ITC 2/55

Information Technology and Control
Vol. 55 / No. 2/ 2026
pp. 402-422
DOI 10.5755/j01.itc.55.2.41390

EfficientNetB4-SECA: A Deep Learning Model for Detecting Leaf Diseases in Tomato and Chili Plants

Received 2025/05/02 | Accepted after revision 2025/09/30

HOW TO CITE: Vu, V., G., Thach, T., D., T., Van, T., D. (2026) EfficientNetB4-SECA: A Deep Learning Model for Detecting Leaf Diseases in Tomato and Chili Plants. *Information Technology and Control*, 55(2), 402-422. https://doi.org/10.5755/j01.itc.55.2.41390

EfficientNetB4-SECA: A Deep Learning Model for Detecting Leaf Diseases in Tomato and Chili Plants

Giang Van Vu

University of People's Security; Ho Chi Minh City, Vietnam; e-mail: vuvangiang@gmail.com;
HUTECH University, Ho Chi Minh City, Vietnam; e-mail: vvgiang23nct@hutech.edu.vn

Thao Dang Thi Thach

University of Information Technology; VNU-HCM, Ho Chi Minh City, Vietnam; e-mail: thaodtt.16@grad.uit.edu.vn;
Faculty of Information Technology; HUTECH University, Ho Chi Minh City, Vietnam; e-mail: dtt.thao@hutech.edu.vn

Hoang Thien Van*

Faculty of Information Technology; HUTECH University; Ho Chi Minh City; Vietnam; e-mail: vt.hoang@hutech.edu.vn

Corresponding author: vt.hoang@hutech.edu.vn

Tomatoes and chilies are essential crops vulnerable to foliar diseases that significantly reduce yield and quality. Accurately identifying these diseases is essential for effective precision agriculture. Despite advances in deep learning, few studies have systematically evaluated state-of-the-art models on crop-specific datasets. To address this gap, we assessed the performance of leading architectures, including EfficientNet-B4, Vision Transformer, InceptionV3, and EfficientNet-B4 variants with attention, on Tomato Leaf Disease and COLD Chili datasets. Motivated by the evaluation results, we propose EfficientNet-B4-SECA, a novel deep learning model tailored for tomato and chili leaf disease classifications. It builds upon the EfficientNet-B4 backbone and incorporates two attention mechanisms: Squeeze-and-Excitation, which enhances key feature channels, and Coordinate Attention, which captures spatial and channel dependencies. Experimental results indicate that EfficientNet-B4-SECA achieves the highest classification accuracy on both datasets, with 84.97% for tomato and 87.83% for chili leaf diseases, surpassing that of all compared models. These findings demonstrate the effectiveness of the model and its potential for real-world deployment in agricultural disease monitoring.

KEYWORDS: EfficientNet-B4; Deep Learning; Leaf Disease Classification; Coordinate Attention; Squeeze-and-Excitation.

1. Introduction

Crop improvement has become a key research focus owing to globally increasing food needs. Tomatoes and chilies, among the most economically important crops, are valued for their nutritional content, distinctive flavour, and commercial viability [15], [24]. However, they are highly susceptible to foliar diseases, such as Bacterial Spot, Early and Late Blight, Septoria Leaf Spot, Target Spot, as well as various fungal and viral infections [10], [4], [18]. These diseases can cause substantial yield losses and reduced crop quality, posing serious threats to both economic stability and food security. Accurate and timely diagnosis of plant diseases is thus essential to prevent outbreaks and support sustainable agriculture [15], [10]. Traditional diagnostic methods relying on expert visual inspection are time-consuming, labour-intensive, and susceptible to human error [24], [10].

In recent years, artificial intelligence (AI), particularly deep learning (DL), has emerged as a promising solution for automating disease detection in precision agriculture [5]. Precision agriculture, integrating sensors, robotics, and AI algorithms, enables early intervention, reduces chemical usage, and optimises resource allocation [15], [24], [7]. DL models such as Convolutional Neural Networks (CNNs), ResNet [17], Vision Transformers (ViT) [12], and EfficientNet [8], [25], [29] have demonstrated strong potential for plant disease classification. However, these models often struggle to capture fine-grained lesions and spatial patterns, particularly in cases of visually similar diseases under varying environmental conditions [5]. Despite recent advances, systematic evaluation of state-of-the-art models on crop-specific datasets remains lacking.

To address this gap, we compared the performance of leading architectures: EfficientNet-B4, ViT, InceptionV3, and EfficientNet-B4 variants with attention, using two benchmark datasets, including Tomato Leaf Disease (TLD) and COLD Chili. Although EfficientNet-B4 offers a favourable balance between accuracy and efficiency, it has mainly been used without attention enhancements, which are specifically tailored for the visual complexity of plant diseases. Motivated by the limitations of existing approaches, we propose EfficientNet-B4-SECA, a novel attention-enhanced architecture that integrates the Squeeze-and-Excitation (SE) and Coordinate Attention (CA) modules into the EfficientNet-B4 backbone. This design improves the ability of the model to capture both critical feature channels and spatial dependencies, resulting in an improved classification performance across diverse disease types and imaging conditions. The main contributions of this study are as follows:

- A comparative analysis of DL models (EfficientNet-B4, ViT, InceptionV3, EfficientNet-B4-CA, and EfficientNet-B4-CBAM) for classifying tomato and chili leaf disease.
- Proposal of the EfficientNet-B4-SECA architecture, integrating the SE and CA attention mechanisms to improve model robustness and accuracy.
- EfficientNet-B4-SECA achieves the highest classification accuracy on both datasets, with 84.97% for tomato and 87.83% for chili leaf diseases, surpassing that of all compared models.

The remaining paper is structured as follows: Section 2 reviews the related studies. Section 3 details the proposed method and model training procedure. Section 4 presents the experimental results and discussion. Finally, Section 5 concludes the paper.

2. Related Works

Recent studies have demonstrated that DL effectively classifies plant diseases with high accuracy. Unlike traditional machine learning methods, which require complex handcrafted feature extraction, DL enables direct learning from raw images with minimal preprocessing. This advantage enhances classification performance as well as reduces reliance on manual image processing.

Rangarajan et al. [21] applied transfer learning with AlexNet and VGG16 to classify tomato leaf diseases using the PlantVillage dataset, which included six diseases and one healthy class. Both models were fine-tuned by modifying the final layers, achieving an accuracy of 97.49% with AlexNet and 97.23% with VGG16, demonstrating the effectiveness of transfer learning in diagnosing plant diseases. Abbas et al. [2] proposed a tomato leaf disease classification method that integrates synthetic images generated using a Conditional GAN with real data for training a DenseNet-121 model using transfer learning. Their approach, evaluated on the PlantVillage dataset, achieved high classification accuracies of 99.51%, 98.65%, and 97.11% for 5, 7, and 10 disease classes, respectively. Abayomi-Alli et al. [1] proposed a novel data augmentation method to improve plant disease classification under low-quality imaging conditions. Their approach involved generation of synthetic images by transforming colour histograms using Chebyshev functions, thereby expanding the trainable colour space and enabling neural networks to better detect key features despite issues such as blurring, motion, or overexposure. Agarwal et al. [3] proposed ToLeD, a CNN model for identifying tomato leaf disease, comprising three convolutional layers, three max-pooling layers, and two fully connected (FC) layers. ToLeD was compared with VGG16, InceptionV3, and MobileNet on the PlantVillage dataset (including nine diseases and one healthy class), achieving an average accuracy of 91.2%, outperforming other models. Atila et al. [8] used EfficientNet to classify 39 classes in the PlantVillage dataset, achieving an accuracy of 99.97% with EfficientNet-B4 and 99.91% with EfficientNet-B5. EfficientNet outperformed AlexNet, ResNet50, VGG16, and InceptionV3 in classification accuracy and overall performance. Rukhsar et al. [22] used InceptionV3 with transfer learning to classify rice leaf diseases. The model, trained on 3,355 images divided into four classes, achieved an accuracy of 90.50% on the training set and 82.52% on the test set, outperforming a standard CNN model (63.40%). This study emphasised the effectiveness of transfer learning and its potential for smart monitoring of plants. Amin et al. [6] proposed a DL model combining features from EfficientNet-B0 and DenseNet-121 to classify maize leaf diseases. The model was trained on a subset of the PlantVillage dataset, which contained four leaf conditions, achieving an accuracy of 98.56%, thus surpassing ResNet152 (98.37%) and InceptionV3 (96.26%) and highlighting its potential for smart agriculture. Pratap et al. [20] introduced EfficientLeafNetB4, an optimised model derived from EfficientNet-B4, for classifying chili leaf diseases using 2,000 images of five classes. The model, fine-tuned to better capture leaf features, achieved an accuracy of 95.82%, outperforming ResNet-50, DenseNet-121, MobileNetV2, and VGG-16. This result demonstrated the effectiveness of the model in disease detection and its potential for application in smart agriculture. Bhuyan et al. [11] developed SE_SPnet, a CNN model with parallel convolutional layers and SE blocks for predicting rice leaf strength. Tested on three disease datasets, SE_SPnet achieved an accuracy of 99.2% using the SGD optimisation algorithm, outperforming other methods because ofits ability to enhance feature extraction and reduce redundancy. Shafik et al. [23] proposed PDDNet, a plant disease detection model based on transfer learning with nine pre-trained CNN models. Tested on the PlantVillage dataset (comprising 15 classes and 54,305 images), PDDNet-LVE achieved an accuracy of 97.79%, outperforming standard CNN models. Singh et al. [25] used the Eff-CBAM (EfficientNet-B4 + CBAM) model to diagnose rice leaf diseases and achieved an accuracy of 93% on the Rice Diseases Image dataset, surpassing that of InceptionV3 and ResNet101. These results indicate that attention mechanisms significantly improve the classification of rice leaf diseases using CNNs. Van et al. [29] improved EfficientNet-B4 by integrating the LeafGabor filter and CA block for chili leaf classification and tested iton the JNUCLS and COLD datasets. The CA block enhanced both spatial and channel attention, whereas the LeafGabor filter highlighted leaf veins and reduced noise. The LGENetB4CA model achieved an accuracy of 89.61% on the JNUCLS dataset and 85.90% on the COLD dataset, outperforming other methods and demonstrating effectiveness in chili leaf classification and disease detection.

As summarised in Table 1, various DL methods have been proposed for plant disease detection in different crops, showing superior performance over traditional methods in terms of classification accuracy, adaptability to different diseases, and automation of recognition processes. Recent DL models, such as SE_SPnet [11], PDDNet [11], LGENetB4CA [29], and EfficientNet [8], have advanced significantly with the integration of pre-trained CNN models with transfer learning, achieving high accuracy rates. These studies achieve high performance as well as outperform traditional methods in disease detection, highlighting their potential in reducing crop yield losses and improving disease management efficiency. EfficientNet [27], particularly EfficientNet-B4, achieves high accuracy owing to its optimised architecture, using a compound scaling strategy to balance depth, width, and resolution [8]. With fewer parameters compared to those in ResNet or DenseNet [6], EfficientNet-B4 reduces computational load while maintaining high accuracy and generalisation across different disease types without overfitting [8]. However, EfficientNet-B4 still faces limitations when applied to leaf disease diagnosis and treatment under varying lighting conditions and environments, particularly for analysing tomato and chili leaf diseases, indicating a need for further improvements to enhance model applicability and robustness [29], [28], [27]. Although EfficientNet-B4 has incorporated SE blocks to improve channel attention, it has not fully utilised spatial attention, limiting its ability to accurately detect disease areas under varied lighting conditions or in images with heavy background noise, thus reducing its accuracy in real-world applications [10]. Singh et al. [25] demonstrated that integrating a Convolutional Block Attention Module (CBAM) into EfficientNet-B4 helps focus the attention of the model on important features for early disease detection, facilitating effective disease management and food security. However, the model still faces challenges in identifying disease areas under diverse lighting conditions or with noisy backgrounds. Thus, we explored advanced DL methods for classifying chili and tomato leaf diseases and present an improved approach to enhance classification performance. In the next section, we discuss our proposed method.

3. Methodology

This section presents in detail the proposed EfficientNet-B4-SECA model for classifying chili and tomato leaf diseases. In image classification tasks, computer vision commonly leverages pre-trained DL models trained on large-scale datasets, and then applies transfer learning to adapt these models to specific domains [15]. EfficientNet is a prominent architecture that achieves a balance between depth, width, and resolution; its performance can be further enhanced by integrating attention mechanisms such as SE, Dual Cross-Attention Module, CA, and CBAM.

Table 1

Summary of the related work.

Reference	Year	Model	Class	Dataset
Rangarajan et al. [21]	2018	AlexNet VGG16	7	PlantVillage (Tomato leaf)
Abbas et al. [2]	2021	DenseNet-121	5, 7, 10	PlantVillage (Tomato leaf)
Agarwal et al. [3]	2020	ToLeD	10	PlantVillage (Tomato leaf)
Atila et al. [8]	2020	EfficientNet	38	PlantVillage
Rukhsar et al. [22]	2022	InceptionV3	4	Rice Diseases Image Dataset
Amin et al. [6]	2022	EfficientNet-B0 DenseNet-121	4	PlantVillage (Corn leaf)
Pratap et al. [20]	2023	Fine-tuning EfficientNet-B4 Adjusting output layers Optimising parameters	5	Chili Leaf Diseases Dataset
Bhuyan et al. [11]	2023	SE_Spnet	3	Rice Leaf Disease Images
Shafik et al. [23]	2024	PDDNet	15	PlantVillage
A. Singh et al. [25]	2024	Eff-CBAM	4	Rice Diseases Image Dataset
Hoang Thien Van et al. [29]	2025	LGENetB4CA LeafGabor Filter	5	COLD Chili

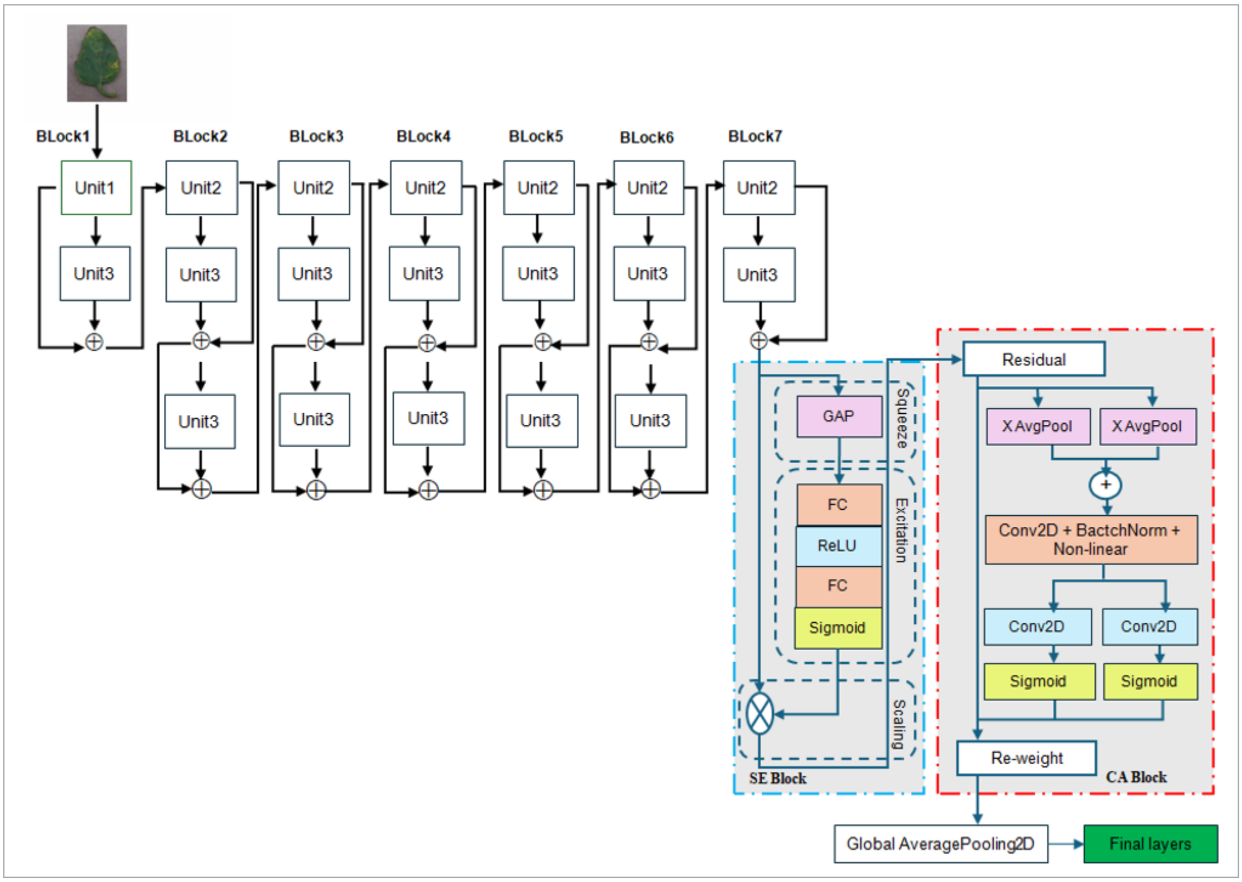
These enhancements allow the model to focus on the most informative image features, making them particularly effective for tasks such as plant species classification [8], [25], [29]. Building on this foundation, we propose the EfficientNet-B4-SECA model, which extends EfficientNet-B4 by incorporating the SECA attention module to address complex tasks such as chili and tomato leaf disease classification (Figure 1). In this architecture, the SE block [14] improves feature representation by emphasising important channels while suppressing less relevant ones. The CA block [13] complements this by introducing spatial information, which allows the model to better capture the relationships between different regions of the image. Together, these components enhance the ability of the model to recognise fine-grained patterns in leaf disease symptoms under diverse and challenging visual conditions.

3.1. EfficientNet-B4

EfficientNet is a family of CNNs, ranging from EfficientNet-B0 to EfficientNet-B7 [8], [25], [27], which are known for significantly improving accuracy with minimal increases in model parameters. This architecture employs the Swish activation function (instead of ReLU) and introduces a compound scaling method that jointly adjusts network depth, width, and resolution. These scaling coefficients are optimised using a grid search strategy to ensure efficient use of computational resources [27], [19]. EfficientNet-B4 starts with a standard convolutional layer, followed by MBConv blocks that combine depth-wise separable convolutions with SE modules to improve computational efficiency. It also employs Batch Normalisation, Global Average Pooling (GAP), Dense layers, and Dropout to optimise classification and prevent overfitting [3]. The architecture includes 24 MBConv blocks,

Figure 1

Architectural model for EfficientNet-B4-SECA.



2 convolutional layers, 1 GAP layer, 1 Dense layer, and 1 classification layer, with the MBConv1 and MBConv6 variants forming the core of the base network. The MBConv structure comprises three main stages:

1. Pointwise Convolution (1×1): Expands or reduces the number of channels based on the expansion ratio,
2. Depth-wise Convolution (k×k): Efficiently extracts spatial features, and
3. SE Block: Recalibrates channel-wise feature importance to optimise weights.

After the SE block, a final 1×1 pointwise convolution restores the feature map size. The input and output of the MBConv block maintain the same dimensions and are connected via a skip connection, which improves training stability and optimisation efficiency [3].

3.2. SE

EfficientNet-B4 integrates SE blocks into each MBConv block to enhance feature learning and boost accuracy with a minimal parameter increase, thus improving image classification performance [8]. The SE block applies channel-wise attention, allowing the network to emphasise important features by learning and adjusting channel weights [14]. The operation of the SE block comprises three main steps:

1. Spatial Information Compression (Squeeze) aggregates spatial information across the entire feature map into a single value per channel using GAP, allowing the model to assess the relative importance of each channel as follows:

$$F_{sq} = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} x_c(i, j), \tag{1}$$

where F_{sq} denotes the squeezed feature at channel c; H and W represent the height and width of the feature map, respectively. Finally, $x_c(i,j)$ is the pixel value at position (i,j) in channel c.

2. Channel Weight Adjustment (Excitation) uses two FC layers with ReLU and Sigmoid activations to learn optimal weights, determining the contribution of each channel to the final prediction as follows:

$$F_{ex} = \sigma(W_2 \delta(W_1 . GAP(x))), \tag{1}$$

where F_{ex} is the excitation feature of the channel, W_1 and W_2 are the weights of the two FC layers, δ denotes the ReLU activation function, σ is the Sigmoid function, and $GAP(x)$ represents the global average pooling of the input x.

3. Channel Weight Scaling multiplies the learned weights with each input channel, allowing the model to emphasise important features and suppress less relevant ones.

3.3. CA

The CA block enhances feature recognition by focusing on key channels and spatial locations while suppressing noise [13]. This improves the sensitivity of the model to important information leading to higher-quality feature extraction. Compared with SE and CBAM [25], CA uniquely captures channel, spatial, and directional information, making it particularly effective in tasks requiring spatial awareness. When integrated into EfficientNet-B4, the CA block improves classification accuracy and optimises spatial feature utilisation. Using one-dimensional pooling with distinct kernels for width and height, CA strengthens the ability of the model to recognise complex image patterns [29], [14]. The width and height attention signals are processed through the following steps:

1. The height tensor z^H and width tensor z^W are first concatenated, then passedthrough an activation layer F and anonlinear activation function to generate tensor s:

$$s = \delta(F([z^H, z^W])), \tag{3}$$

2. The tensor s is then split into s^H and s^W, corresponding to height and width. Each of these is processed by separate convolutional layers F_H and F_W to produce attention weights g^H and g^W. These weights refine the input feature tensor X, generating the final output:

$$u_c(i, j) = x_c(i, j) \times g_c^H(i) \times g_c^W(j), \tag{4}$$

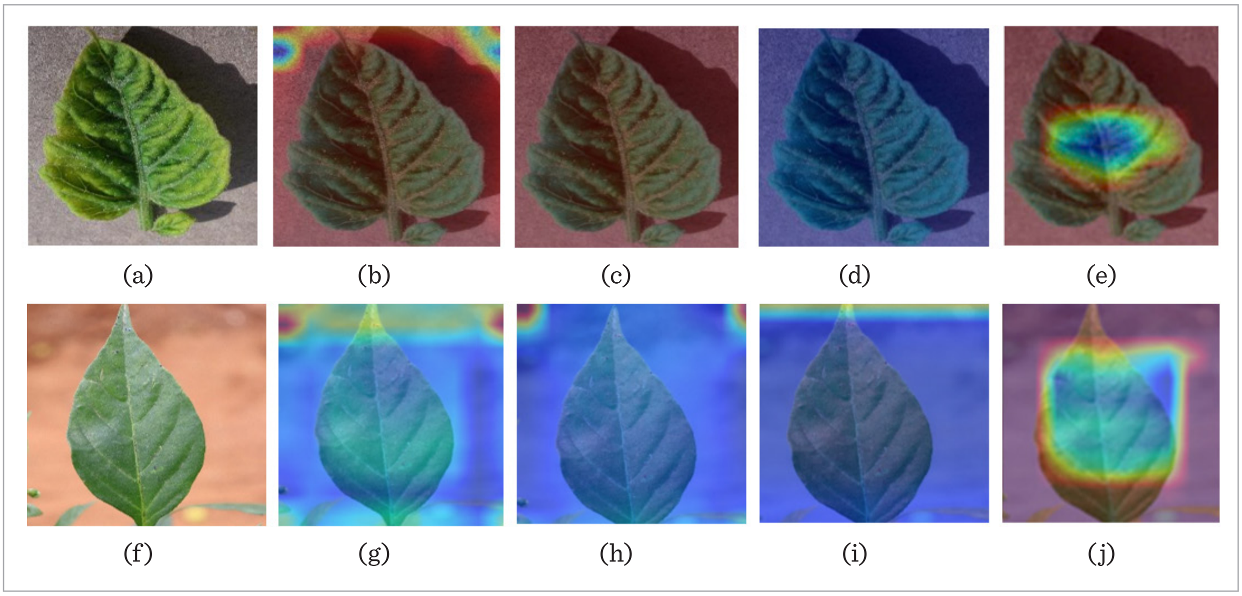
where $g_c^H(i)$ and $g_c^W(j)$ are the attention weights for height and width, respectively.

3.4. EfficientNet-B4-SECA

The EfficientNet-B4 model comprises seven blocks, each utilising MBConv, an optimised depth-wise

Figure 2

Grad-CAM visualisations using the proposed method (Dense_15 and Dense_16 layers in the CA Block): (a) Tomato leaf images; (b–e) visualisations of the block7a_se_excite, block7b_se_excite, Dense_15, Dense_16, and top_conv layers. (f) Chili leaf images; (g–j) visualisations of the same corresponding layers as indicated in the tomato example.



separable convolution that reduces parameters while maintaining efficiency. To fully leverage its capabilities, the proposed EfficientNet-B4-SECA model is initialised with pre-trained weights from ImageNet and fine-tuned for chili and tomato leaf disease classification. The model follows these processing steps (see Table 2):

1. **Block 1:** Starts with a standard convolutional layer to process 300×300-pixel input images.
2. **Blocks 2–6:** Each block includes two MBConv3 units and one MBConv2 unit, allowing the network to extract increasingly complex features.
3. **Block 7:** Refines features further using convolution, activation, and normalisation layers, followed by the integration of SE and CA blocks. The SE block enhances important channel-wise features, whereas the CA block introduces spatial awareness to capture inter-regional relationships.
4. Attention-enhanced features are passed through GAP to reduce dimensionality. This is followed by Batch Normalisation, a Dense layer with 256 units and ReLU activation, and a Dropout layer to mitigate overfitting.
5. Finally, a Dense output layer with Softmax activation generates class probabilities for classification.

Figure 2 shows the visualisations of diseased leaves from the final layers of the EfficientNet-B4-SECA model on tomato and chili leaf datasets.

Table 2

EfficientNet-B4-SECA for classifying foliar diseases of crop plants.

The modified model: EfficientNet-B4-SECA	
Layer type	**Output shape**
Input (Image Dataset)	300 × 300 × 3
EfficientNet-B4 Base Model	**10 × 10 × 1792**
Squeeze-and-Excitation (SE)	**10 × 10 × 1792**
Coordination Attention (CA)	**10 × 10 × 1792**
Global Average Pooling	1792
Batch Normalisation	1792
Dense (256 units, ReLU)	256
Dropout	256
Dense (Output classes, Softmax)	5

The overall architecture aims to strike a balance between computational efficiency and classification performance, particularly in challenging leaf disease recognition tasks. Compared with baseline models such as EfficientNet-B4 and its attention-augmented variants (e.g., CA and CBAM), the proposed EfficientNet-B4-SECA achieves higher accuracy and more stable training behaviour. This improvement stems from the complementary roles of the SE and CA blocks. While SE emphasises informative channel-wise features, CA introduces spatial awareness, enabling the model to capture long-range dependencies across the image. By contrast, models with only CA or CBAM typically enhance either spatial or channel information in isolation. Integration of both mechanisms in SECA enhances feature representation and contributes to the improved generalisation observed in our experiments.

4. Experimental Results

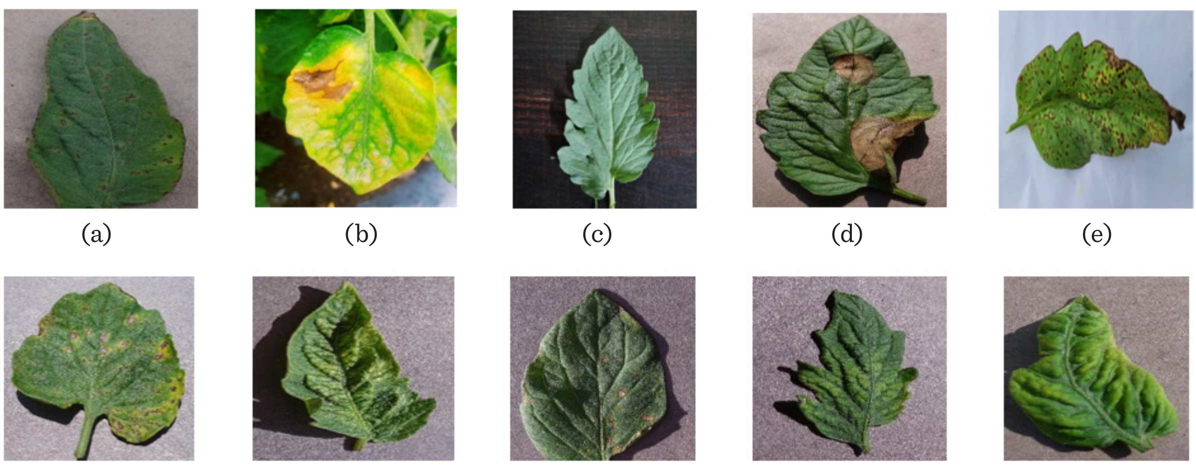
This section presents the TLD and COLD Chili datasets, along with the evaluation metrics and experimental results. We compare the performance of various DL models, including ViT, InceptionV3, EfficientNet-B4, EfficientNet-B4-CA, EfficientNet-B4-CBAM, and EfficientNet-B4-SECA, to demonstrate that EfficientNet-B4-SECA achieves a superior performance in leaf disease classification. The models were trained with a learning rate of 0.0001 and a batch size of 16 for 40 epochs, using early stopping with a patience of 5 epochs and a stopping threshold of 0.0 based on validation accuracy. K-Fold Cross-Validation (k=4)] was used to prevent overfitting and enhance generalisation, with each subset serving as the validation set once. All experiments were conducted on Google Colab using L4 GPUs. Two experiments were performed to evaluate the proposed model against state-of-the-art DL methods.

4.1. Dataset

This study uses two datasets for classifying tomato and chili leaf diseases: TLD [26] and COLD Chili [16]. The TLD dataset contains 1,646 high-resolution images (256×256 pixels) of tomato leaves captured under diverse conditions. Each image represents a distinct health status, including both healthy and diseased samples. The images were captured on a plain, uncluttered background, which improves the visibility of leaf edges and disease symptoms. The dataset was divided into 10 classes, including 1 healthy class and 9 disease categories: Bacterial Spot, Early Blight,

Figure 3

Diseases in the TLD dataset: (a) Bacterial_spot, (b) Early_blight, (c) healthy, (d) Late_blight, (e) Leaf Fungus, (f) Septoria_leaf_spot, (g) Spider_mites Two_spotted_spider_mite, (h) Target_Spot, (i) Tomato_mosaic_virus, and (j) Tomato_Yellow_Leaf_Curl_Virus.



(a) (b) (c) (d) (e)

(f) (g) (h) (i) (j)

Figure 4
Diseases in the COLD Chili Dataset: (a) Cercospora, (b) healthy, (c) mites_and_thrips, (d) nutritional deficiency, and (e) powdery mildew.

(a) (b) (c) (d) (e)

Late Blight, Leaf Mold, Septoria Leaf Spot, Spider Mite Damage, Mosaic Virus, Target Spot, and Yellow Leaf Curl Virus. These classes show significant variations in colour, texture, and lesion shape, providing a reliable foundation for training DL models for classification and disease detection tasks (Figure 3) [4], [26]. The COLD Chili Dataset comprises high-quality images of chili leaves collected during the Kharif season in Chilwadigi village, located in the Koppal district of Karnataka, India. All photographs were captured in a natural rural environment using a Canon Mark II D digital camera, ensuring high resolution and realistic conditions for model training. The dataset comprises 532 raw images that were cropped and resized to 224×224 pixels. It includes five classes: one healthy class (69 images) and four diseased classes [Cercospora (152 images), Mites and Thrips (107 images), Nutritional Deficiency (102 images), and Powdery Mildew (102 images)] (Figure 4) [16]. Although the TLD dataset is well structured, it presents challenges such as class imbalance and visual similarities among diseases.

By contrast, the COLD Chili dataset reflects environmental variations, offering additional challenges for model generalisation. Both datasets are useful for agricultural computer vision but pose different challenges. All images were resized to 300 × 300 pixels to ensure consistency. Augmentation was selectively applied to training set classes with fewer than 300 samples. Data augmentation introduces diverse examples in the training data and helps improve DL performance [29], [9]. We used techniques such as flipping, rotation, width shift, height shift, and zooming. The parameters were as follows: a rotation range of 20 degrees, width and height shifts of 0.1, and a zoom range of 0.2. Tables 3 and 4 present the properties and parameters of the datasets. After augmentation, each class in the training set contained 300 samples.

Table 3
Parametersof the COLD Chili dataset.

Class (Label)	Training Set	Testing Set	Total
Healthy	45	24	69
Cercospora	87	65	152
Mites and Trips	76	31	107
Nutritional Deficiency	68	34	102
Powdery Mildew	67	35	102

Table 4
Parameters of the TLD dataset.

Class (Label)	Training Set	Testing Set	Total
Healthy	125	32	157
Bacterial Spot	86	22	108
Early Blight	238	60	298
Late Blight	88	23	111
Leaf Fungus	300	80	380
Septoria Leaf Spot	96	25	121
Spider Mite	87	22	109
Mosaic Virus	79	20	99
Target Spot	125	32	157
Yellow Leaf Curl Virus	84	22	106

4.2. Evaluation Metrics

To evaluate the proposed model, we used the following metrics:

- **Accuracy (A):** The overall percentage of correct predictions across all classes.
- **Precision (P):** The ratio of correctly predicted samples for each class to all samples predicted for that class, indicating the false positives.
- **Recall (R):** The ratio of correctly predicted samples for each class to all actual samples of that class, indicating the false negatives.
- **F1score (F1):** The harmonic mean of Precision and Recall, providing a balance between the two, particularly useful for imbalanced datasets.
- All metrics were calculated using macro-averaging, thus treating each class equally regardless of class size as follows:

$$Avg_A = \frac{1}{classes}\sum_i^{classes} A(i),\quad A(i) = \frac{TP(i)+TN(i)}{TP(i)+TN(i)+FP(i)+FN(i)} \tag{5}$$

$$Avg_R = \frac{1}{classes}\sum_i^{classes} R(i),\quad R(i) = \frac{TP(i)}{FN(i)+TP(i)} \tag{6}$$

$$Avg_P = \frac{1}{classes}\sum_i^{classes} P(i),\quad P(i) = \frac{TP(i)}{TP(i)+FP(i)} \tag{7}$$

$$Avg_{F_1} = \frac{1}{classes}\sum_i^{classes} F_1(i),\quad F_1(i) = 2\times\frac{Pr(i)\times R(i)}{Pr(i)+R(i)}, \tag{8}$$

where, with class i, TP is True Positive, TN is True Negative, FP is False Positive, and FN is False Negative.

4.3. Experimental Results on the TLD Dataset

In this experiment, we trained models on the TLD dataset to evaluate the classification performance of state-of-the-art DL architectures (InceptionV3, ViT, EfficientNet-B4, EfficientNet-B4-CA, Effi-

Figure 5

Feature visualisations using Grad-CAM: (a) Original image from the TLD dataset; (b) conv2d_84 feature map of InceptionV3; (c) last layer feature map of ViT; top_conv feature maps of (d) EfficientNet-B4, (e) EfficientNet-B4-CA, and (g) EfficientNet-B4-SECA; and (f) conv2d feature map from the CBAM block in EfficientNet-B4-CBAM.

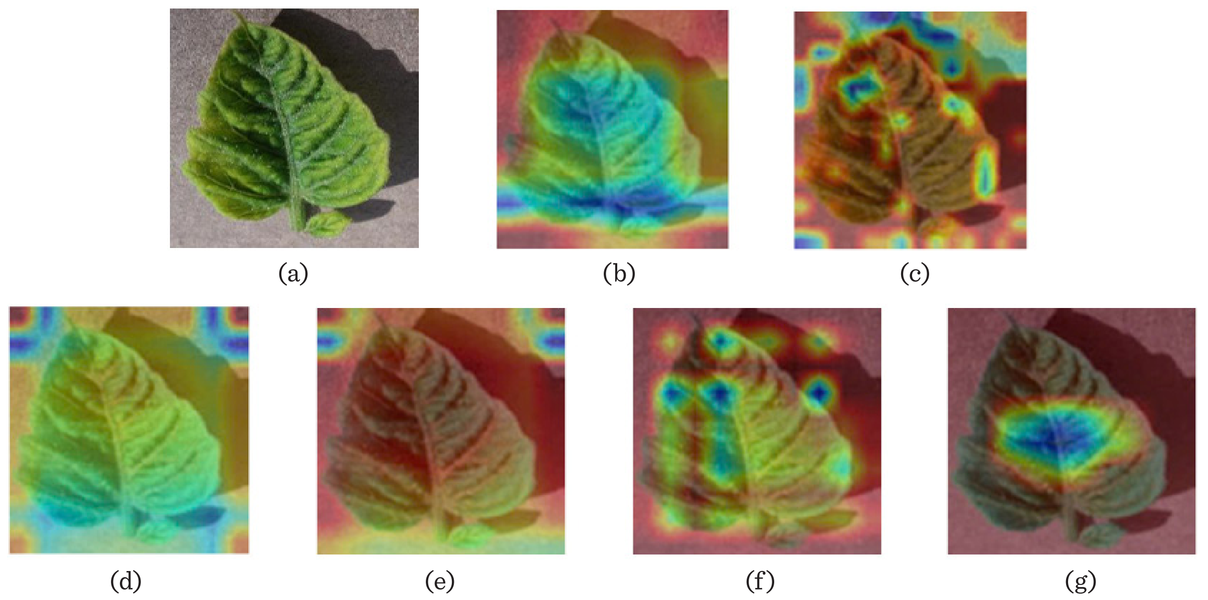
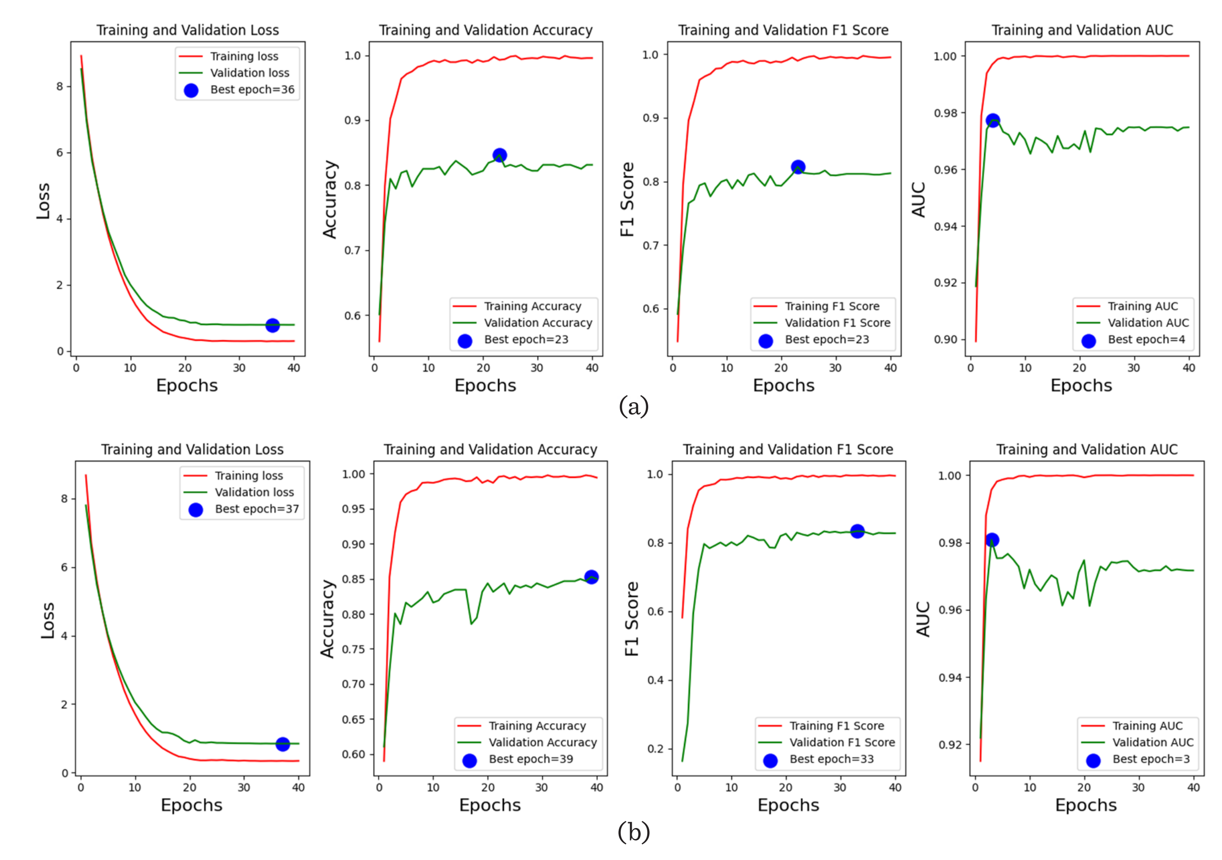


Figure 6

Training and evaluation performance of the models on the TLD dataset (Fold 1), including (a) EfficientNet-B4 and (b) EfficientNet-B4-SECA.



cientNet-B4-CBAM) and our proposed model, EfficientNet-B4-SECA. To improve generalisation and minimise overfitting, we applied a 4-fold cross-validation, with each fold used once for validation and the remaining three used for training. Figure 5 presents the feature visualisations generated using Grad-CAM for the evaluated models. Figure 6 compares the training behaviour of EfficientNet-B4 (a) and EfficientNet-B4-SECA (b) on the TLD dataset (Fold 1). Both models showed decreasing loss curves, but EfficientNet-B4-SECA achieved better generalisation. Its validation accuracy, F1 score, and area under the curve (AUC) were more stable and closely followed the training curves. Contrastingly, EfficientNet-B4 showed larger gaps between training and validation metrics, indicating overfitting. The integration of SE and CA blocks helped improve model robustness and reduced overfitting.

Table 5 presents the cross-validation accuracy results for the evaluated methods on the TLD dataset. The proposed EfficientNet-B4-SECA model achieved accuracies ranging from 82.66% to 84.06%, demonstrating consistently strong performance across all folds. It recorded the highest average accuracy of 83.24%, outperforming other models such as EfficientNet-B4-CBAM (82.78%), the original EfficientNet-B4, InceptionV3, and ViT. In addition to its superior accuracy, EfficientNet-B4-SECA also showed excellent stability, with a low standard deviation (STD = 0.68), which was significantly lower than that of EfficientNet-B4 (0.85), InceptionV3 (2.36), and ViT (1.52). This low variation across folds indicates a more consistent and reliable training process. Furthermore, the model achieved a narrow and high 95% confidence interval [82.16, 84.32], reinforcing its robustness and generalisability.

Table 5

Four-fold cross-validation accuracy of various methods on the TLD Dataset (%).

Methods	Fold1	Fold2	Fold3	Fold4	Mean	STD-DEV	Confidence interval (CI)
ViT	77.09	80.43	80.00	79.87	79.35	1.52	[76.92, 81.77]
InceptionV3	77.71	83.23	81.56	81.76	81.07	2.36	[77.31, 84.82]
EfficientNet-B4	81.11	82.3	82.81	81.13	81.84	0.85	[80.48, 83.20]
EfficientNet-B4-CBAM	82.35	82.3	83.12	83.33	82.78	0.53	[81.94, 83.61]
EfficientNet-B4-CA	81.73	82.61	83.12	82.39	82.46	0.58	[81.55, 83.38]
EfficientNet-B4-SECA	82.66	83.54	84.06	82.70	83.24	0.68	[82.16, 84.32]

Table 6

Comparing the performance of DL models on the TLD Dataset.

Methods	Accuracy (%)	F1 score (%)	Training time (second)	Test time (second/sample)
ViT	80.06	80.16	1765.92	0.0439
InceptionV3	82.82	82.76	8936.27	0.0491
EfficientNet-B4	84.05	84.05	20096.98	0.0828
EfficientNet-B4-CBAM	84.05	84.01	19564.22	0.0767
EfficientNet-B4-CA	84.66	84.73	19767.77	0.0859
EfficientNet-B4-SECA	**84.97**	**84.86**	**19673.85**	**0.0920**

As shown in Table 6, EfficientNet-B4-SECA achieved the highest classification accuracy (84.97%) and F1 score (84.86%), outperforming baseline EfficientNet-B4 (84.05%), EfficientNet-B4-CBAM (84.05%), EfficientNet-B4-CA (84.66%), InceptionV3 (82.82%) and ViT (80.06%). These results show that EfficientNet-B4 combined with the CBAM block or the CA block achieves higher accuracy than the baseline EfficientNet-B4, and EfficientNet-B4 integrated with both the SE and CA blocks achieves the highest accuracy. The training time (19,673.85 seconds) and testing time per sample (0.0920 seconds) of EfficientNet-B4-SECA were only slightly higher than those of EfficientNet-B4-CA (19,767.77 and 0.0859 seconds, respectively). By contrast, the InceptionV3 model and particularly ViT, exhibited significantly faster training and testing times. However, their classification accuracy was notably lower, with ViT and InceptionV3 achieving an accuracy of 80.06% and 82.82%, respectively, both falling well behind the accuracy of EfficientNet-B4-SECA (84.97%). Figures 7, 8, 9, and 10 illustrate the accuracy of the proposed model compared to that of other methods, using the confusion matrix and AUC-receiver operating characteristic (ROC) curve. Figure 9b presents the confusion matrix, which highlights that the proposed model achieved a perfect classification accuracy for disease classes with prominent visual characteristics, such as Spider Mites (25/25), Tomato Mosaic Virus (22/22), and Tomato Yellow Leaf Curl Virus (20/20). Additionally, the proposed method demonstrated a higher success rate for recognising the Leaf Fungus class (67/78) compared to that with other methods such as ViT (55/78), EfficientNet-B4 (61/78), and EfficientNet-CBAM (64/78), underscoring its strong generalisation ability. However, the confusion matrix also revealed some classification ambiguities between disease classes with visually similar symptoms, such as Early Blight and Healthy. Specifically, the proposed model classified Early Blight with an accuracy of 45/60, which was slightly lower than that of ViT (48/60), EfficientNet-B4 (46/60), and EfficientNet-CBAM (46/60). The proposed model achieved an accuracy of 21/32 for Healthy, which is slightly lower than that of EfficientNet-B4 (23/32), and higher than ViT (19/32), EfficientNet-CBAM (20/32). This difference may be attributed to the subtle visual distinctions between healthy leaves and leaves with early-stage infection.

Figure 7

Confusion matrix of the (a) ViT and (b) EfficientNet-B4 models on the TLD Dataset.

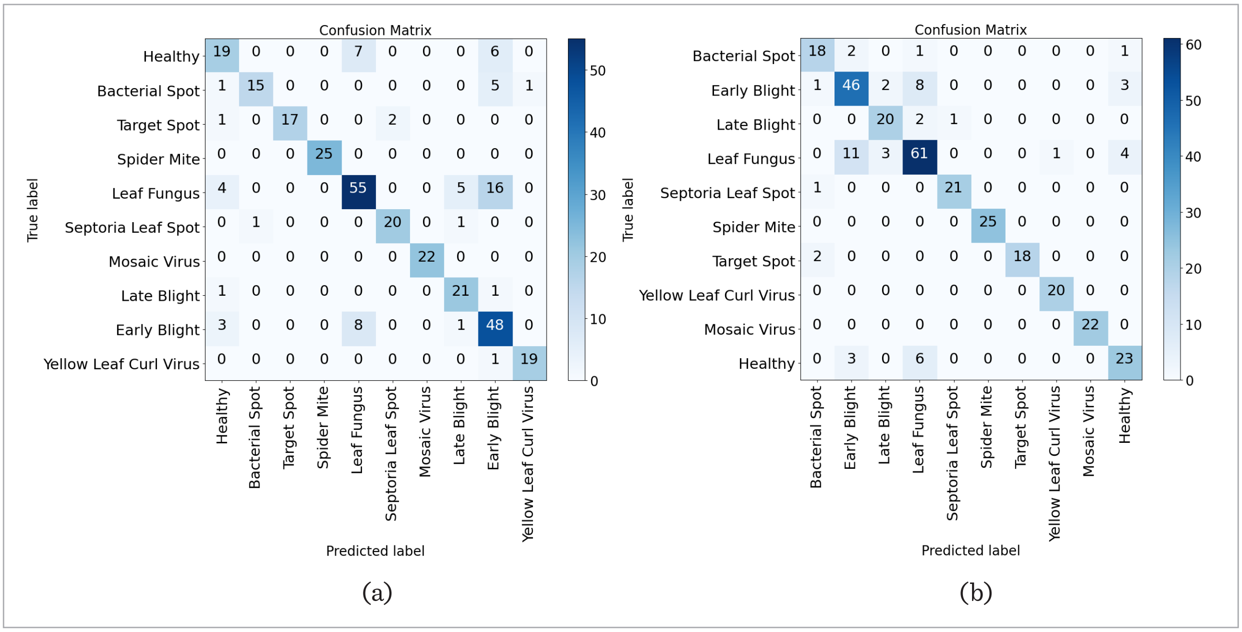


Figure 8

AUC-ROC of the (a) ViT and (b) EfficientNet-B4 models on the TLD Dataset.

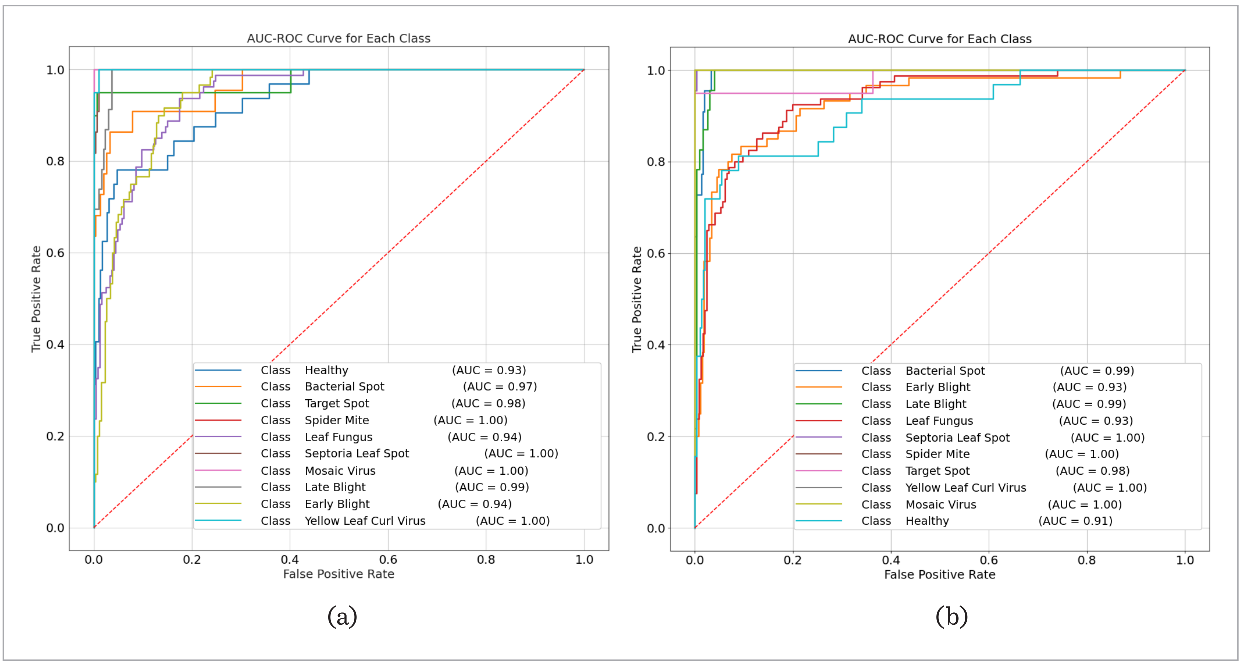
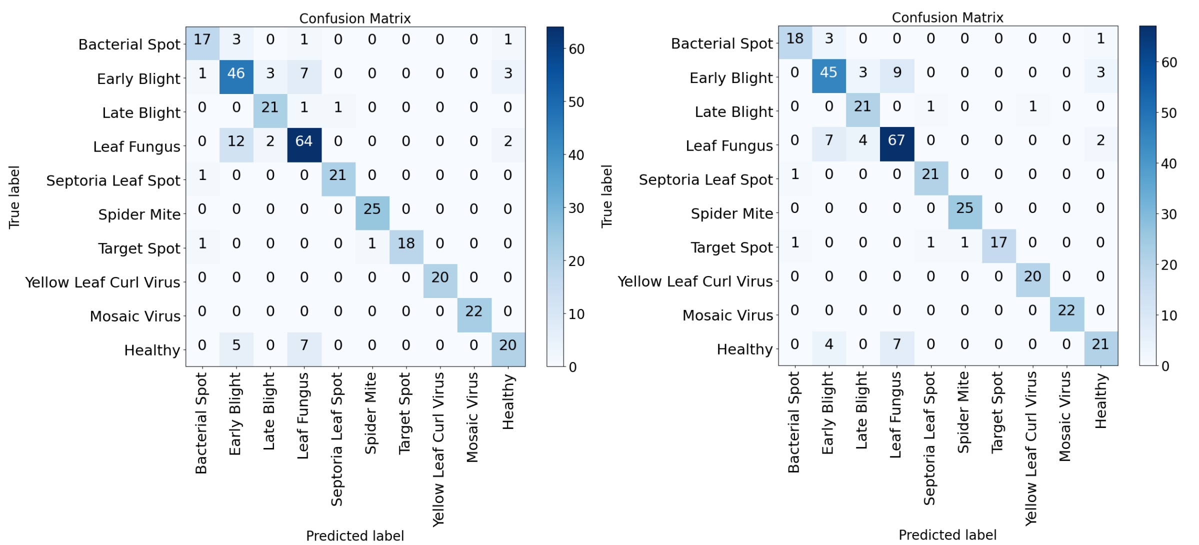


Figure 9

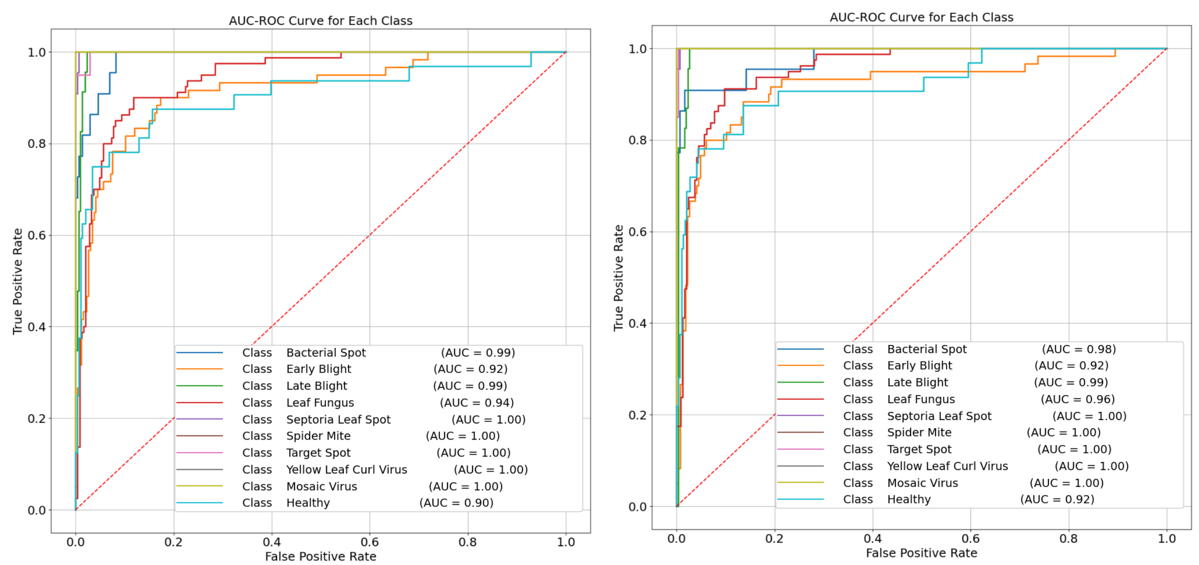
Confusion matrix of the (a) EfficientNet-B4-CBAM and (b) EfficientNet-B4-SECA models on the TLD Dataset.



(a) (b)

Figure 10

AUC-ROC of the (a) EfficientNet-B4-CBAM and (b) EfficientNet-B4-SECA models on the TLD Dataset.



(a) (b)

Overall, the experimental results demonstrate that the EfficientNet-B4-SECA model outperforms other methods in terms of accuracy and F1score, as well as exhibits effective class discrimination and strong generalisation capabilities. These findings highlight the potential of this model for real-world applications in detecting plant diseases.

4.4. Experimental Results on the COLD Chili Dataset

In the second experiment, we evaluated the performance of various DL models on the COLD Chili dataset. Figure 11 presents the feature visualisations of models, generated using GradCAM. Figure 12 compares the training dynamics of EfficientNet-B4-CA (a) and EfficientNet-B4-SECA (b) on the COLD Chili dataset (Fold 2). Both models exhibit steadily decreasing training and validation loss curves, indicating effective convergence. However, EfficientNet-B4-SECA demonstrates superior generalisation, as evidenced by a higher validation accuracy, F1score, and AUC. By contrast, EfficientNet-B4-CA shows larger gaps between training and validation metrics, particularly in the F1score and AUC, suggesting potential overfitting. This highlights the benefit of incorporating SE and CA attention mechanisms together. Table 7 presents the K-fold cross-validation results, wherein the

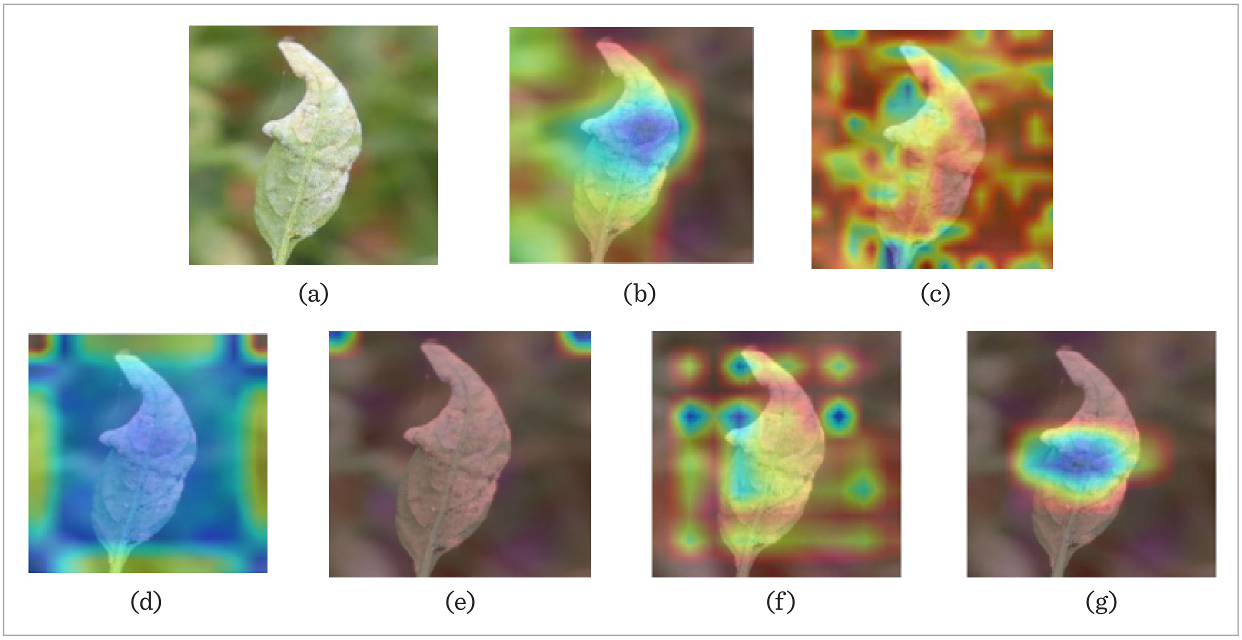
Table 7

Four-fold cross-validation accuracy of different methods on the COLD Chili Dataset (%).

Methods	Fold1	Fold2	Fold3	Fold4	Mean	STD	Confidence interval (CI)
ViT	80.56	82.86	89.71	83.82	84.24	3.90	[78.04, 90.44]
InceptionV3	84.72	82.86	79.41	83.82	82.70	2.32	[79.01, 86.40]
EfficientNet-B4	86.11	84.29	88.24	80.88	84.88	3.12	[79.92, 89.84]
EfficientNet-B4-CBAM	87.5	85.71	83.82	77.94	83.74	4.15	[77.14, 90.35]
EfficientNet-B4-CA	81.94	88.57	86.76	83.82	85.27	2.96	[80.56, 89.98]
EfficientNet-B4-SECA	**84.72**	**88.57**	**83.82**	**85.29**	**85.60**	**2.07**	**[82.31, 88.89]**

Figure 11

Feature visualisations using Grad-CAM: (a) Original image from the Chili Leaf Disease dataset; (b) conv2d_84 feature map of InceptionV3; (c) last layer feature map of ViT; top_conv feature maps of (d) EfficientNet-B4, (e) EfficientNet-B4-CA, and (g) EfficientNet-B4-SECA; and (f) conv2d feature map from the CBAM block in EfficientNet-B4-CBAM.



(a) (b) (c) (d) (e) (f) (g)

Figure 12

Training and evaluation performance of the proposed model on each fold of the COLD Chili Dataset.

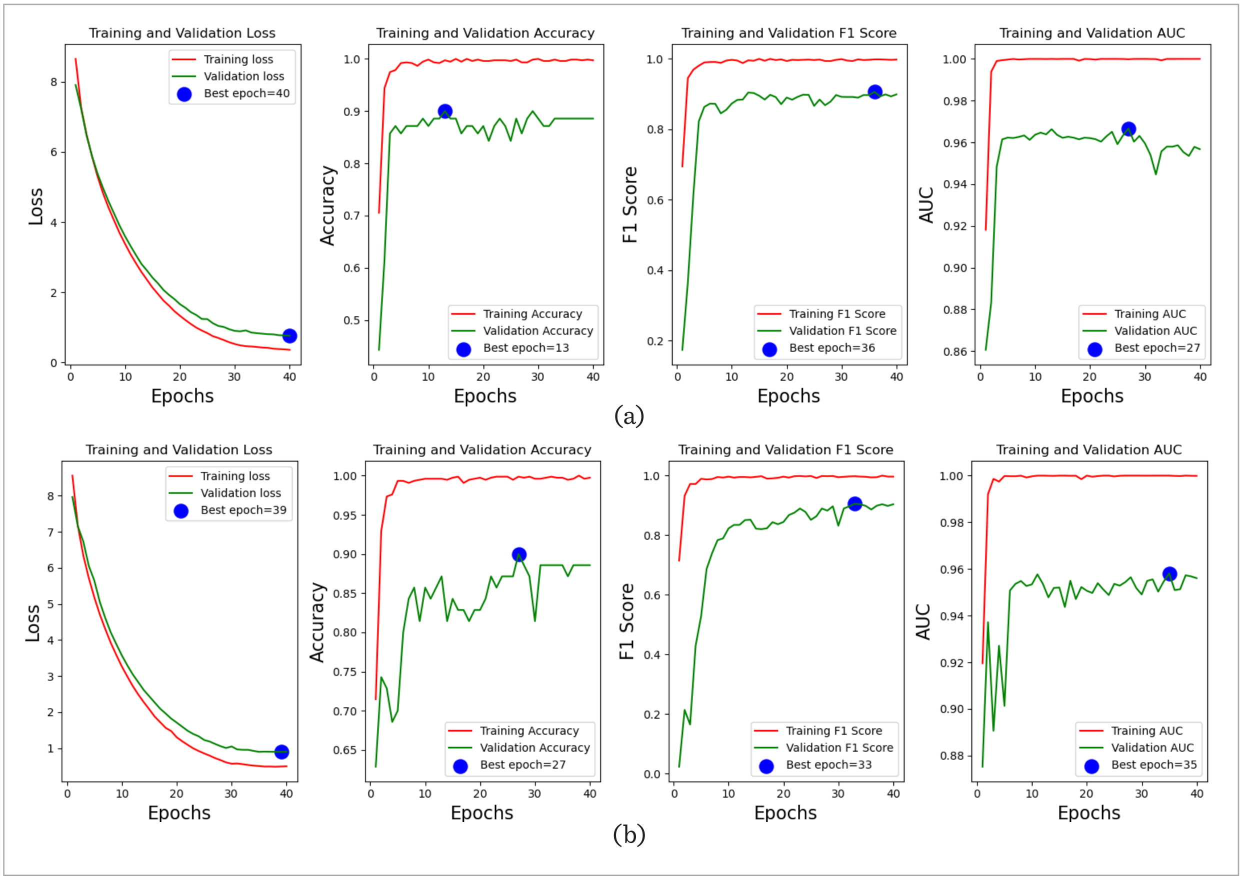


Table 8

Comparing the performance of various DL models on the COLD Chili Dataset.

Methods	Accuracy (%)	F1 score (%)	Training time (second)	Test time (second/sample)
ViT	83.07	80.06	1169.59	0.0260
InceptionV3	77.78	78.33	4205.33	0.0265
EfficientNet-B4	83.60	82.86	9571.44	0.0429
EfficientNet-B4-CBAM	85.71	85.65	9933.02	0.0423
EfficientNet-B4-CA	86.24	86.16	9643.66	0.0529
EfficientNet-B4-SECA	**87.83**	**87.9**	**9652.30**	**0.0529**

proposed EfficientNet-B4-SECA model achieved the highest average accuracy of 85.60%, outperforming ViT (84.24%), InceptionV3 (82.70%), EfficientNet-B4 (84.88%), EfficientNet-B4-CBAM (83.74%), and EfficientNet-B4-CA (85.27%). Its low STD (2.07) and narrow 95% confidence interval [82.31, 88.89] indicate a consistent performance and strong generalisation across folds.

Table 8 presents the performance of various DL models on the COLD Chili dataset. Figures 13–16 illustrate the confusion matrices and AUC-ROC curves for InceptionV3, EfficientNet-B4, EfficientNet-B4-CA, and EfficientNet-B4-SECA, respectively. Among all evaluated models (Table 8), EfficientNet-B4-SECA achieved the highest classification accuracy of 87.83%, outperforming the baseline EfficientNet-B4 (83.60%), EfficientNet-B4-CBAM (85.71%), EfficientNet-B4-CA (86.24%), InceptionV3 (77.78%) and ViT (83.07%). These results demonstrate that combining the SE and CA mechanisms enhances the ability of the model to extract spatial and contextual features. This improvement allows more accurate classification of chili leaf diseases, particularly in challenging scenarios involving complex lighting conditions, diverse backgrounds, and variations in leaf shape. Figures 13, 14, 15, and 16 illustrate the accuracy of the proposed model compared to that of other methods, using the confusion matrix and AUC-receiver operating characteristic (ROC) curve. EfficientNet-B4-SECA also showed a robust performance across multiple disease categories. As shown in the confusion matrix (Figure 15b), the proposed model correctly classified most samples from critical classes, including Powdery Mildew (35 out of 35), Cercospora (51 out of 65), and Nutritional Deficiency (33 out of 34). However, some confusion was observed between visually similar classes, such as Healthy and Mites_and_thrips, indicating that further improvements are requiredfor fine-grained classification. The AUC-ROC curve in Figure 16b supports the effectiveness of the proposed model effectiveness, showing high AUC scores for all classes: Powdery Mildew (1.00), Mites_and_thrips (0.98), Nutritional Deficiency (0.97), Healthy (0.95), and Cercospora (0.94). These scores confirm the high precision and confidence of the model in prediction.

Regarding computational efficiency, the training time of EfficientNet-B4-SECA was 9,652.3 seconds, which was comparable to that ofEfficientNet-B4 (9,571.44 seconds), EfficientNet-B4-CBAM (9,933.02 seconds), and EfficientNet-B4-CA (9,643.66 seconds). The testing time per sample was 0.0529 seconds, which was equal to that of EfficientNet-B4-CA, and slightly higher than EfficientNet-B4-CBAM (0.0423 seconds), EfficientNet-B4 (0.0429 seconds), InceptionV3 (0.0265 seconds), and ViT (0.0260 seconds). These results indicate that, despite incorporating two attention mechanisms, EfficientNet-B4-SECA maintains reasonable training and inference times

Figure 13

Confusion matrix of the (a) InceptionV3 and (b) EfficientNet-B4 models on the COLD chili dataset.

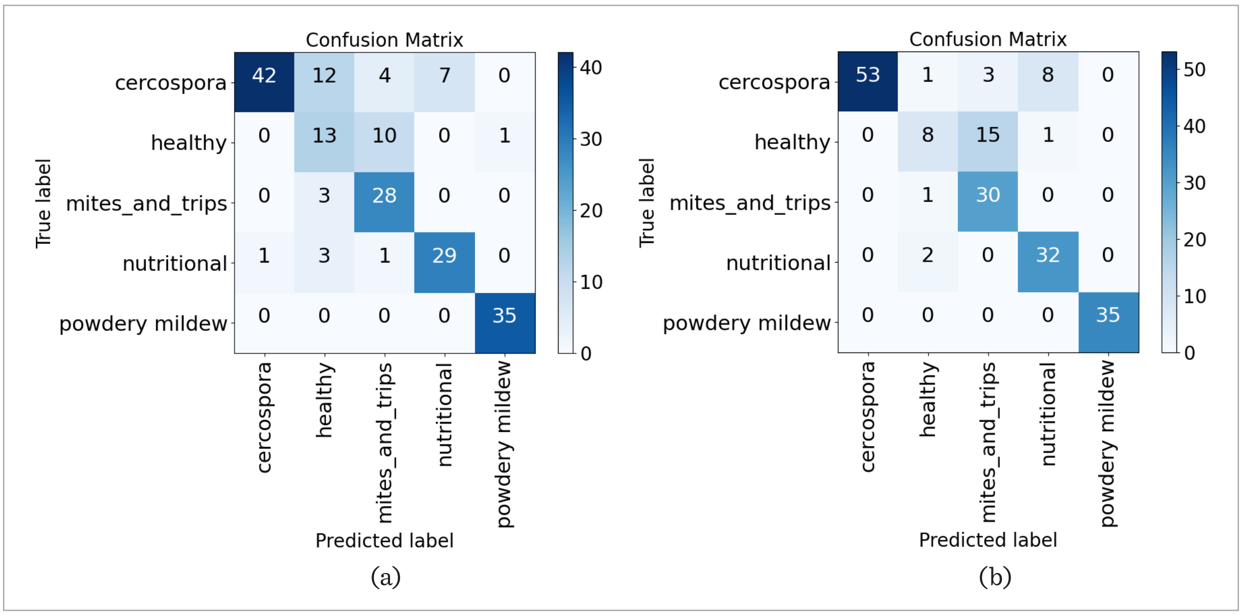


Figure 14

AUC-ROC of the (a) ViT and (b) EfficientNet-B4 models on the COLD chili dataset.

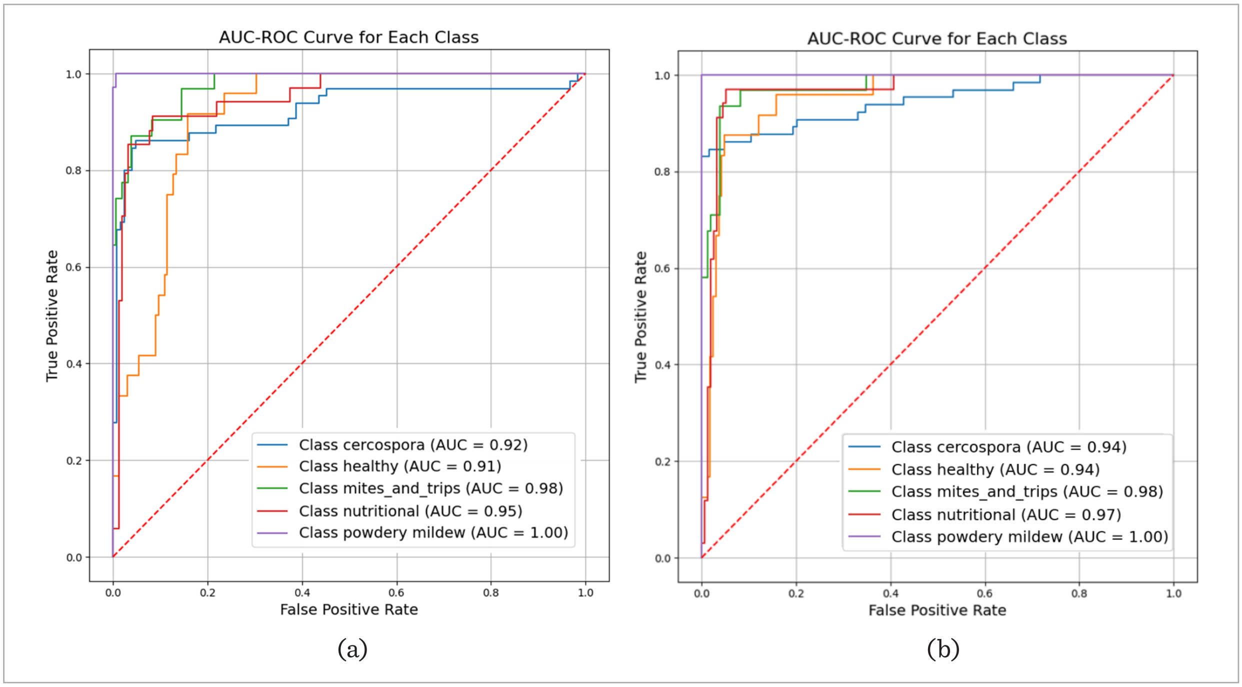
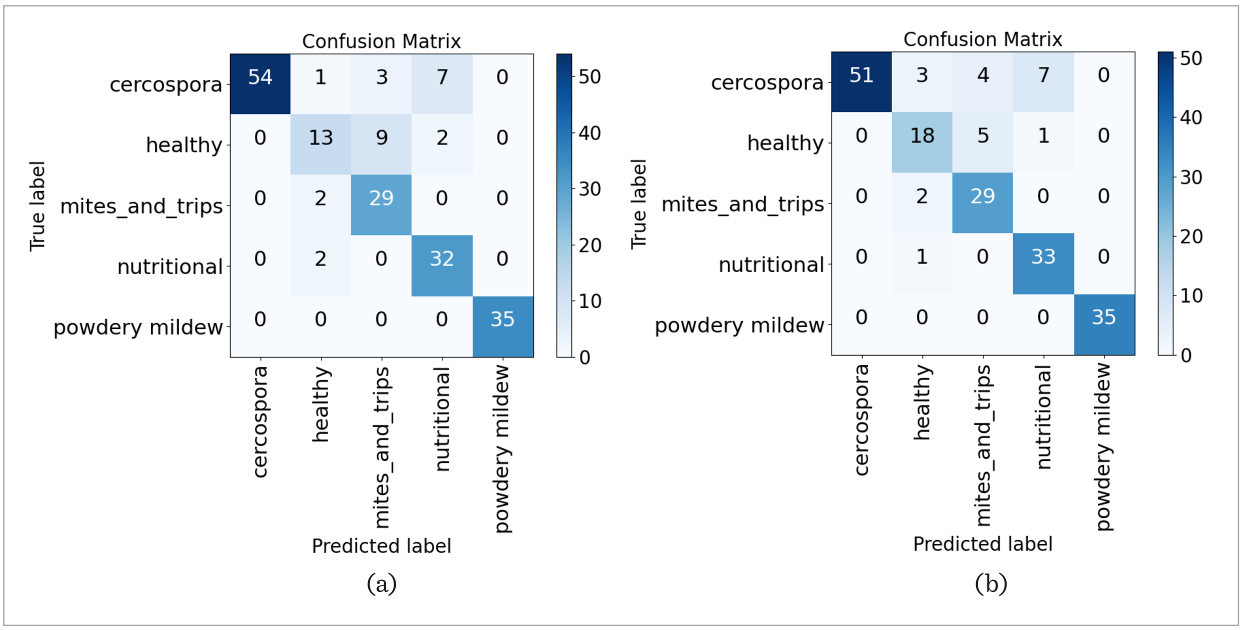


Figure15

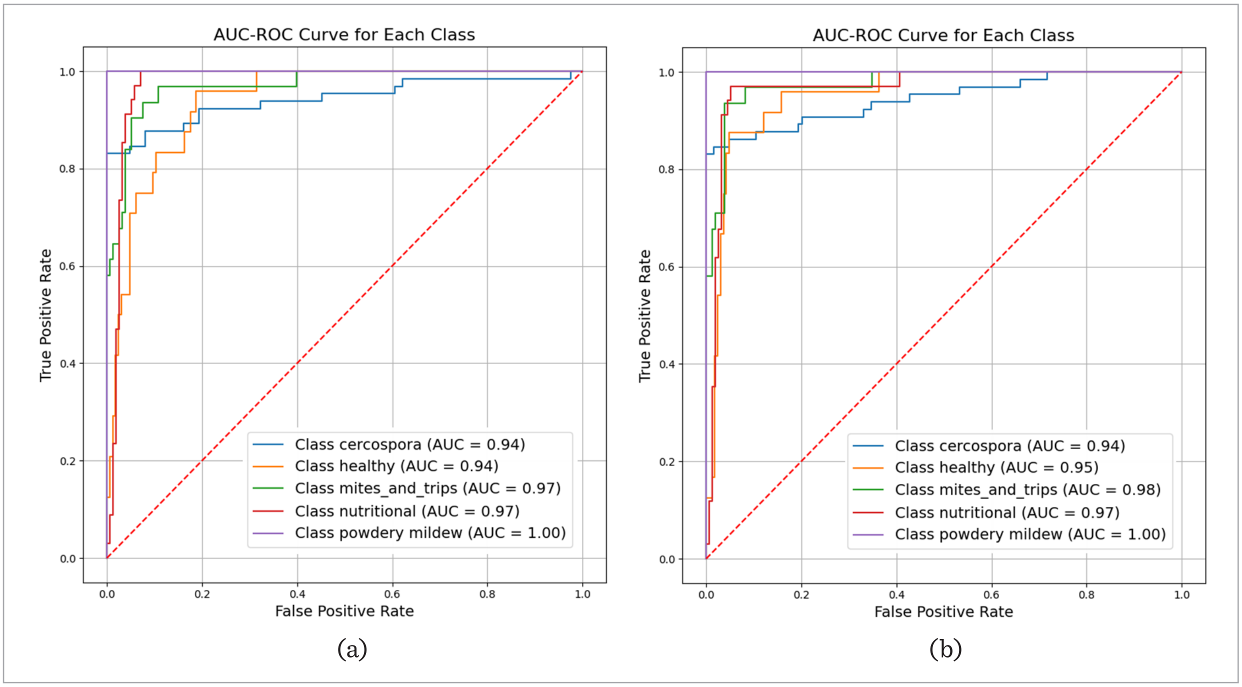
Confusion matrix of the (a) EfficientNet-B4-CA and (b) EfficientNet-B4-SECA models on the COLD chili dataset.



while delivering a superior classification performance. Overall, the proposed EfficientNet-B4-SECA model demonstrates high accuracy, strong generalisation, and efficient computation, making it a promising solution for real-world chili leaf disease classification tasks.

Figure 16

AUC-ROC of the (a) EfficientNet-B4-CA and (b) EfficientNet-B4-SECA models on the COLD chili dataset.



4.5. Discussion

In this study, we investigate the challenge of accurately classifying foliar diseases in tomatoes and chilies—two widely cultivated crops. The variability in lesion appearance, including differences in shape, texture, and colour, pose significant difficulties for automated detection systems. We thus evaluated several DL models on two datasets with varying visual complexity: TLD and COLD Chili. The experimental results showed that the proposed model achieved a superior performance compared to that of existing methods. Particularly, it attained an accuracy of 84.97% on the TLD Dataset and 87.83% on the COLD Chili Dataset. These results surpass those of ViT, InceptionV3, EfficientNet-B4, EfficientNet-B4-CA, and EfficientNet-B4-CBAM. The improvement confirmed that the combination of SE and CA mechanisms enhanced both spatial feature extraction and contextual understanding. However, the model still showed occasional misclassifications, particularly in classes with visually similar symptoms, suggesting the need for further enhancements. Although the total training and testing durations were slightly longer than some baseline models, the per-sample inference time remained efficient with 0.0920 seconds on the tomato dataset and 0.0529 seconds on the chili dataset. This high level of classification accuracy, combined with an acceptable inference speed, highlights the potential of the model for deployment in real-world agricultural applications such as mobile or drone-based monitoring platforms, where accurate and reliable plant disease diagnosis is essential for timely intervention and effective crop management.

5. Conclusion

In this study, we investigated DL approaches for classifying leaf diseases in tomato and chili plants. We proposed the EfficientNet-B4-SECA model, which enhances the EfficientNet-B4 architecture by integrating the SE and CA mechanisms. These modules improve the ability of the model to focus on relevant features and capture spatial dependencies critical for disease recognition. Experiments conducted on the TLD and COLD Chili datasets showed that EfficientNet-B4-SECA achieved classification accuracies of 84.97% and 87.83%, respectively. It outperformed several strong baseline models, including ViT, InceptionV3, EfficientNet-B4, and its attention-augmented variants. These results demonstrate the potential of the proposed model to support automated plant disease diagnosis in precision agriculture.

Prediction accuracy can be enhanced by future research and developing more advanced DL architectures integrated with improved attention mechanisms. Further, we aim to apply this approach to broader plant recognition tasks using the images of leaves, fruits, and flowers. The goal is to contribute to the development of intelligent agricultural systems that support early disease detection, quality assessment, and sustainable crop management.

Conflict of Interest

The authors declare no conflict of interest.

References

1. Abayomi-Alli, O. O., Damaševičius, R., Misra, S., Maskeliūnas, R. Cassava Disease Recognition from Low-Quality Images Using Enhanced Data Augmentation Model and Deep Learning. Expert Systems, 2021, 38(7), e12746. https://doi.org/10.1111/exsy.12746
2. Abbas, A., Jain, S., Gour, M., Vankudothu, S. Tomato Plant Disease Detection Using Transfer Learning with C-GAN Synthetic Images. Computers And Electronics in Agriculture, 2021, 187, 106279. https://doi.org/10.1016/j.compag.2021.106279
3. Agarwal, M., Singh, A., Arjaria, S., Sinha, A., Gupta, S. ToLeD: Tomato Leaf Disease Detection Using Convolutional Neural Network. Procedia Computer Science, 2020, 167, 293-301. https://doi.org/10.1016/j.procs.2020.03.225
4. Ahmed Imtiaz, A., Swapnil, F. B. I., Masud, S. R., Karmaker, D. Tomato Leaf Dataset: A Dataset for Multiclass Disease Detection and Classification. Data In Brief, 2025, 60, 111520. https://doi.org/10.1016/j.dib.2025.111520
5. Al-Shamasneh, A. R., Ibrahim, R. W. Classification of Tomato Leaf Images for Detection of Plant Disease Using Conformable Polynomials Image Features. MethodsX, 2024, 13, 102844. https://doi.org/10.1016/j.mex.2024.102844
6. Amin, H., Darwish, A., Hassanien, A. E., Soliman, M. End-To-End Deep Learning Model for Corn Leaf Disease Classification. IEEE Access, 2022, 10, 31103-31115. https://doi.org/10.1109/ACCESS.2022.3159678
7. Ashok, S., Kishore, G., Rajesh, V., Suchitra, S., Sophia, S. G. G., Pavithra, B. Tomato Leaf Disease Detection Using Deep Learning Techniques. In International Conference on Communication and Electronics Systems (ICCES), 2020, 979-983. https://doi.org/10.1109/ICCES48766.2020.9137986
8. Atila, Ü., Uçar, M., Akyol, K., Uçar, E. Plant Leaf Disease Classification Using EfficientNet Deep Learning Model. Ecological Informatics, 2021, 61, 101182. https://doi.org/10.1016/j.ecoinf.2020.101182
9. Bello, I., Fedus, W., Du, X., Cubuk, E. D., Srinivas, A., Lin, T.-Y., Shlens, J., Zoph, B. Revisiting ResNets: Improved Training and Scaling Strategies. Advances In Neural Information Processing Systems, 2021, 34, 22614-22627. https://doi.org/10.48550/arXiv.2103.07579
10. Bezabih, Y. A., Salau, A. O., Abuhayi, B. M., Mussa, A. A., Ayalew, A. M. CPD-CCNN: Classification of Pepper Disease Using a Concatenation of Convolutional Neural Network Models. Scientific Reports, 2023, 13, 15581. https://doi.org/10.1038/s41598-023-42843-2
11. Bhuyan, P., Singh, P. K., Das, S. K., Kalla, A. SE_SPNet: Rice Leaf Disease Prediction Using Stacked Parallel Convolutional Neural Network with Squeeze-and-Excitation. Expert Systems, 2023, 40(7), e13304. https://doi.org/10.1111/exsy.13304
12. Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weisseborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., Houlsby, N. An Image Is Worth 16x16 Words: Transformers for Image Recognition at Scale. In International Conference on Learning Representations, 2021. https://doi.org/10.48550/arXiv.2010.11929
13. Hou, Q., Zhou, D., Feng, J. Coordinate Attention for Efficient Mobile Network Design. In IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, 13708-13717. https://doi.org/10.1109/CVPR46437.2021.01350
14. Hu, J., Shen, L., Sun, G. Squeeze-and-Excitation Networks. In IEEE Conference on Computer Vision and Pattern Recognition, 2018, 7132-7141. https://doi.org/10.1109/CVPR.2018.00745

15. Li, L., Zhang, S., Wang, B. Plant Disease Detection and Classification by Deep Learning-A Review. IEEE Access, 2021, 9, 56683-56698. https://doi.org/10.1109/ACCESS.2021.3069646

16. M. P., Aishwarya; Reddy, P. Chilli Dataset. Mendeley Data, 2024, V2. https://doi.org/10.17632/tf9dtfz9m6.2

17. Mukti, I. Z., Biswas, D. Transfer Learning Based Plant Diseases Detection Using ResNet50. In International Conference on Electrical Information and Communication Technology (EICT), 2019, 1-6. https://doi.org/10.1109/EICT48899.2019.9068805

18. Nageswararao Naik, B., Malmathanraj, R., Palanisamy, P. Detection and Classification of Chilli Leaf Disease Using a Squeeze-and-Excitation-Based CNN Model. Ecological Informatics, 2022, 69, 101663. https://doi.org/10.1016/j.ecoinf.2022.101663

19. Padhi, J., Korada, L., Dash, A., Sethy, P. K., Behera, S. K., Nanthaamornphong, A. Paddy Leaf Disease Classification Using EfficientNetB4 with Compound Scaling and Swish Activation: A Deep Learning Approach. IEEE Access, 2024, 12, 126426-126437. https://doi.org/10.1109/ACCESS.2024.3451557

20. Pratap, V. K., Kumar, N. S. High-Precision Multiclass Classification of Chili Leaf Disease Through Customized EfficientNetB4 from Chili Leaf Images. Smart Agricultural Technology, 2023, 5. https://doi.org/10.1016/j.atech.2023.100295

21. Rangarajan, A. K., Purushothaman, R., Ramesh, A. Tomato Crop Disease Classification Using Pre-Trained Deep Learning Algorithm. Procedia Computer Science, 2018, 133, 1040-1047. https://doi.org/10.1016/j.procs.2018.07.070

22. Rukhsar, Upadhyay, S. K. Deep Transfer Learning-Based Rice Leaves Disease Diagnosis and Classification Model Using InceptionV3. In International Conference on Computational Intelligence and Sustainable Engineering Solutions (CISES), 2022, 493-499. https://doi.org/10.1109/CISES54857.2022.9844374

23. Shafik, W., Tufail, A., De Silva Liyanage, C., Apong, R. A. A. H. M. Using Transfer Learning-Based Plant Disease Classification and Detection for Sustainable Agriculture. BMC Plant Biology, 2024, 24, 136. https://doi.org/10.1186/s12870-024-04825-y

24. Singh, V., Sharma, N., Singh, S. A Review of Imaging Techniques for Plant Disease Detection. Artificial Intelligence in Agriculture, 2020, 4, 229-242. https://doi.org/10.1016/j.aiia.2020.10.002

25. Singh, A., Kumar, V. A Deep Learning Attention Approach for Diagnosis of Rice Diseases. In International Conference on Electronics, Communication and Signal Processing (ICECSP), 2024, 1-5. https://doi.org/10.1109/ICECSP61809.2024.10698030

26. Solapure, V., Jawale, A. Tomato Leaf Disease Dataset. Mendeley Data, 2024, V1. https://doi.org/10.17632/zfv4jj7855.1

27. Tan, M., Le, Q. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. In International Conference on Machine Learning, 2019, 97, 6105-6114. ttps://doi.org/10.48550/arXiv.1905.11946

28. Tanim, S. A., Aurnob, A. R., Anik, Z. H., Hossain, M. I. Precise Detection of Tomato Leaf Diseases Using Deep Learning Approach with EfficientNet. In International Conference on Computer and Information Technology (ICCIT), 2023, 1-6. https://doi.org/10.1109/ICCIT60459.2023.10441130

29. Van, H. T., Vu, G. V., Thai, T. T., Vo, B., Chung, Y. S. LGENetB4CA: A Novel Deep Learning Approach for Chili Germplasm Differentiation and Leaf Disease Classification. Computers And Electronics in Agriculture, 2025, 233, 110149. https://doi.org/10.1016/j.compag.2025.110149