lemurs is represented by means of the matrix. Moreover, the population is depicted in the form of a matrix and given in Eq. (10).

$$F = \begin{bmatrix} b_1^1 & b_1^2 & \dots & b_1^{vd} \\ b_2^1 & b_2^2 & \dots & b_2^{vd} \\ \vdots & \vdots & \vdots & \vdots \\ b_{sc}^1 & b_{sc}^2 & \dots & b_{sc}^{vd} \end{bmatrix} \quad (10)$$

Here, sc depicts the candidate solution vd depicts the decision variable, and F provides the set of lemurs.

Then, the decision variable g among the solutions h is expressed as in Eq. (11).

$$b_h^g = rnd(\cdot) \times ((U_b - L_b) + L) \quad (11)$$

$$\forall h \in (1, 2, \dots, mn), \quad \forall g \in (1, 2, \dots, vd)$$

Here, the term $rnd(\cdot)$ is depicted as a distributed random number $(1, 2, \dots, mx_in)$, where it represents an integer number. Then, the variables' lower and upper bounds are depicted as $(U_b - L_b)$.

Moreover, when the fitness value is low and then the decision variables are changed. In addition to that, the lemurs are organised based on the value of fitness, and thus it has aided in improving the performance along with the iteration through considering the global best lemur gbl as well as choosing the best nearest lemur for each of the lemurs bnl .

Further, the decision variable g among the solutions h is validated as the value of iteration by utilizing the two options: a) the value is elected by gbl , and b) the value is selected by means of bnl . It is given in Eq. (12).

$$Z_h^g = \begin{cases} b(h, g) + abs(b(h, g)) - b(bnl, c) * (rnd - 0.5) * 2, rnd < rr \\ b(h, g) + abs(b(h, g)) - b(gbl, c) * (rnd - 0.5) * 2, rnd < rr \end{cases} \quad (12)$$

Here, $b(h, g)$ is depicted as the current lemur, $b(bnl, c)$ is given as the best nearest lemur, rr is denoted as the rate of risk of all lemurs, and rnd is indicated as the random number [0, 1]. Thus, the AVLO model's representation is in Algorithm 1.

Parameters, iterations, populations, and dimensions are included.			
Validate the fitness value			
For $t = 1$ to IT_{MX}			
For $i = 1$ to POP_N			
While $it < IT$ do			
Position one update takes place using AVOA in Eqs. (7) and (9)			
Else			
Position second update takes place using LO in Eqs. (11) and (12)			
Formulate the position-based derivation $ps = mean(ps1, ps2) + \frac{bs}{wf}$			
End while			
End for			
End for			
End			
Return the better solution			

Algorithm 1. AVLO

Then, the flowchart depiction for the AVLO model is given in Fig. 3.

Dilated, adaptive, and attention-based mask RCNN for segmentation to detect the plant disease

Model of mask RCNN

In general, the mask RCNN⁴³ model is regarded as an image detection and segmentation model as well as it is considered as one of the Faster RCNN models. The diseased regions from the images are more effectively captured by the mask RCNN model as compared to the ResNet, EfficientNet, or transformer-based models like Swin Transformer. In addition, the boundary accuracy of the mask RCNN is very high, so it provides feasible results in the task of segmentation. Especially under the small dataset, the noisy and the complex background in the agricultural images are effectively balanced by the mask RCNN. In order to attain the corresponding features map on the starting phase itself, the images get forwarded to the trained FPN and then given to the ResNext101 model, where the ResNext101 model is the kind of CNN model that is used to enhance the rate of accuracy and then lower the hyper-parameter impact without influencing the performance of the given model. Consequently, the fixed Region of Interest (RoI) has been subjected to the RPN over the binary classification as well as the Bounding-Box regression process. In addition to that, there are four essential phrases in the Mask RCNN model.

- Backbone:** In general, it is used to pre-train the parameters as well as to attain the trained model by means of initial parameters. It has the potential to extract the relevant map features through original images. It has adopted the structural model in a fixed form, like DenseNet and so on.
- Feature Pyramid Network (FPN):** In general, it is completely opposite to that of the backbone phrase, and it aims to extract the multi-scale feature maps fully.
- Region Proposal Network (RPN):** The ultimate goal of this phrase is to offer and select the rough identification rectangle. It is followed up by the FPN phrases, and its parameters are propagated. The RoI involved in this phase is regarded as the align step, and that utilizes the bilinear interpolation rather than the rounding operation, which is utilized by means of RoI pooling that can improve the RoI accuracy rate. RPN has diverse scales that are normalized into the same dimension of RoI as 7×7 .

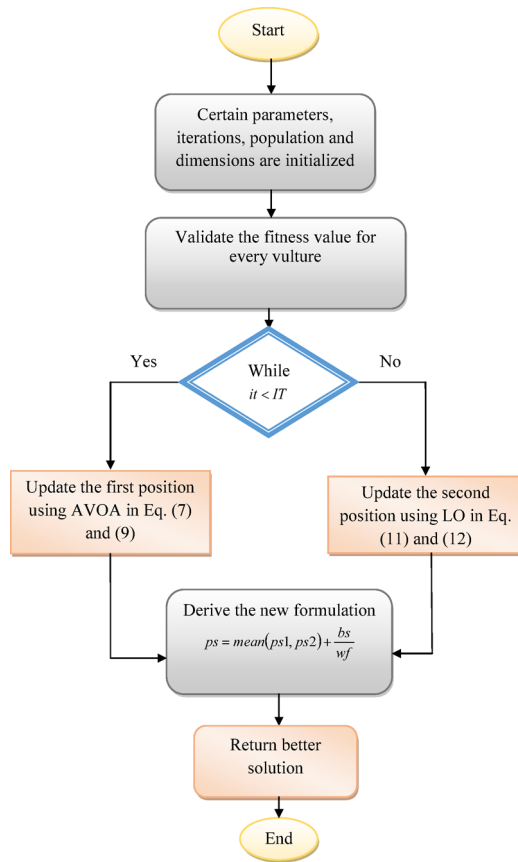


Fig. 3. Flowchart depiction for AVLO model.

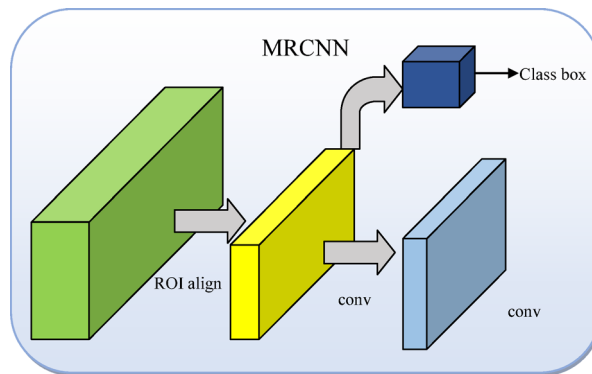


Fig. 4. Basic depiction for mask RCNN model.

- (d) Function Branches: On considering the candidate RoI that is generally refined by means of a RoI alignment as well as involves three functional branches, such as categorization, identification, and segmentation process. Moreover, the segmentation branches are connected to the fully convolutional layer has provided a binary mask. Thus, the basic depiction of the mask RCNN model is given in Fig. 4.

Dilated and attention in mask RCNN

In order to carry out this phase, the initially aggregated images PD_{zz} are subjected as input to this model. For retrieving the most adequate or relevant information through the given image pixels, the process of segmentation has been carried out in general, and it has the potential to improvise the capability of the model without any lack in the entire process. The conventional segmentation approaches provide inaccuracy in the segmentation operation if it faces images with varying backgrounds and overlapping leaf images. So, the DAA-MRCNN is proposed in this work, which is made up of dilated convolution, attention modules, and an adaptive concept for improving the accuracy of the segmentation process. The combination of these modules in the suggested DAA-

MRCNN enables to precise isolation of the diseased areas from the images to get the effective results in plant disease detection.

Dilation: The convolution process is regarded as the process of dilation, in other words, because it is more similar to the process of convolution, but the dilation process involves the pixel shifting to ensure a huge number of areas in the given images. In addition to that, the convolution layer has attained a huge number of receptive fields without impacting the size of feature maps, which has aided in improving the information regarding the features. In general, the convolution has acquired the kernel with holes as well as the dilated rate is regarded as the essential parameter for distinguishing the dilated from normal. The receptive field is improved in the suggested model with the incorporation of the dilation module. The addition of the dilation does not increase the number of parameters while expanding the receptive field in the model. In this work, the fixed dilation rates (1, 2, 4, 8) are used in the suggested model. In order to capture the finer details in the images, dilation rates of 1 and 2 are adopted. Likewise, the global contextual awareness is improved by the higher dilation rates (4, 8). In the different layers of the hybrid models, the combination of dilation rates (2, 4, and 6) is used so that the large and the small contextual cues are more effectively captured using the suggested model. During the training process, the dilation rates are taken as a fixed value.

Attention: The most significant process involved in the mechanism of attention process is to intent on the specified part of the images in order to perform the process of classification adequately. It is detailed in Eq. (13).

$$ATT(y, z) = \sum_{i=1}^Z \sigma_i(y, z_i) E_i \quad (13)$$

Here, y and $\sigma_i(y, z_i)$ are the query task, similarity function among queries, and its respective key as well as E_i , z_i is the keys.

In addition to that, the softmax function involved in the attention mechanism has also been determined and it is expressed in Eq. (14).

$$sf(j_i) = \frac{\exp(j_i)}{\sum_{i'} \exp(j_{i'})} \quad (14)$$

Dilation and attention with mask RCNN: Thus, the dilation in the mask RCNN model has been made by replacing the layer of convolution with dilation to offer better performance. Consequently, by using integrated attention values, attention that it has been guided through a set of input images that are attained.

Thus, the segmented images have been attained as the outcome of this phase, and it is termed as ss^{mrc} and its process is detailed in Fig. 5.

Suggested DAA-MRCNN for segmentation

In the agricultural imagery, the end-to-end classification model trained on the raw images is a computationally efficient and simpler operation. The classifier sometimes produces misclassified results due to the variance in the background and lighting conditions of the raw input images. When the diseases occupy a tiny portion of the images, the detection of plant disease from the entire image frame is a difficult operation. Instead of focusing on the lesion of the images, the correlation of the background elements is mainly focused on by the end-to-end classifier. So, this research adopted the two-stage pipeline that consists of segmentation and classification by

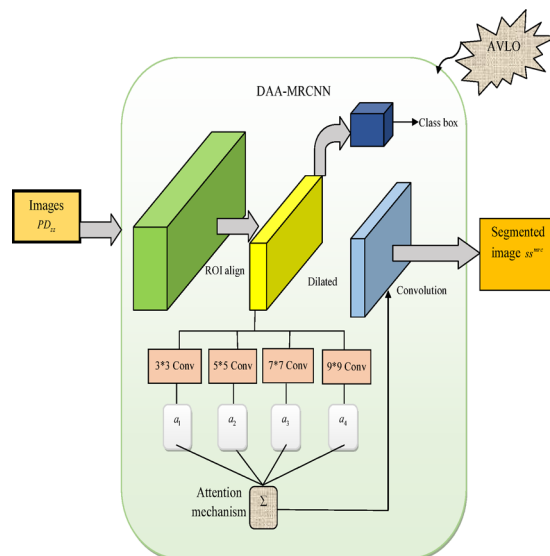


Fig. 5. Depiction of the segmentation process carried out using the DAA-MRCNN model.

addressing the aforementioned issues. To perform the segmentation process, the new heuristic model, termed as DAA-MRCNN, has been implemented. The suggested DAA-MRCNN model removes the irrelevant context effectively by focusing on the diseased region of the images, and this model improves the specificity and the precision by extensively concentrating on the relevant areas of the images. To tackle some of the limitations of the mask RCNN model, a new formulation is derived. In general, this model is simple to train, and it has better flexibility and efficiency. Consequently, it has acquired more time for detection. Therefore, with the aid of the AVLO algorithmic model, certain parameters in MRCNN have been optimized to enrich the performance. It is derived in Eq. (15).

$$of_1 = \arg \min_{\{af_{mrc}, hn_{mrc}, ep_{mrc}\}} \left(\frac{1}{dc + jc} \right) \quad (15)$$

Here, the terms af_{mrc} , hn_{mrc} , ep_{mrc} indicated the activation function among [0, 4], hidden neuron count in [5–255], and the number of epochs in [5 to 55] in MRCNN are tuned using AVLO. Further, the terms dc and jc are in Eq. (16) and (17).

$$dc = \frac{2 \times aa}{(aa + bb) + (aa + bz)} \quad (16)$$

$$jc = \frac{aa}{aa + bb + bz} \quad (17)$$

Here, the terms aa , az , bb , bz are the “true positive, true negative, false positive, and false negatives”.

Identifying the plant disease using dilated, adaptive, and attention-based multiscale densenet with AVLO

Multiscale densenet

Here, the segmented images ss^{mrc} are given to the final classification phase to obtain the classified outcomes.

Multi-scale⁴⁴: Here, both the down-sampling as well as the dense blocks have been acquired along the path of down-sampling and thus made the multi-scale DenseNet model. While alleviating the computational expenses, the down-sampled feature maps have enabled the networking block of the Densenet to model over a wider frequency range dependency as well as longer contexts. Further, in order to recover the original resolutions by means of a lower resolution feature map, the concept of an upsampling layer has been considered as the transposed convolution, and the size of its filter is similar to that of the pooling. For permitting both the backwards and the forward flow of the signals without influencing the lower resolution blocks and thus the inter-block skip connection has been introduced, which is directly interconnected with the same scale to two dense blocks.

DenseNet: The efficient feature reuse operation of the DenseNet makes it as the well effective structure for the classification of plant diseases. Here, the vanishing gradient issues are effectively mitigated; also it promotes the reuse of features via the layer-to-layer connection between the models. The subtle features in the images are preserved more effectively to get valuable results in plant disease detection. Unlike the ResNet, EfficientNet, or transformer-based models like Swin Transformer models, the DenseNet model does not require heavy parameters, and it also provides effective results even for the highly variable agricultural images. In accordance with the feed-forward networking model, the outcomes of the layer m have been validated through $k_m = I_m(k_{m-1})$, in which the term k_0 is termed as the input of the network and the non-linear transformation that includes functions of the operations and is termed as $I_m(\cdot)$. To tackle the limitations of the deep models, the ResNet has deployed a skip connection and it is expressed in Eq. (18).

$$k_m = I_m(k_{m-1}) + k_{m-1} \quad (18)$$

Further, the skip connection has permitted the networking model to directly pass the gradient to the preceding layer. Here, the DenseNet has further enhanced the flow of information among the layers by replacing the output of the preceding layer. It is derived in Eq. (19).

$$k_m = I_m([k_{m-1}, k_{m-2}, \dots, k_0]) \quad (19)$$

Here, the concatenation process is given as $[\cdot \cdot \cdot]$. Moreover, the dense connectivity has ensured to reuse of the features that have been validated through preceding layers. It has neglected the determination of the most similar features over the diverse layers that made the networking model more effective. Here, I_m has been comprised of BN that has been followed up through ReLU and convolution along with the feature map \cdot . In order to perform the image recognition tasks, a pooling layer that collects the maps to the lower dimensions and local activation and thus it has become significant for acquiring global information effectively.

Dilation: In general, the process of dilation over the DenseNet model has been utilized to retrieve multi-level information from the images. Thus, the dilated dense block has acquired the ability to manage the same spatial resolution of the images. Through enlarging the small $n \times n$ kernel filter $n + (n - 1)(dl - 1)$, where the term dl is depicted as the dilation ratio, the size of the receptive field has been increased to the same size. Thus, the dilated layer has retrieved the multi-level information from the images.

Attention: The attention mechanism has been considered as a resource allocation in general. To learn the feature weight in accordance with the loss and thus to make the significant feature maps with greater weight is

the intent of this process. Thus, it has compressed the features along with the spatial attention as well as then turned every dimensional feature channel that has global receptive fields to some extent.

Dilation and attention-based multi-scale DenseNet: For attaining the feature maps along with the representative information through the image, the concept of dilation convolution is involved, and the convolution layer in the DenseNet model gets replaced with the dilation layer. This attention mechanism adaptively chose the relevant features through different feature maps in the domain.

Thus, this process is depicted in Fig. 6.

Recommended DAA-MDeNet for classification

The robust feature extractions from the segmented images are necessary to get accurate results in the classification of plant disease. The multiscale features are not effectively captured by the conventional models, which may limit their capability in accurately detecting the plant disease. In this work, the DAA-MDeNet is adopted, which leverages the attention mechanism and multiscale dense connectivity for improving the feature representation from the images. The challenging visual patterns and the variability among the diseased samples are accurately covered by the proposed DAA-MDeNet to get accurate results in the disease classification. To perform the classification process, the new heuristic model, termed DAA-MDeNet, has been implemented in this model. For tackling the restriction of the DenseNet model, a new formulation is derived. The DenseNet model has offered better computational efficiency. But this model has faced the replication of data multiple times. Therefore, with the aid of the AVLO algorithmic model, certain parameters in DenseNet have been optimized to enrich the performance. It is derived in Eq. (18).

$$of_1 = \arg \min_{\{af_{d_{ns}}, hn_{d_{ns}}, ep_{d_{ns}}\}} \left(\frac{1}{acc} + fnr + fdr + fpr \right) \quad (20)$$

Here, the terms $af_{d_{ns}}$, $hn_{d_{ns}}$, $ep_{d_{ns}}$ indicated the activation function among [0, 4], hidden neuron count among [5 to 255], and the number of epochs among [5 to 50] DenseNet are optimized using AVLO. Further, the terms acc , fdr , fnr and fpr are equated in Eqs. (21) to (24).

$$acc = \frac{(aa + az)}{(aa + az + bb + bz)} \quad (21)$$

$$fnr = \frac{bz}{az + aa} \quad (22)$$

$$fdr = \frac{bb}{bb + bz} \quad (23)$$

$$fpr = \frac{bb}{bb + az} \quad (24)$$

Thus, in the end, the final outcomes are classified as well as the process is given in Fig. 7.

Imaging results

The acquired resultant images over the detection process of plant disease are given in Fig. 8.

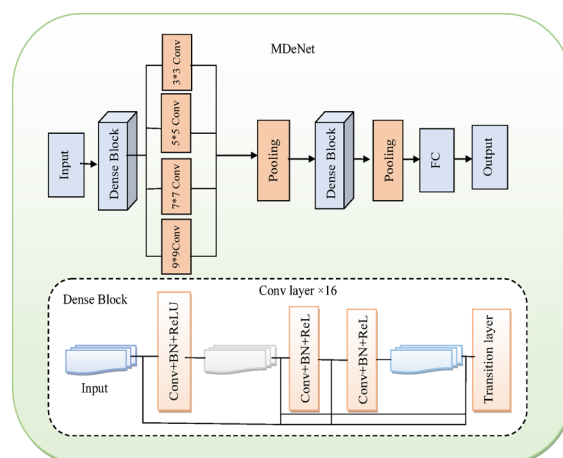


Fig. 6. Depiction of the DAA-MDeNet model.

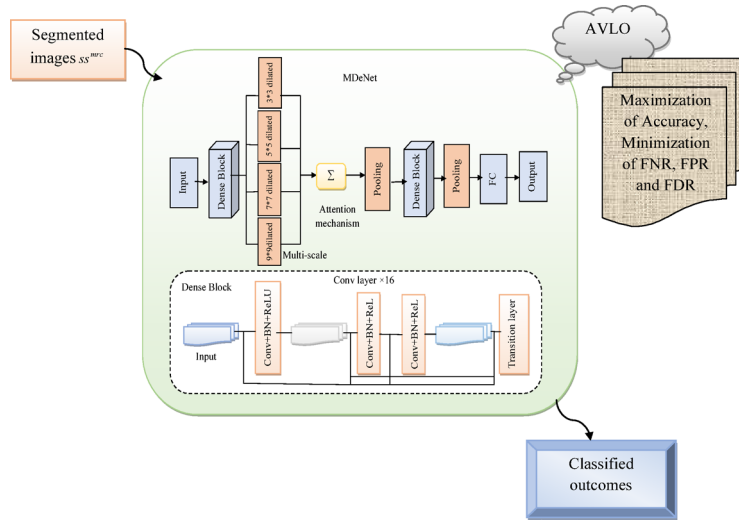


Fig. 7. Depiction of the classification process made using the DAA-MDeNet model.

Results and discussion

Simulation setup

The model was executed in Python, and its extensive results were obtained. Dingo Optimizer (DO)-DAA-MDeNet⁴⁵, Eurasian Oystercatcher Optimizer (EOO)-DAA-MDeNet⁴⁶, Residual Network (ResNet)⁴⁷, VGG16⁴⁸, CNN⁴⁹, and Modified CNN (MCNN)⁵⁰ models are used for assimilations. Here, 10 are population size, 6 as length of the chromosome, and 50 as the maximum number of iterations that are used in this model. The code for the implementation of developed model is available at the link: <https://github.com/kalicharan8u/Plant-Disease-Detection-using-Mask-RCNN-with-Multiscale-DenseNet>

Performance measures

An analysis of the processed model is carried out as follows:

- (a) Specificity in Eq. (25)

$$sp = \frac{az}{az + bb} \tag{25}$$

- (b) Sensitivity in Eq. (26)

$$sy = \frac{aa}{aa + bz} \tag{26}$$

- (c) Precision in Eq. (27)

$$pr = \frac{aa}{aa + bb} \tag{27}$$

- (d) NPV in Eq. (28)

$$npv = \frac{bz}{bz + az} \tag{28}$$

- (e) MCC in Eq. (29)

$$MCC = \frac{aa \times az - bb \times bz}{\sqrt{(aa + bb)(aa + bz)(az + bb)(az + bz)}} \tag{29}$$

- (f) F1-score in Eq. (30)

$$f1 - S = 2 \cdot \frac{sy \cdot pr}{pr + sy} \tag{30}$$

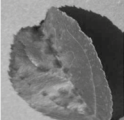
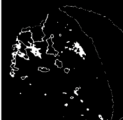
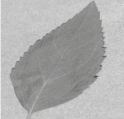



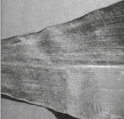
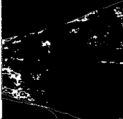
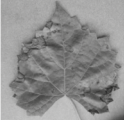



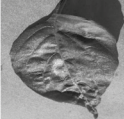

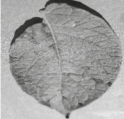

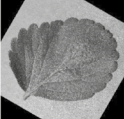
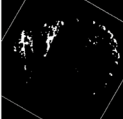
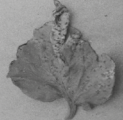

Dataset	Original image	Segmented images
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		

Fig. 8. Resultant images for plant disease detection framework.

Determination of classification process using algorithms and classifiers

By varying the number of batch sizes from [4 to 64] as well as the learning percentage from [35 to 88], the determination of the classification process in the plant disease detection model by varying both the algorithm and classifier is given in Figs. 9 and 10 accordingly. On considering the value of the classifier, at batch size 4 for accuracy, the AVLO-DAA-MDeNet model has a better value of 16%, 10%, 14% and 16% maximized than the Resnet, VGG16, DenseNet, and MDeNet models. It is similar to the case of algorithm comparison, where the AVLO-DAA-MDeNet model shows better value for positive measures than DO-DAA-MDeNet, EOO-DAA-MDeNet, AVOA-DAA-MDeNet, and LO-DAA-MDeNet and improves the classification process.

Statistical determination over the segmentation process using algorithms and classifiers

By varying the statistical measures, the determination of the segmentation process in the plant disease detection model by varying both the algorithm and classifier is given in Fig. 11. The graphical depiction shows a maximized value of accuracy, dice-coefficient, and Jaccard-coefficient value for the newly designed AVLO-DAA-MRCNN model, and thus better-segmented outcomes have aided the final classification process.

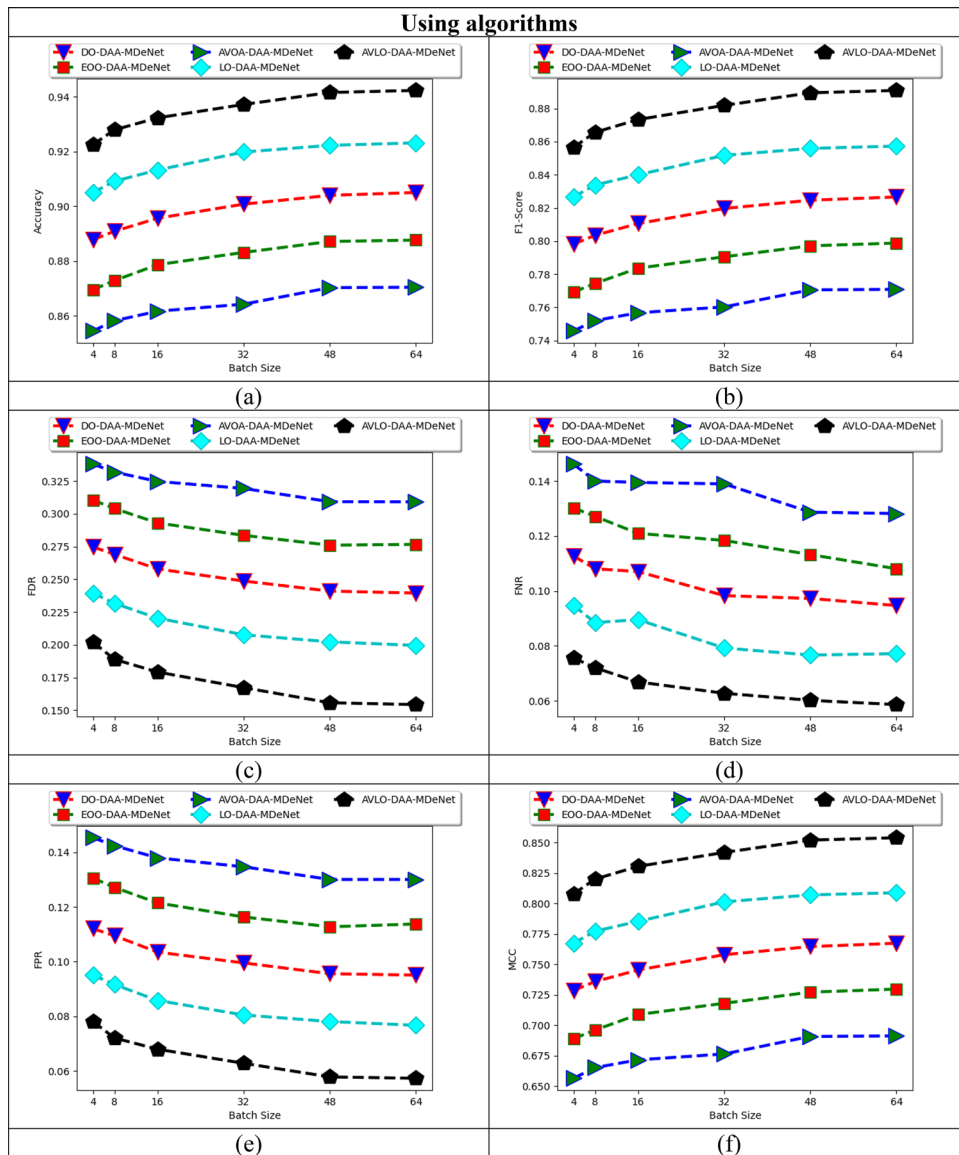


Fig. 9. Analyzing the classification process over the plant disease framework in terms of batch size by assimilating with algorithms and classifiers regarding (a) Accuracy, (b) F1-Score, (c) FDR, (d) FNR, (e) FPR, and (f) MCC.

Analysis on the confusion matrix for the plant disease classification model

Here, Fig. 12 shows the analysis of the confusion matrix for all ten datasets, and Fig. 13 shows the ROC and convergence analysis for dataset 1 for the proposed disease detection model. This validation has further enriched the entire performance of the given model and proved its efficiency also.

Overall analysis for plant disease classification model

Table 2 shows the validation that was made over the classification process of the plant disease detection framework in terms of both algorithms and classifiers for dataset 1. On considering the tabular values of accuracy over algorithm comparison, the proposed AVLO-DAA-MDeNet model has a better value of 6%, 3%, 2%, and 1% than DO-DAA-MDeNet, EOO-DAA-MDeNet, AVOA-DAA-MDeNet, and LO-DAA-MDeNet and improved the classification process.

Overall analysis for plant disease classification model using algorithms

Table 3 shows the statistical analysis made over the classification process of the plant disease detection framework using algorithms for dataset 1. When assimilating the proposed AVLO-DAA-MDeNet model with all other conventional algorithm models like DO-DAA-MDeNet, EOO-DAA-MDeNet, AVOA-DAA-MDeNet and LO-DAA-MDeNet, the proposed model has offered better outcomes.

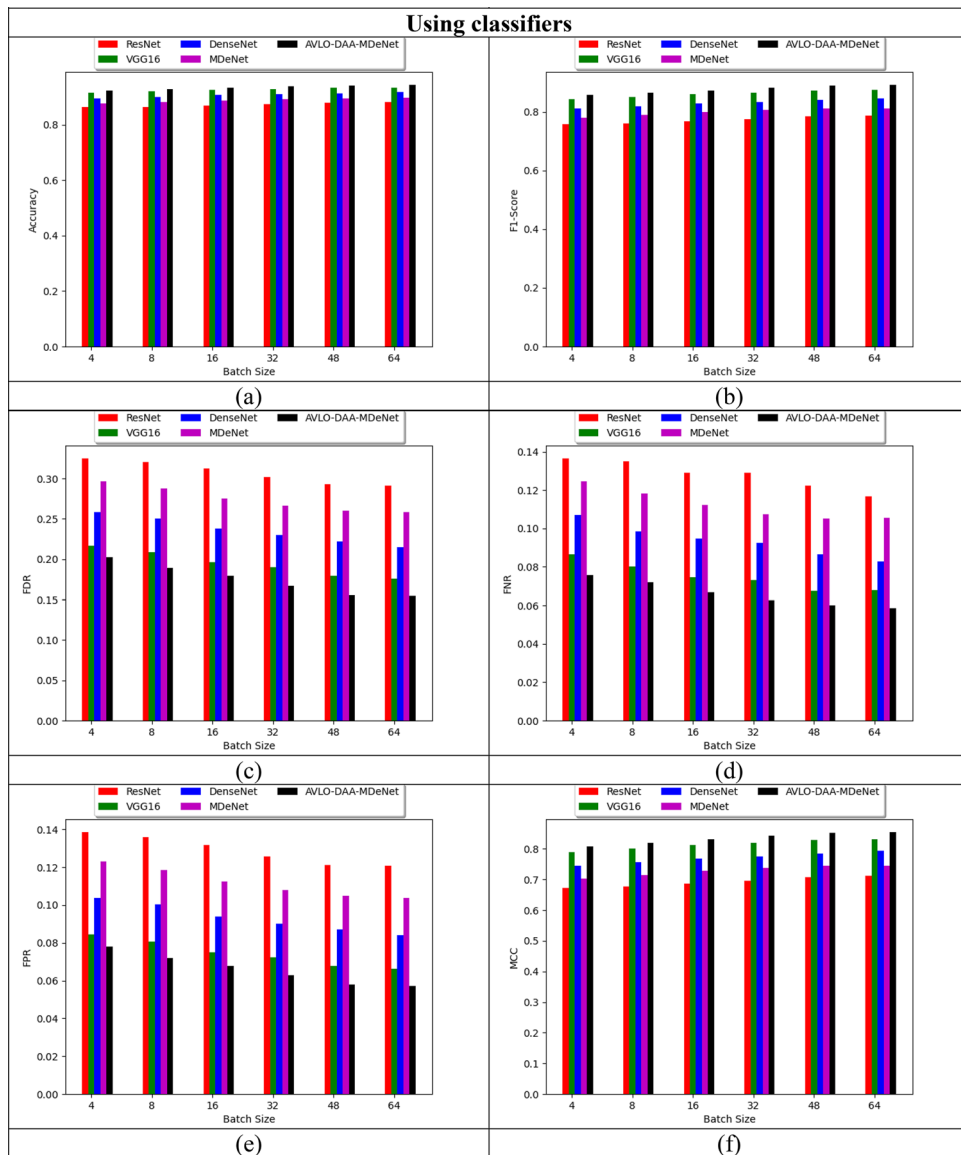


Fig. 9. (continued)

Accuracy report for plant disease classification model using classifiers in terms of datasets

The reports by means of accuracy for all ten datasets by assimilating with various conventional algorithms are shown in Table 4. On considering the values of Apple, the proposed AVLO-DAA-MDeNet model has acquired a better value of 6%, 1%, 3%, and 4% higher than Resnet, VGG16, DenseNet, and MDeNet models. It is the same for all other datasets; the proposed model has offered a better value.

Ablation study of the suggested model

In order to prove the interpretability of the presented plant disease detection and the segmentation model, an ablation study is conducted in Table 5. As noted from the Table 5, the incorporation of dilated convolutions, adaptive attention, and optimizers helps to get the accuracy value up to 94.58% but there is a reduction in the accuracy value is occurs when the DenseNet (88.04%) model is used alone for the plant disease classification process. Likewise, the incorporation of the dilation further improves the accuracy of the DenseNet model up to 89.91%, which proves the interpretability of the proposed model in the classification of plant diseases.

Comparison of the suggested model with the recent architectures

The suggested AVLO-DAA-MdeNet-based plant disease classification model is compared with the baseline approaches in Table 6 to claim the superior performance of the proposed approach in the plant disease classification task. It can be clearly seen from Table 6 is that the presented AVLO-DAA-MdeNet-based plant disease classification model provides an accuracy of 88.4% in the classification of the plant disease, which is comparatively superior to the Vision Transformers (85.424%), ConvNeXt (80.832%), and Swin Transformer

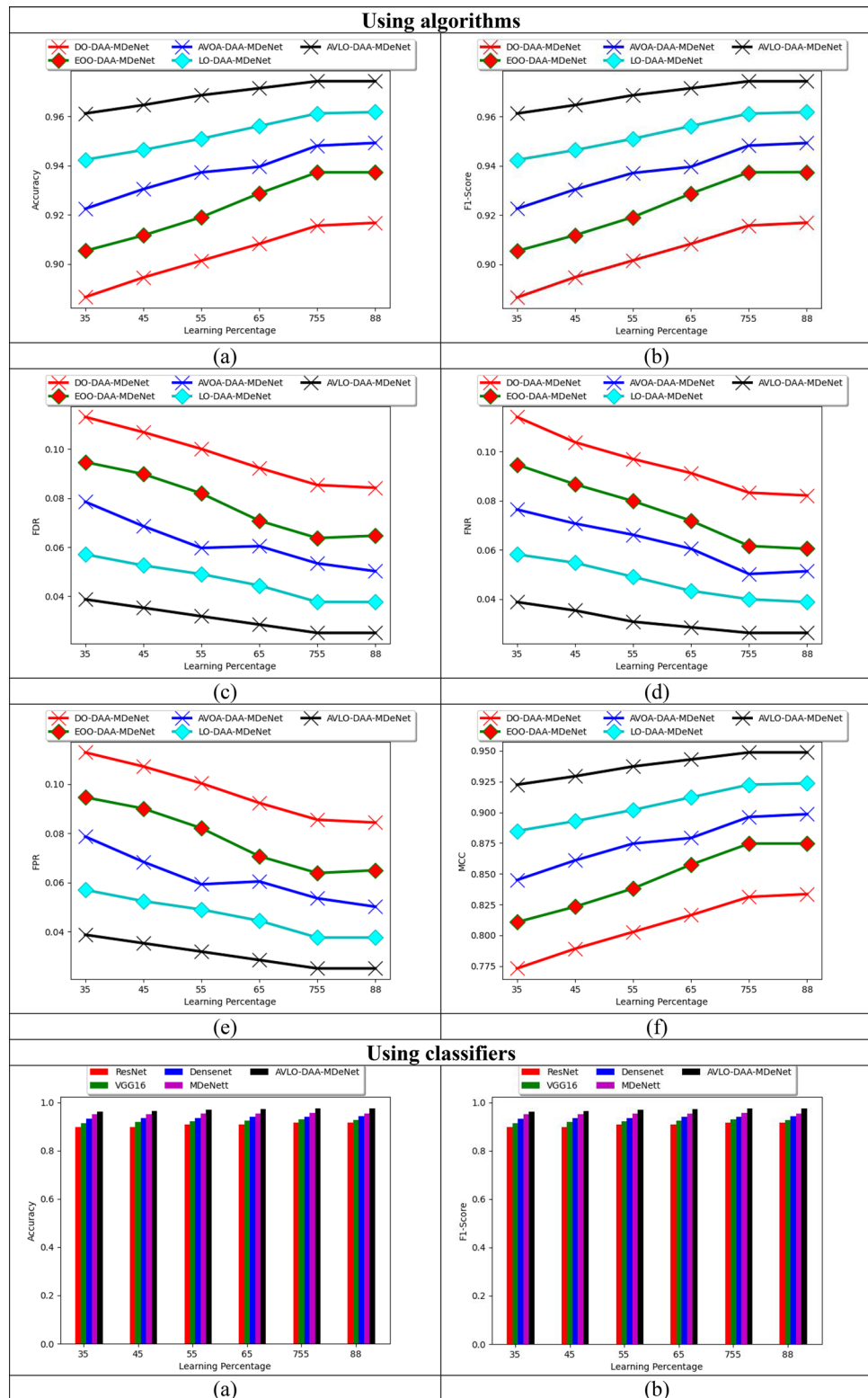


Fig. 10. Analyzing the classification process over the plant disease framework in terms of learning percentage by assimilating with algorithms and classifiers regarding (a) Accuracy, (b) F1-Score, (c) FDR, (d) FNR, (e) FPR, and (f) MCC.

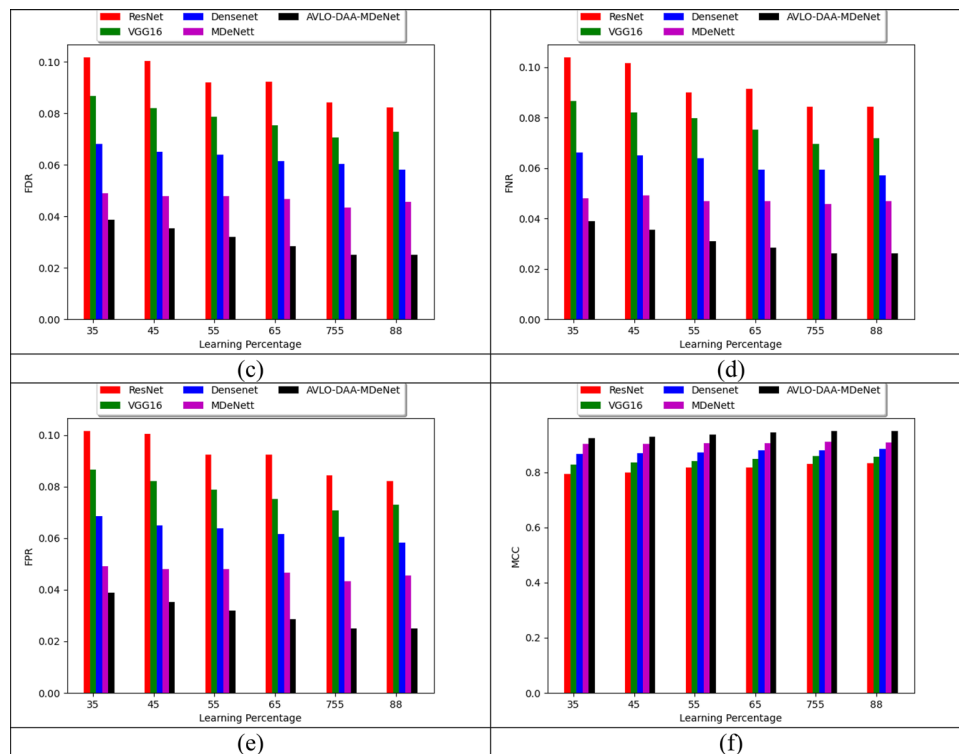


Fig. 10. (continued)

(83.88%). Thus, the comparative results in Table 6 prove the effectiveness of the proposed model in plant disease classification.

Generalizability and robustness of the suggested model, varying lighting, occlusions, or camera angles

The suggested generalizability and robustness of the presented model under varying lighting, occlusion or camera angles are given in Table 7. According to Fig. 14a, the generalizability of the presented AVLO-DAA-MdeNet model is measured in terms of accuracy metrics, which proves the higher generalization of the presented AVLO-DAA-MdeNet in the varying occlusion, lighting and camera angles. In addition, the robustness of the presented AVLO-DAA-MdeNet is measured in terms of the noise coefficient. Here, the correlation of the presented AVLO-DAA-MdeNet model is decreased towards the increase in noise coefficient, which effectively confirms the robustness of the presented AVLO-DAA-MdeNet against various lighting, occlusion or camera angles.

Comparison of the presented model with other multi-scale strategies

The suggested AVLO-DAA-MdeNet-based plant disease detection is compared with the other multiscale approaches, such as ASPP and FPN to confirm the lesion capturing capability of the presented model. Here, the proposed AVLO-DAA-MdeNet model reaches the accuracy value of 92.4% which is comparatively higher than the other multiscale approaches, such as ASPP (91.184%), FPN (85.256%) across varying scales. The accuracy improvement up to 7.2% confirms the capability of the presented AVLO-DAA-MdeNet in capturing the lesion across varying scales.

Attention maps of the proposed models

In both DAA-MRCNN and DAA-MDeNet models, the attention maps are used to emphasize the diagnostically relevant regions. The attention maps given in Fig. 15 help to illustrate that the model focuses on diseased areas versus healthy or background regions. Here, the integration of the attention maps can effectively focus on the diseased region, which greatly improves the accuracy of the plant disease detection operation. The visual explanations provided by the attention map would increase trust among agronomists and farmers and support real-world adoption in precision agriculture.

Comparison of the AVLO with other evolutionary algorithms

Table 8 provides the comparison on the convergence speed and the solution quality of the AVLO to prove the effectiveness over the other algorithms such as Adam, RMSprop, PSO, and GA. As noted from Table 8, the proposed AVLO is the improved version of the optimisation algorithm that provides a higher convergence speed and solution quality compared to the other approaches.

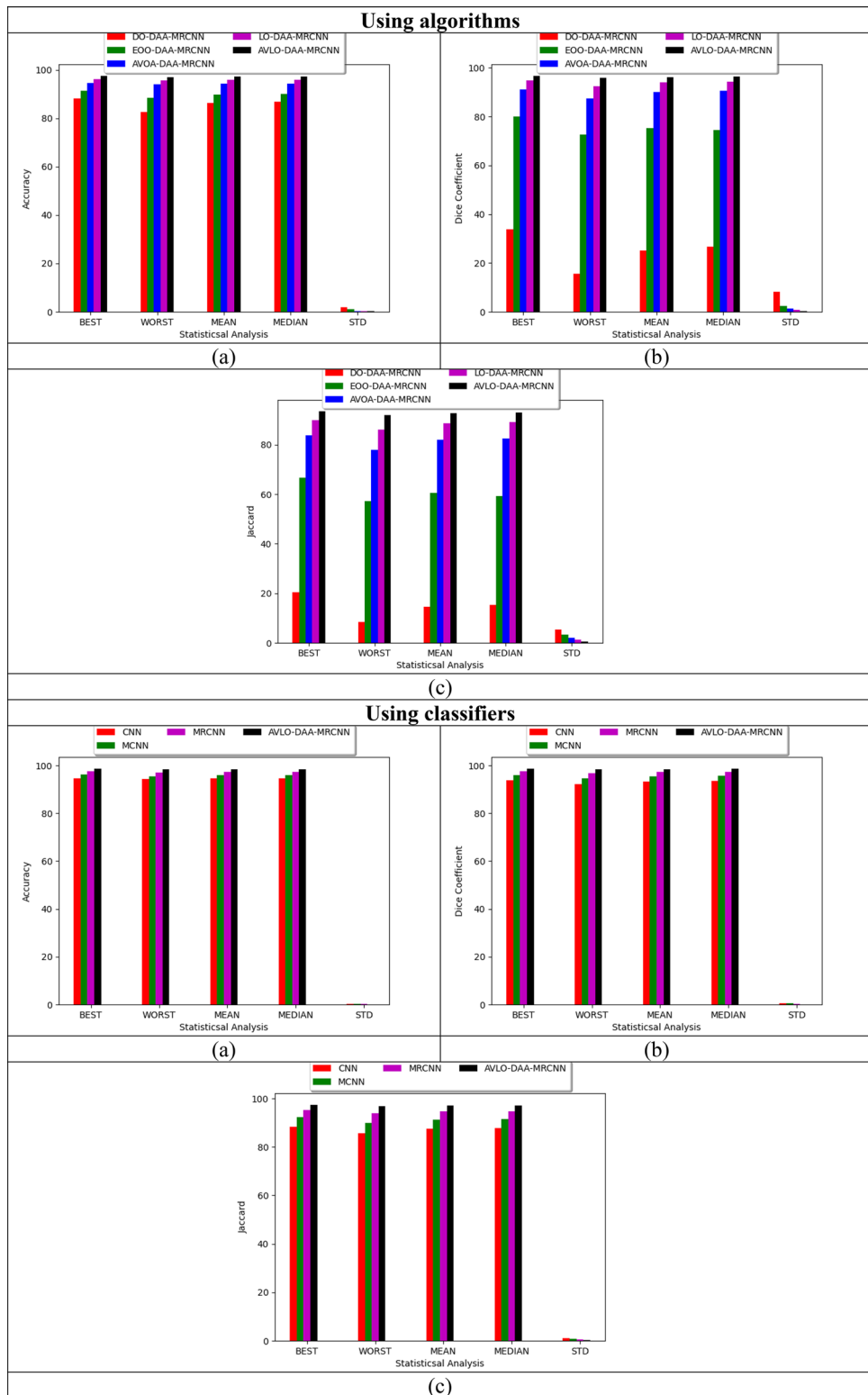


Fig. 11. Statistical analysis over the segmentation process of proposed plant disease classification framework by assimilating with algorithms and classifiers regarding (a) Accuracy, (b) dice-coefficient and (c) Jaccard coefficient.

Accuracy margin of the proposed model compared to the direct classification on raw images
 Table 9 shows the accuracy comparison of the two-stage pipeline (proposed model) and the direct classification on the raw images. As given in Table 9, the two-stage pipeline provides the accuracy value of 95.6% but the direct

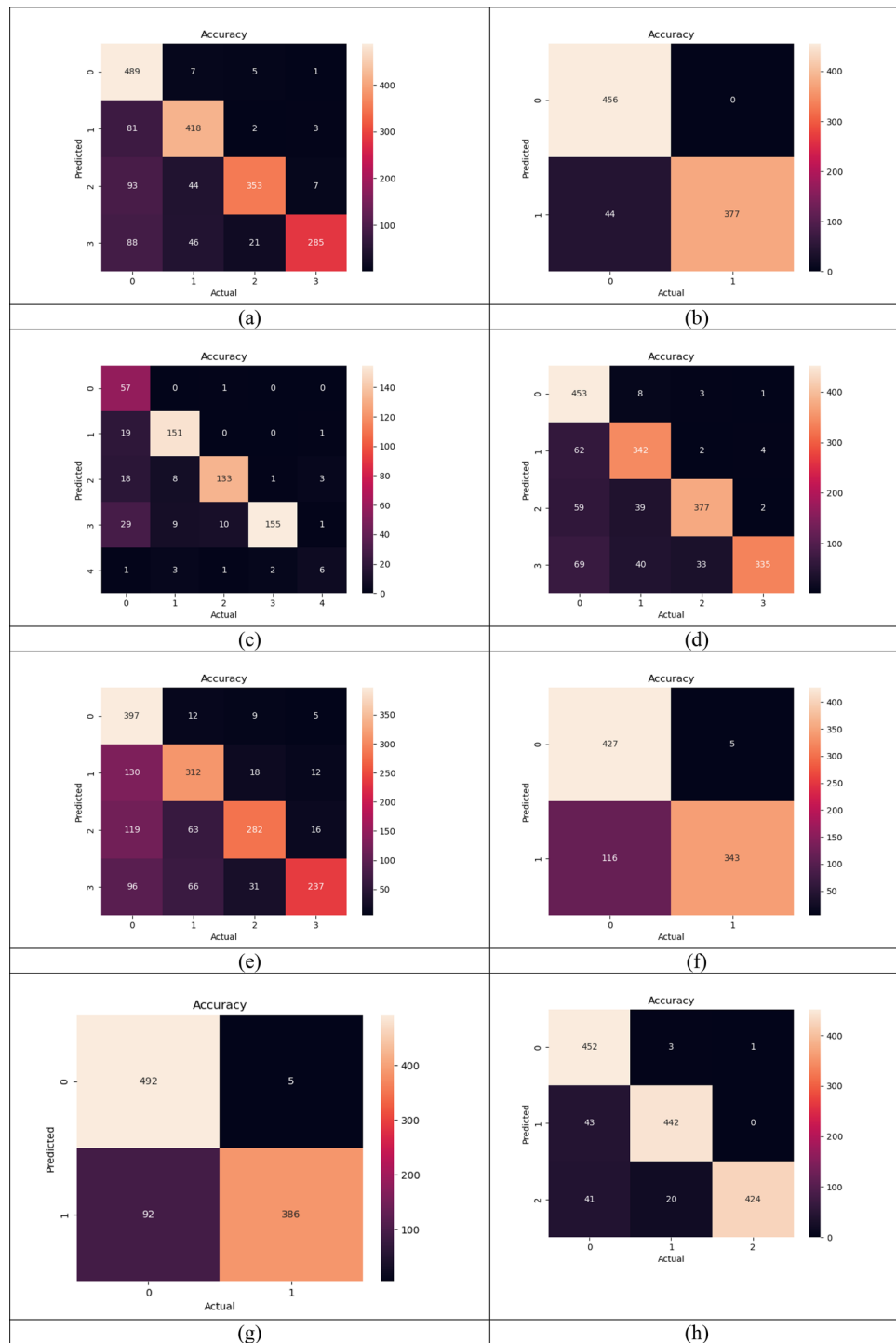


Fig. 12. Confusion analysis over the proposed plant disease classification framework regarding the dataset for (a) Apple, (b) Cherry, (c) Citrus, (d) Corn, (e) Grape, (f) Peach, (g) Pepper, (h) Potato, (i) Strawberry and (j) Tomato.

classification on the raw images provides the accuracy of 93.4%. This proves the extent of the presented two-stage pipeline over the other models.

Conclusion

In this research work, an improved plant disease classification model was developed using a hybrid deep learning architecture. From the standard publicly available database, the images were collected in the first step. Then, the attained images were fed to segmentation process for acquiring the segmented images that have aided the plant disease detection process to the next level, which has the potential to offer a more accurate and time-

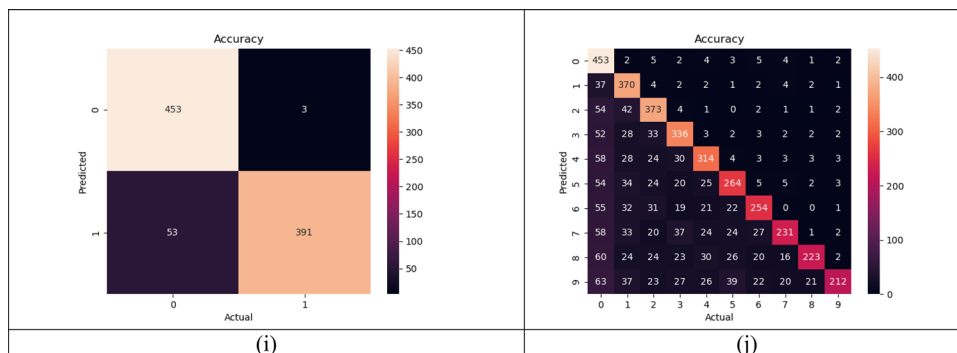


Fig. 12. (continued)

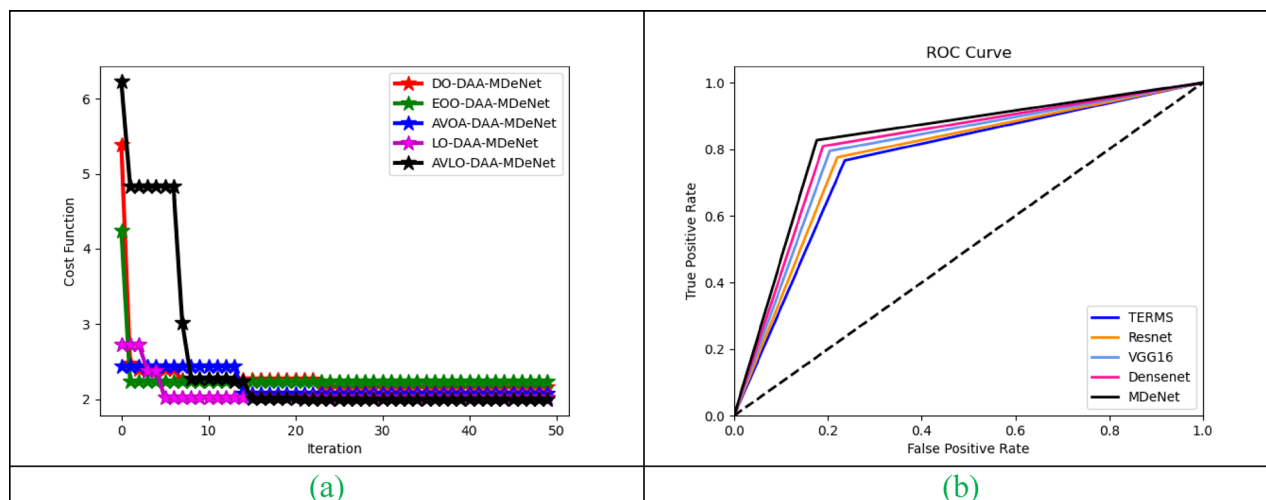


Fig. 13. Analysis of the proposed plant disease classification framework for dataset 1 regarding (a) ROC and (b) Convergence.

saving process. For performing the segmentation process, the newly developed DAA-MRCNN was used. Then, the segmented images were forwarded to the final phase of classification for classifying the diseased plant. In this phase, the newly designed DAA-MDeNet effectively performs the classification process. Further, by implementing the new AVLO algorithm for optimizing the parameters in both the DAA-MRCNN and DAA-MDeNet models for enhancing the performance of the given model. The determination of the detection process was carried out by assimilating it with other models, where the performance measures are used for validation. For the value of precision, the proposed AVLO-DAA-MDeNet model has acquired a better value of 6%, 4%, 3% and 1% higher than DO-DAA-MDeNet, EEO-DAA-MDeNet, AVOA-DAA-MDeNet and LO-DAA-MDeNet models. Timely and reliable assessments for detecting the disease in the plant for improving the protection activity are the future scopes.

Limitations and future scope: The class imbalance issues occur across species or disease categories, and these issues are solved in future work through data augmentation, loss weighting, or sampling techniques such as focal loss or SMOTE for improving the performance of the developed model in the underrepresented classes. Despite the high accuracy, the longer inference time and the higher computational complexity may limit the real-time performance of the developed model under the low-power edge environment. The suggested model is one of the effective approaches, but it is not feasible in a resource-constrained environment since it demands higher memory and processing power. The need for future work is highlighted because of the trade-off between the deployability and accuracy of the proposed model. In future, the practical use of the suggested model in the real-world agricultural setting is enabled by focusing on the deployment of pruning, quantization, and knowledge distillation approaches. Furthermore, the real-world deployment and feasibility of the suggested model on the real-world agricultural environment is assessed by evaluating the performance of the developed model in the edge platforms such as NVIDIA Jetson Nano and Raspberry Pi, etc.

Comparison on algorithms					
Measures	DO-DAA-MdeNet ⁴⁵	EOO-DAA-MdeNet ⁴⁶	AVOA-DAA-MdeNet ⁴¹	LO-DAA-MdeNet ⁴²	AVLO-DAA-MdeNet
“Accuracy”	91.56214	93.72862	94.81186	96.12315	97.43444
“Sensitivity”	91.67617	93.84265	94.9829	96.00912	97.37742
“Specificity”	91.44812	93.6146	94.64082	96.23717	97.49145
“Precision”	91.46758	93.62912	94.65909	96.22857	97.48858
“FPR”	8.551881	6.385405	5.359179	3.762828	2.508552
“FNR”	8.323831	6.157355	5.017104	3.990878	2.622577
“NPV”	91.44812	93.6146	94.64082	96.23717	97.49145
“FDR”	8.532423	6.370876	5.340909	3.771429	2.511416
“F1-Score”	91.57175	93.73576	94.82072	96.11872	97.43297
“MCC”	83.1245	87.45747	89.62424	92.24653	94.86893
Classifier comparison					
Measures	Resnet ⁴⁷	VGG16 ⁴⁸	Densenet ⁴⁴	MDeNet ⁴⁴	AVLO-DAA-MDeNet
“Accuracy”	91.56214	92.98746	94.01368	95.55302	97.43444
“Sensitivity”	91.56214	93.04447	94.0707	95.439	97.37742
“Specificity”	91.56214	92.93044	93.95667	95.66705	97.49145
“Precision”	91.56214	92.9385	93.96355	95.65714	97.48858
“FPR”	8.437856	7.069555	6.04333	4.332953	2.508552
“FNR”	8.437856	6.95553	5.929304	4.561003	2.622577
“NPV”	91.56214	92.93044	93.95667	95.66705	97.49145
“FDR”	8.437856	7.061503	6.036446	4.342857	2.511416
“F1-Score”	91.56214	92.99145	94.01709	95.54795	97.43297
“MCC”	83.12429	85.97497	88.02742	91.10628	94.86893

Table 2. Proposed plant disease classification models overall analysis for dataset 1 by means of algorithms and classifiers.

Measures	DO-DAA-MdeNet ⁴⁵	EOO-DAA-MdeNet ⁴⁶	AVOA-DAA-MdeNet ²⁷	LO-DAA-MdeNet ⁴²	AVLO-DAA-MDeNet
“BEST”	2.160891	2.230286	2.075638	2.017225	1.991723
“WORST”	5.385244	4.234116	2.437597	2.733341	6.223734
“MEAN”	2.285784	2.270363	2.176987	2.074722	2.477693
“MEDIAN”	2.160891	2.230286	2.075638	2.017225	1.991723
“STD”	0.450372	0.280536	0.162519	0.180945	1.058724

Table 3. Proposed plant disease classification models: statistical determination using algorithms.

Plant name	Resnet ⁴⁷	VGG16 ⁴⁸	Densenet ⁴⁹	MDeNet ⁴⁴	AVLO-DAA-MDeNet
Apple	87.9439	93.38652	91.40504	89.62944	94.31292
Cherry	91.84721	92.98746	94.35576	95.32497	97.49145
Citrus	90.67323	91.82266	92.93924	93.56322	96.42036
Corn	86.83707	88.66867	90.32258	92.14051	93.01531
Grape	80.29086	78.97507	81.5928	83.71191	84.27978
Peach	85.12907	83.72615	81.76207	87.03704	88.04714
Pepper	87.64103	93.4359	91.38462	89.79487	94.20513
Potato	91.67835	92.98738	94.27302	95.4418	97.45208
Strawberry	90.77778	91.83333	92.83333	93.61111	96.55556
Tomato	86.77863	88.54308	90.27045	92.15703	93.06216

Table 4. Proposed plant disease classification models accuracy report for the dataset by means of classifiers.

Term	DenseNet	Dilated DenseNet	Attention DenseNet	Multiscale DenseNet	Attention Multiscale DenseNet	AVLO-DAA-MDeNet
Accuracy	88.04	89.91	87.39	90.58	92.02	94.58

Table 5. Ablation study of the presented plant disease classification model.

Terms	Vision transformers ⁵¹	ConvNeXt ⁵²	Swin transformer ⁵³	AVLO-DAA-MDeNet
Accuracy	85.424	80.832	83.88	88.408
Sensitivity	85.47212	80.71926	83.93536	88.35551
Specificity	85.37492	80.94699	83.82353	88.46154
Precision	85.63492	81.20816	84.10859	88.65045
FPR	14.62508	19.05301	16.17647	11.53846
FNR	14.52788	19.28074	16.06464	11.64449
NPV	85.37492	80.94699	83.82353	88.46154
FDR	14.36508	18.79184	15.89141	11.34955
F1_score	85.55344	80.96297	84.02189	88.50274
MCC	0.708458	0.616637	0.677576	0.768149

Table 6. Comparison of the suggested plant disease classification model with the baseline approaches.

Terms	ASPP ⁵⁴	FPN ⁵⁵	AVLO-DAA-MDeNet
Accuracy	91.184	85.256	92.416
Sensitivity	91.18018	85.24771	92.47397
Specificity	91.18793	85.26452	92.35638
Precision	91.41095	85.61242	92.56159

Table 7. Comparison of the suggested plant disease classification model with other multiscale strategies.

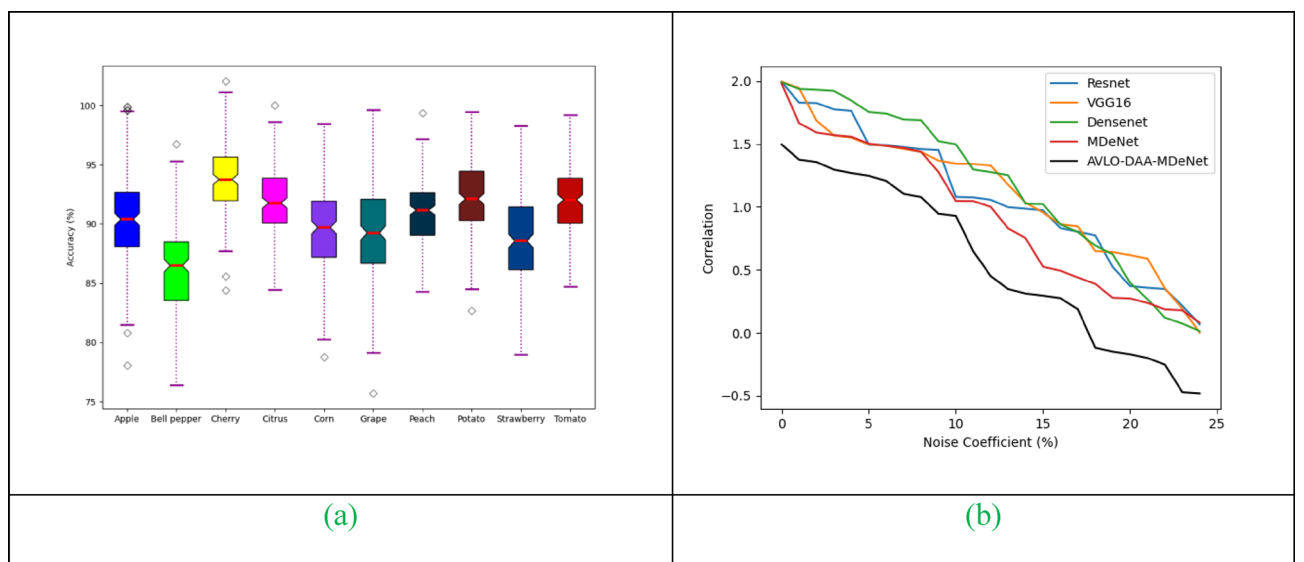


Fig. 14. Analysis of the proposed plant disease classification framework regarding “(a) Generalizability and (b) Robustness.”

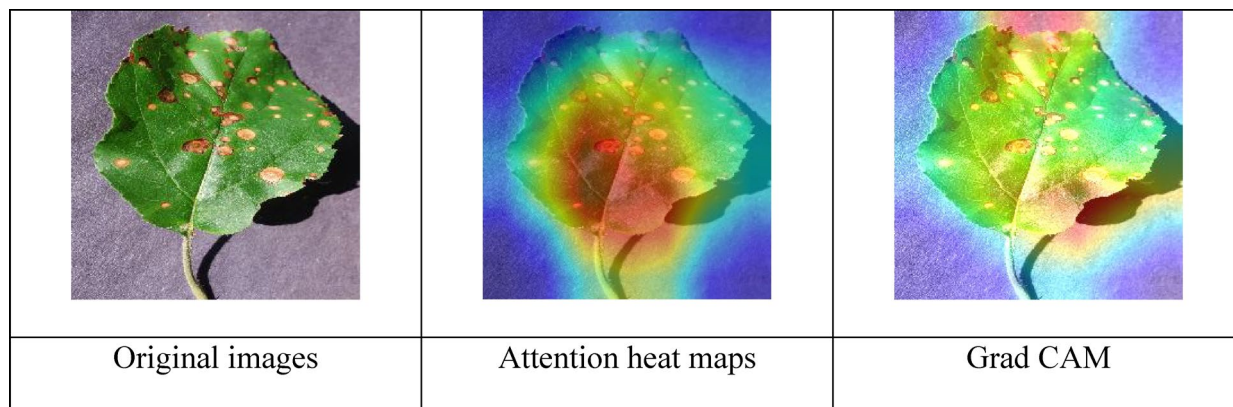


Fig. 15. Attention maps of the proposed plant disease classification framework.

Algorithms	Convergence speed	Solution quality	Search efficiency	Scalability	Generalization
GA	150	– 0.0001	Low	Poor	Low
Adam	200	– 0.0012	High	Moderate	Moderate
RMSprop	220	– 0.0025	Low	Good	High
PSO	300	– 0.0018	Moderate	Moderate	Moderate
Proposed AVLO	350	– 0.002	Very high	Excellent	Excellent

Table 8. Comparison of the AVLO with other evolutionary algorithms.

Terms	Direct classification on raw images	Two-stage pipeline
Accuracy	93.464	95.8

Table 9. Accuracy of the two-stage pipeline and direct comparison.

Data availability

In case of benchmark data: The data underlying this article are available in the dataset link as: <https://www.kaggle.com/datasets/lavaman151/plantifydr-dataset>.

Code availability

The code for the implementation of the developed model is available at the link <https://github.com/kalicharan8u/Plant-Disease-Detection-using-Mask-RCNN-with-Multiscale-DenseNet-> and it has been given in Section "Simulation setup".

Received: 8 July 2025; Accepted: 27 October 2025

Published online: 23 November 2025

References

- Ashourloo, D., Matkan, A. A., Huete, A., Aghighi, H. & Mobasheri, M. R. Developing an index for detection and identification of disease stages. *IEEE Geosci. Remote Sens. Lett.* **13**(6), 851–855. <https://doi.org/10.1109/LGRS.2016.2550529> (2016).
- Asha Rani, K. P. & Gowrishankar, S. Pathogen-based classification of plant diseases: A deep transfer learning approach for intelligent support systems. *IEEE Access* **11**, 64476–64493. <https://doi.org/10.1109/ACCESS.2023.3284680> (2023).
- Noon, S. K., Amjad, M., Qureshi, M. A. & Mannan, A. Handling severity levels of multiple co-occurring cotton plant diseases using improved YOLOX model. *IEEE Access* **10**, 134811–134825. <https://doi.org/10.1109/ACCESS.2022.3232751> (2022).
- Delnevo, G., Girau, R., Ceccarini, C. & Prandi, C. A deep learning and social IoT approach for plants disease prediction toward a sustainable agriculture. *IEEE Internet Things J.* **9**(10), 7243–7250. <https://doi.org/10.1109/JIOT.2021.3097379> (2022).
- Liu, Z. et al. Internet of Things (IoT) and machine learning model of plant disease prediction-blister blight for tea plant. *IEEE Access* **10**, 44934–44944. <https://doi.org/10.1109/ACCESS.2022.3169147> (2022).
- Zeng, Q., Ma, X., Cheng, B., Zhou, E. & Pang, W. GANs-based data augmentation for citrus disease severity detection using deep learning. *IEEE Access* **8**, 172882–172891. <https://doi.org/10.1109/ACCESS.2020.3025196> (2020).
- Kumar, M., Kumar, A. & Palaparthi, V. S. Soil sensors-based prediction system for plant diseases using exploratory data analysis and machine learning. *IEEE Sens. J.* **21**(16), 17455–17468. <https://doi.org/10.1109/JSEN.2020.3046295> (2021).
- Masood, M. et al. MaizeNet: A deep learning approach for effective recognition of maize plant leaf diseases. *IEEE Access* **11**, 52862–52876. <https://doi.org/10.1109/ACCESS.2023.3280260> (2023).
- Shovon, M. S. H. et al. PlantDet: A robust multi-model ensemble method based on deep learning for plant disease detection. *IEEE Access* **11**, 34846–34859. <https://doi.org/10.1109/ACCESS.2023.3280260> (2023).

10. Chen, J., Chen, W., Zeb, A., Yang, S. & Zhang, D. Lightweight inception networks for the recognition and detection of rice plant diseases. *IEEE Sens. J.* **22**(14), 14628–14638. <https://doi.org/10.1109/JSEN.2022.3182304> (2022).
11. Li, F. et al. A remote sensing and airborne edge-computing based detection system for pine wilt disease. *IEEE Access* **9**, 66346–66360. <https://doi.org/10.1109/ACCESS.2021.3073929> (2021).
12. Tian, L. et al. VMF-SSD: A novel V-space based multi-scale feature fusion SSD for apple leaf disease detection. *IEEE/ACM Trans. Comput. Biol. Bioinf.* **20**(3), 2016–2028. <https://doi.org/10.1109/TCBB.2022.3229114> (2023).
13. Patle, K. S. et al. IoT enabled, leaf wetness sensor on the flexible substrates for in-situ plant disease management. *IEEE Sens. J.* **21**(17), 19481–19491. <https://doi.org/10.1109/JSEN.2021.3089722> (2021).
14. Ashourloo, D., Aghighi, H., Matkan, A. A., Mobasheri, M. R. & Rad, A. M. An Investigation into machine learning regression techniques for the leaf rust disease detection using hyperspectral measurement. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **9**(9), 4344–4351. <https://doi.org/10.1109/JSTARS.2016.2575360> (2016).
15. Zhu, X. et al. LAD-Net: A novel light weight model for early apple leaf pests and diseases classification. *IEEE/ACM Trans. Comput. Biol. Bioinf.* **20**(2), 1156–1169. <https://doi.org/10.1109/TCBB.2022.3191854> (2023).
16. Schor, N. et al. Robotic disease detection in greenhouses: Combined detection of powdery mildew and tomato spotted wilt virus. *IEEE Robot. Automat. Lett.* **1**(1), 354–360. <https://doi.org/10.1109/LRA.2016.2518214> (2016).
17. Nie, X., Wang, L., Ding, H. & Xu, M. Strawberry verticillium wilt detection network based on multi-task learning and attention. *IEEE Access* **7**, 170003–170011. <https://doi.org/10.1109/ACCESS.2019.2954845> (2019).
18. Saini, R., Patle, K. S., Kumar, A., Surya, S. G. & Palaparthi, V. S. Attention-based multi-input multi-output neural network for plant disease prediction using multisensor system. *IEEE Sens. J.* **22**(24), 24242–24252. <https://doi.org/10.1109/JSEN.2022.3219601> (2022).
19. Moupojou, E., Tagne, A., Retraint, F. & Tadonk, A. FieldPlant: A dataset of field plant images for plant disease detection and classification with deep learning. *IEEE Access* **11**, 35398–35410 (2023).
20. Saleem, M. H., Potgieter, J. & Arif, K. M. A performance-optimized deep learning-based plant disease detection approach for horticultural crops of New Zealand. *IEEE Access* **10**, 89798–89822 (2022).
21. Sunil, C. K. & Jaidhar, C. D. Cardamom plant disease detection approach using EfficientNetV2. *IEEE Access* **10**, 789–804 (2022).
22. Hosny, K. M., El-Hady, W. M., Samy, F. M., Vrochidou, E. & Papakostas, G. A. Multi-class classification of plant leaf diseases using feature fusion of deep convolutional neural network and local binary pattern. *IEEE Access* **11**, 62307–62317 (2023).
23. Amin, H., Darwish, A., Hassanien, A. E. & Soliman, M. End-to-end deep learning model for corn leaf disease classification. *IEEE Access* **10**, 31103–31115 (2022).
24. Vishnoi, V. K., Kumar, K., Kumar, B., Mohan, S. & Khan, A. A. Detection of apple plant diseases using leaf images through convolutional neural network. *IEEE Access* **11**, 6594–6609 (2023).
25. Zhao, Y., Chen, Z., Gao, X., Song, W. & Xion, Q. Plant disease detection using generated leaves based on DoubleGAN. *IEEE/ACM Trans. Comput. Biol. Bioinf.* **19**(3), 1817–1826 (2021).
26. Ahmad, M., Abdullah, M., Moon, H. & Han, D. Plant disease detection in imbalanced datasets using efficient convolutional neural networks with stepwise transfer learning. *IEEE Access* **9**, 140565–140580 (2021).
27. Hassan, E., Saber, A., El-Sappagh, S. & El-Rashidy, N. Optimized ensemble deep learning approach for accurate breast cancer diagnosis using transfer learning and grey wolf optimization. *Evol. Syst.* **16**(2), 59 (2025).
28. Hassan, E., Saber, A., El-Kenawy, E.-S. M., Bhatnagar, R. & Shams, M. Y. Early detection of black fungus using deep learning models for efficient medical diagnosis. In *2024 International Conference on Emerging Techniques in Computational Intelligence (ICETCI)* 426–431 (IEEE, 2024).
29. Hassan, E., Saber, A., El-Hafeez, T. A., Medhat, T. & Shams, M. Y. Enhanced dysarthria detection in cerebral palsy and ALS patients using WaveNet and CNN-BiLSTM models: A comparative study with model interpretability. *Biomed. Signal Process. Control* **110**, 108128 (2025).
30. Chouhan, S. S., Singh, U. P., Saxena, A. & Jain, S. Assessing the importance and need of artificial intelligence for precision agriculture. In *Artificial Intelligence Techniques in Smart Agriculture* 1–6 (Springer Nature Singapore, Singapore 2024).
31. Chouhan, S. S., Kaul, A. & Singh, U. P. Image segmentation using fuzzy competitive learning based counter propagation network. *Multimed. Tools Appl.* **78**(24), 35263–35287 (2019).
32. Sharma, A., Patel, R. K., Pranjali, P., Panchal, B. & Chouhan, S. S. Computer vision-based smart monitoring and control system for crop. In *Applications of Computer Vision and Drone Technology in Agriculture 4.0* 65–82 (Springer Nature Singapore, Singapore 2024).
33. Chouhan, S. S., Singh, U. P. & Jain, S. Performance evaluation of different deep learning models used for the purpose of healthy and diseased leaves classification of Chirimoya (*Annona Chirimola*) plant. *Neural Comput. Appl.* **37**(6), 4531–4544 (2025).
34. Mahmood, T. et al. A brief survey on breast cancer diagnostic with deep learning schemes using multi-image modalities. *IEEE Access* **8**, 165779–165809 (2020).
35. Rehman, A., Mahmood, T. & Saba, T. Robust kidney carcinoma prognosis and characterization using Swin-ViT and DeepLabV3+ with multi-model transfer learning. *Appl. Soft. Comput.* **170**, 112518 (2025).
36. Ali, A. M., Benjdira, B., Koubaa, A., Boulila, W. & El-Shafai, W. TESR: two-stage approach for enhancement and super-resolution of remote sensing images. *Remote Sens.* **15**(9), 2346 (2023).
37. Mahmood, T., Saba, T., Rehman, A. & Alamri, F. S. Harnessing the power of radiomics and deep learning for improved breast cancer diagnosis with multiparametric breast mammography. *Expert Syst. Appl.* **249**, 123747 (2024).
38. Mahmood, T., Rehman, A., Saba, T., Wang, Y. & Alamri, F. S. Alzheimer’s disease unveiled: Cutting-edge multi-modal neuroimaging and computational methods for enhanced diagnosis. *Biomed. Signal Process. Control* **97**, 106721 (2024).
39. Mahmood, T., Saba, T. & Rehman, A. Breast cancer diagnosis with MFF-HistoNet: a multi-modal feature fusion network integrating CNNs and quantum tensor networks. *J. Big Data* **12**(1), 60 (2025).
40. Mahmood, T., Saba, T., Al-Otaibi, S., Ayesha, N. & Almasoud, A. S. AI-Driven microscopy: Cutting-edge approach for breast tissue prognosis using microscopic images. *Microsc. Res. Tech.* **88**(5), 1335–1359 (2025).
41. Abdollahzadeh, B., Gharehchopogh, F. S. & Mirjalili, S. African vultures optimization algorithm: A new nature-inspired metaheuristic algorithm for global optimization problems. *Comput. Ind. Eng.* **158**, 107408 (2021).
42. Abasi, A. K. et al. Lemurs optimizer: A new metaheuristic algorithm for global optimization. *Appl. Sci.* **12**(19), 10057 (2022).
43. Cai, L., Long, T., Dai, Y. & Huang, Y. Mask R-CNN-based detection and segmentation for pulmonary nodule 3D visualization diagnosis. *IEEE Access* **8**, 44400–44409 (2020).
44. Takahashi, N. & Mitsufuji, Y. Multi-scale multi-band densenets for audio source separation. In: *2017 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)* (2017).
45. Bairwa, A. K., Joshi, S. & Singh, D. Dingo optimizer: A nature-inspired metaheuristic approach for engineering problems. *Math. Probl. Eng.* **2021**, 2571863 (2021).
46. Bader, A. S., Jummar, W., Jasem, F. M. & Al-Yousif, M. Eurasian oystercatcher optimiser: New meta-heuristic algorithm. *J. Intell. Syst.* **31**, 332–344 (2022).
47. Liang, J. Image classification based on RESNET. *J. Phys: Conf. Ser.* **1634**, 012110 (2020).
48. Tammina, S. Transfer learning using VGG-16 with deep convolutional neural network for classifying images. *Int. J. Sci. Res. Publ.* **9**, 143–150 (2019).
49. Xu, G., Shen, X., Chen, S., Zong, Y. & Zhang, C. A deep transfer convolutional neural network framework for EEG signal classification. *IEEE Access* **7**, 112767–112776 (2019).

50. Hasan, D. A. & Abdulazeez, A. M. A modified convolutional neural networks model for medical image segmentation. *Test Eng. Manage.* **20**, 22 (2020).
51. Chang, B., Wang, Y., Zhao, X., Li, G. & Yuan, P. A general-purpose edge-feature guidance module to enhance vision transformers for plant disease identification. *Expert Syst. Appl.* **237**, 121638 (2024).
52. KP, A. R. & Gowrishankar, S. Convnext-based mango leaf disease detection: Differentiating pathogens and pests for improved accuracy. *Int. J. Adv. Comput. Sci. Appl.* **14**(6), 2023.
53. Kalpana, P., Anandan, R., Hussien, A. G., Migdady, H. & Abualigah, L. Plant disease recognition using residual convolutional enlightened Swin transformer networks. *Sci. Rep.* **14**(1), 8660 (2024).
54. Ma, W. et al. Crop disease detection against complex background based on improved atrous spatial pyramid pooling. *Electronics* **12**(1), 216 (2023).
55. Chen, S., Yang, W. & Chen, X. Automatic recognition of agriculture pests with balanced feature pyramid network. *Appl. Eng. Agric.* **40**(5), 525–535 (2024).

Acknowledgements

I would like to express my very great appreciation to the co-authors of this manuscript for their valuable and constructive suggestions during the planning and development of this research work.

Author contributions

Conceptualization: K.S. and S.T.; Methodology: M.K.S., S.S.; validation: J.S.P. and H.K.S.; Formal analysis: H.K.S., S.S. and G.K.; investigation: H.K.S., S.S. and J.S.P.; resources: K.S. and M.K.S.; data curation: G.K. and J.S.P.; writing original draft preparation: K.S., S.T. and M.K.S.; writing review and editing: H.K.S. and G.K.; visualization: K.S. and J.P.S.; supervision: H.K.S. and J.P.S.; project administration: M.K.S. and S.T.

Funding

Open access funding provided by Manipal University Jaipur. This research did not receive any specific funding.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to K.S. or H.K.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025