REVIEW ARTICLE

Agricultural Science Digest, Volume 46 Issue 4: 553-563 (August 2026)

A Hybrid Approach using ResNet 50 and EfficientNetB0 for Attention-enhanced Deep Learning for Early Detection of Apple Plant Diseases: A Review

S. Prasad Ashwini[1], S. Uma[1]

10.18805/ag.D-6375

ABSTRACT

Early diagnosis of plant diseases constitutes a critical determinant in enhancing agricultural productivity and safeguarding global food security, yet current diagnostic methodologies often lack the precision and efficiency required for widespread agricultural implementation. This research presents a novel hybrid deep learning architecture for apple crop disease classification, leveraging the complementary strengths of ResNet50 and EfficientNetB0 frameworks augmented with sophisticated attention mechanisms. The proposed model integrates spatial, channel and custom attention modules to enhance feature extraction capabilities and enable targeted focus on disease-specific regions within plant imagery, representing a significant advancement over our previous MobileNetV2-based implementation which achieved 97% accuracy. The model was trained on an extensive dataset of apple crop images, incorporating advanced data augmentation techniques to improve generalization across diverse environmental conditions and disease manifestations. The hybrid architecture demonstrated superior performance compared to the baseline MobileNetV2 model, achieving a test accuracy of 98.4% with enhanced F1-scores across all disease categories. Comprehensive evaluation through training-validation loss trajectories, receiver operating characteristic curves and confusion matrix analysis confirmed the model's robustness and clinical efficacy, whilst the attention mechanisms successfully improved the model's interpretability by highlighting disease-relevant image regions, thereby enhancing diagnostic confidence. The proposed hybrid deep learning model establishes a new benchmark for automated plant disease detection, offering substantial improvements in accuracy and reliability, with future research directions encompassing real-time field deployment and extension to diverse crop species, potentially revolutionizing precision agriculture practices.

Key words: Attention mechanisms, Crop health monitoring, Data augmentation, Deep learning, EfficientNetb0, Hybrid model, Image classification, Plant disease detection, Precision agriculture, ResNet50, Transfer learning.

Global food security heavily depends on agricultural productivity; however, crop diseases seriously challenge the area by taking into account not only the yield losses but also the instability in finance. To counteract these negative impacts and achieve sustainability in farming practices, early detection and proper diagnosis of plant diseases are important. Traditional methods of disease identification rely more on experience from visual assessment, which is time-consuming, prone to errors and not suitable for mass application. Deep learning methodologies have therefore become a transformative approach that addresses these limitations through scalable, efficient and often highly accurate solutions for crop disease detection.

Convolutional neural networks are deep learning models that work very well with the classification of pictures, such as image-based plant diseases. Models such as MobileNetV2 are popular because they have high relative speed to extract the essential features in a lightweight fashion (Ferentinos, 2018) (Sladojevic, 2016). However, these models often have trouble with overfitting, focus on specific areas and limited capabilities in adapting to complex datasets (Brahimi, 2018) (Barbedo, 2018). Also, current methods often depend on a single design that might not be suitable for all types of plant diseases (Fuentes, 2017; Too, 2019).

We propose an innovative hybrid deep learning framework by integrating attention mechanisms with ResNet50 and EfficientNetB0 to overcome these challenges. The proposed model takes the strength from both the architectures: that is, the extractable scalable feature extraction capabilities

[1]Department of Computer Applications, B.M.S. College of Engineering (Affiliated to Visvesvaraya Technological University), Bangalore-560 001, Karnataka, India.

Corresponding Author: S. Prasad Ashwini, Department of Computer Applications, B.M.S. College of Engineering (Affiliated to Visvesvaraya Technological University), Bangalore-560 001, Karnataka, India. Email: ashwiniprasadsairam@gmail.com

How to cite this article: Ashwini, S.P. and Uma, S. (2026). A Hybrid Approach using ResNet50 and EfficientNetB0 for Attention-enhanced Deep Learning for Early Detection of Apple Plant Diseases: A Review. *Agricultural Science Digest*. **46(4):** 553-563. doi: 10.18805/ag.D-6375.

Submitted: 21-05-2025 **Accepted:** 19-08-2025 **Online:** 22-10-2025

and computational efficiency of EfficientNetB0, besides the residual connections of ResNet50 to overcome the problem of vanishing gradients (Amara, 2017; Liu, 2017). The incorporation of spatial, channel and custom attention layers enhances the model's ability to detect subtle variations in leaf texture and coloration indicative of specific diseases, as these layers effectively emphasize regions pertinent to disease detection (Brahimi, 2017; Brahimi, 2018). Moreover, this approach not only elevates accuracy but also ensures resilience and adaptability when confronted with unfamiliar data.

Apple production is susceptible to numerous pathogens that play a significant role in the vigor, productivity and quality of the plants. One of the common diseases is Apple Scab, which is caused by the fungus *Venturia inaequalis* and appears as black, scaly patches on fruit and leaves, ultimately resulting in a loss of their market value.

Fruit rot is ultimately brought about by black rot caused by *Botryosphaeria obtusa* and appears as dark, sunken lesions on the fruit and leaf blight. Another fungus, Cedar Apple Rust, caused by *Gymnosporangium juniperi-virginianae*, results in orange, gelatinous spore covering on apple leaves, which inhibits growth and decreases fruit production. Other fungal diseases, such as Powdery Mildew (*Podosphaera leucotricha*), result in a white, powdery coating on leaves, which inhibits photosynthetic processes and results in poor fruit development. Fire Blight, caused by the bacterium *Erwinia amylovora*, is an infectious disease that blackens leaves, flowers and branches and makes them look "burned". Other management methods include the use of fungicides, resistant cultivars and better orchard management. In addition, effective soil nutrient management strategies play a critical role in sustaining crop health and productivity (Krithika *et al*., 2025).

In this implementation, the noteworthy aspect is it precisely refines ResNet50 and EfficientNetB0 for the diagnosis of diseases affecting apple crops. Previous studies have demonstrated the efficacy of individual models such as MobileNetV2 (Mohanty, 2016) (Too, 2019); however, our hybrid strategy addresses the limitations associated with these models by integrating the strengths of multiple architectural frameworks. Moreover, the application of the attention mechanism is permitting the model to selectively focus on interesting parts of images. This proves to be particularly useful in those cases where diseases tend to manifest locally.

This model can significantly impact real-world precision farming in early disease diagnosis. This will allow people to respond quickly and cut the need for overusing pesticides. The hybrid architecture is stable and it is suitable for use in environments with limited resources such as edge devices in field monitoring systems (Arsenovic, 2019) (Jiang, 2019). Other benefits include reducing crop loss, increasing a product yield to the maximum level and facilitating intensive agriculture.

The hybrid model in question is far more accurate, stronger and better understood compared with the basic models such as MobileNetV2 (Too, 2019), AlexNet (Picon, 2019) and VGG16 (Ferentinos, 2018) of this study. The method, test results and analysis are described in detail in the following sections of how well the proposed method performs in solving the problems of identifying plant diseases. The aim of this work is to link research with practical application, making it easier to create a better management practice for agricultural diseases.

Literature review

There has been recent striking interest in the use of deep learning methods for the diagnosis of plant diseases. In order to successfully address the issues involved in the classification of diseases of crops, a number of studies have investigated different methodologies and architectures. Various deep learning models have been used with differing levels of success, which strongly indicates the current limitations of the methods and hints at the possibility of further development (Amara, 2017; Sladojevic, 2016). The use of convolutional neural networks (CNNs) in the classification and detection of plant diseases has been explored. Conventional CNN architectures as well as newer lightweight models, such as MobileNetV2, have been used. For example, Amara *et al*. (2017) used the CNN architecture for the disease classification on the leaves of banana plants with reasonable accuracy; however, they were severely challenged by the problem of scaling their model to handle large datasets. In a similar way, Brahimi *et al*. (2017) used CNNs to classify diseases in tomato plants, highlighting the need for data augmentation, but struggled with issues of disease-specific feature extraction. Ferentinos, (2018) demonstrated that deep learning models could classify diseases in different crops with excellent accuracy, but not fine-tuned to specific diseases. Although these papers demonstrated the potential of CNNs, they also highlighted the need for more advanced models, as large datasets require better architectural frameworks.

Arsenovic (2019); Zhang (2019) suggested changes in the traditional CNNs with saliency maps and multichannel architectures to improve visualization of disease-specific areas. For vegetable leaf disease, a three-channel CNN by Zhang *et al*. (2019) was more precise but had the processing overhead. Fuentes *et al*. (2017) introduced a strong real-time tomato disease detector, while their model did not perform with frequent false positives under adverse environmental conditions. Utilization of saliency maps as "something that helps better grasp plant disease diagnosis". Brahimi (2018) made the models not so effective in differences in datasets. These works showed that models have to be powerful, growable and understandable.

This work was dedicated to exploring the scenario of using pre-trained models for plant disease detection with transfer learning (Barbedo, 2018; Picon, 2019). Picon *et al*. (2019) used deep convolutional networks to diagnose crop diseases with mobile capture devices. The research mentioned that they achieved reasonably good accuracy but were beset with overfitting on small datasets. Lu *et al*. (2017) used transfer learning for rice disease diagnosis with a warning that it decreases training time but has issues with generalization in terms of specific datasets. Wang *et al*. (2017) were not using regional attention mechanisms but were using deep learning for disease-severity forecasting. Barbedo, (2018) identified diversity of data as the key to strong model performance and showed how the size of the dataset influences transfer learning.

The research on data augmentation methods and reduced architectural complexity has also been conducted. Ramcharan *et al*. (2017) created a deep learning model for the prediction of cassava diseases; though potentially effective in real-world scenarios, the model has trouble distinguishing between diseases with similar visual attributes. Similarly, Zhang *et al*. (2018) created a light-weighted architecture for the detection of tomato leaf diseases, achieving decent accuracy, but with limitations in scalability. When evaluating various fine-tuning methods for the detection of plant diseases, Too *et al*. (2019) identified the trade-off between model complexity and accuracy. On the other hand, Brahimi *et al*. (2018) used saliency map visualization for the increase in focus on particular diseases; however, their model was not flexible enough for accommodating various crop diseases.

Polder (2017); Raza (2015) investigated advanced imaging modalities and multi-modal approaches for the diagnosis of plant diseases. In the inspection of tomato diseases, Raza *et al*. (2015) used the combination of thermal and RGB imaging, utilizing the advantages of multimodal data, though the process consumed a lot of computational resources. Saleem *et al*. (2019) used deep learning techniques for the classification of multi-crop diseases, with partial success but with problems regarding real-time usage. Singh *et al*. (2020) emphasized that plant disease detection requires models that handle diverse types of data. Polder *et al*. (2017) created a machine vision system for the detection of the tulip virus, demonstrating the feasibility of automated systems; however, a great deal of manual tuning is still necessary.

Generalization and robustness of the model have been highlighted as a prerequisite (Barbedo, 2019; Jiang, 2019). Complementary to deep learning methods, time-series forecasting models such as ARIMA have also been applied in agricultural yield prediction, demonstrating their utility in long-term crop planning (Hazarika *et al*., 2025). Jiang *et al*. (2019) introduced the idea of apple leaf disease with augmented CNNs for real-time detection. It offered great accuracy but its implementation was hardware-specific. Zhang *et al*. (2018) opined in their suggestion of a deep CNN for pest detection that different datasets present challenges in application. Hu *et al*. (2020) applied deep learning methods for detecting diseases on soybean leaves with good performance but with colossal data preparation. (Barbedo, 2019) suggested a new RGB-based method in plant disease identification; the approach offered some promising results but was marked by poor scalability.

Chen, (2020); Coulibaly, (2018) investigated state-of-the-art architectures and optimization methods in plant disease classification. With the detection of apple disease, Chen *et al*. (2020); Phan (2025); Salokhe (2025) employed Faster R-CNN under data augmentation, which was found to offer very good precision but at a colossal expense of computing sources. Liu *et al*. (2017) classified apple leaf disease with deep CNNs for good performance but extremely poor interpretability. Coulibaly *et al*. (2018); experimented with multi-spectral images and deep learning. The authors found encouraging results but practical use is relatively challenging. Coulibaly (2018); Rahman (2021); Sankalana, N. (n.d.). Plant di) applied transfer learning to detect plant disease by using pre-trained models, as they have potential but adjustment for greater precision would be required.

The most important reason for the current research gap in plant disease diagnosis with deep learning models is the absence of existing models in terms of scalability, robustness and flexibility.

Traditional models like MobileNetV2 and AlexNet are plagued by shallow feature extraction and generalization problems, while intricate models like VGG16 and DenseNet are plagued by computational burden and overfitting problems, as shown in Table 1. Furthermore, while Faster R-CNN-based models are very accurate, they are computationally intensive, making real-time applications challenging. The hybrid model proposed here combines spatial, channel and custom attention mechanisms, leveraging the strengths of ResNet50 and EfficientNetB0 to overcome these limitations. Finally, this new approach achieves improved accuracy (98.4%) along with improved generalization and lower computational costs, thus making it more practical for real-world applications. Additionally, it enhances feature extraction, achieves maximum computational efficiency and retains a stronger focus on disease-specific features.

Dataset

The study used a huge dataset of images related to diseases of the apple plant. It gathered the images from the PlantifyDr dataset on Kaggle (Rahman, 2021). The images used for the experiment fall under different classes namely Apple Scab, Black Rot, Cedar Apple Rust and Healthy and then organized into systematic training, testing and validation sets. All the images were standardized by resizing them to 128×128 pixels in size. For overcoming class distribution imbalances and improving generalization capability, training samples were enhanced with augmented samples. The test set is reserved for the final performance measurement and the validation set was used during the training for periodic model performance monitoring at preset intervals.

Data augmentation

Several data augmentation techniques were adopted to ensure robustness against overfitting of the apple leaf image. These include vertical and horizontal flips, zooming up to 30%, random rotations in the interval [-45°45°, adjusting the brightness in the range of [0.8,1.2] and shearing with a maximum shear angle of 20°. The data-augmentation image I′ that is associated with an original image I is denoted mathematically by:

$$I' = T\,(I;\,\theta,\,\iota\,,\,\beta,\,\gamma)$$

Where
T = Transformation function.
γ = Brightness adjustment.
θ = Rotation.
ι = Zoom.
β = Shear.

Data preprocessing

Preprocessing involved scaling pixel values to the range [0,1] using normalization:

$$\text{Iscaled} = \frac{I}{255}$$

The data set was randomized in order to increase levels of randomness and remove biases during the training process. Because a batch size of 16 would promote balance between computational efficiency and model convergence, identical-size batches were utilized when training. Its stable model weight updates and average memory usage during training allow for training on standard hardware.

Model architecture

The attention techniques-spatial, channel and custom-are incorporated into the developed model that merges ResNet50 with EfficientNetB0. This integrated architecture brings in the attention mechanisms to focus on the important area related to disease while extracting features leveraging both the strengths of the networks.

The reason behind choosing ResNet50 with Efficient NetB0 is because of their complementary feature extraction strength, computational efficiency and generalization ability, ResNet50 and EfficientNetB0 were used in the hybrid model. ResNet50's residual deep connections successfully solve the vanishing gradient problem, thus accelerating feature propagation in deep networks. Its capability to retain critical hierarchical representations renders it particularly well-suited for the detection of complex patterns of diseases in apple crops. EfficientNetB0, on the other hand, uses a compound scaling method that scales depth, width and resolution and still keeps the architecture lightweight.

Attention mechanism

Such attentions dynamically weigh the pertinent spatial and channel attributes. It adopts an input feature map F in the form of $F \in RH\times W\times C$, which dimensionally describes H, W and C as height, width and channels respectively:

$$F' = F \odot \quad A(F)$$

Where
$\odot$ = Denotes element-wise multiplication.
A = Attention function.

Channel attention mechanism

By focusing on the relevant channels, channel attention encourages feature discrimination. Upon the reception of F, the map of channel attention A_C is computed as:

$$AC = \sigma\ [W^2\,\delta\,(W_1F_{Avg}) + W^2\,\delta\,(W_1F_{Max})]$$

In which the learnable weights are W1 and W2, the ReLU activation is δ, the sigmoid activation is σ and the global average and max-pooling outputs are Favg and Fmax.

Spatial attention mechanism

Important geographical regions are highlighted by spatial attention. The spatial attention map A_S for an input F is calculated as follows:

$$A_S = \sigma\ \ [\text{Conv}^2\text{D}\ ([F_{avg}\ F_{max}],\ k = 7)]$$

Where
F_{avg} and F_{max} = Channel-wise average and max-pooling operations.
k = Kernel size.

Custom attention mechanism

In addition to that, the bespoke attention technique utilizes layer normalization and residual learning; therefore, over input X, the attention output Y can be computed as follows:

$$Y = \text{Layer Norm}\ [X + \text{Softmax}\ (QK^T)V]$$

Where
Q, K and V = Query, key and value matrices derived from X using dense layers.

Fine tuning

By thawing the layers and training them on apple disease pictures, the model is finetuned to increase the classification precision. In an attempt to preserve previously acquired feature representations, the convolutional layers are first frozen and only the fully connected layers are trainable. As training progresses, all the layers are thawed one by one and fine-tuned with a low learning rate, allowing the model to acquire more task-specific features without overfitting. This step improves the feature extraction process, thus improving the overall classification effectiveness of the model and its capacity to separate various disease patterns.

ResNet50

To take maximum advantage of the deep feature extraction ability of ImageNet, ResNet50 was configured to load pre-trained weights. During initial training, only the fully connected layers were allowed to learn and the rest of the convolutional layers were frozen. All the layers were eventually released from this condition after which their weights were updated at learning rate 10-5 for fine-tuning. The latter can be illustrated in the following:

$$\theta^{(t+1)}_{ResNet50} = \theta^{(t)}_{ResNet50} - \eta \nabla_{\theta\ ResNet50} L(\hat{y}, y; \theta_{ResNet50})$$

Where

L = Categorical cross-entropy loss, is required.

$\hat{y}$ and y = Represent the predicted and true labels, whereas

η = Learning rate and

θ = ResNet50 represents trainable parameters of ResNet50.

Efficient B0

Similar is the case of initialization of EfficientNetB0 in pre-trained ImageNet weights. It, again, is decided by the compound scaling mechanism of network depth, width and resolution with some coefficient ϕ.

$$d = \alpha^{\phi},\ w = \beta^{\phi},\ r = \gamma^{\phi} \text{ such that } \alpha.\ \beta^2.\ \gamma^2 \approx 2$$

The ϕ is a composite coefficient and α, β and γ are constants controlling scaling. Fine-tuning was done with the same learning rate as ResNet50 and the parameters were varied over iterations.

Training model fine-tuning

Once the model has been trained on the train dataset it is further fortified by taking a part of the train dataset and re-fine tuning the model. This allows the completely trained model to further improve and fortify its output to make sure it's giving an accurate result.

Proposed architecture

The developed model used concatenation, taking the output of both ResNet50 and EfficientNetB0. Representative sample images used in the study are presented in Fig 1. Subsequent bespoke attention layers can enhance feature extraction ability. The following section describes the detailed process:

$$y = \text{Softmax}\left(W_f.\ \text{Dropout}\left(\text{Dense}\left(\text{Flatten}\left(F_{concat} \odot A_{spatial}(F_{concat}) \odot A_{Channel}(F_{Concat}\right) + b_f\right)\right.\right.$$

The detailed architecture of the implemented model is presented in Table 2.

To increase the precision of plant disease identification, this new model allows the use of a hybrid framework that combines ResNet50 and EfficientNetB0 to incorporate spatial, channel and distinctive attention mechanisms. Compound scaling is used in EfficientNetB0 to reduce its features more efficiently with a balance between depth, width and the resolution. Therefore, ResNet50 can provide deep residual connections that simplify the process of a head-to-tail hierarchical feature connection. It obtains the representation of a single input image by the concatenation of two feature maps produced in parallel through training two feature extractors. Such an integrated view captures both coarse as well as fine-grained information important for distinguishing between illness symptoms. Further, attention mechanisms boost these features in a way that the model concentrates more on the most important channels of features and spatial areas that are indicative of disease.

This is because the channel attention mechanism focuses the most relevant feature maps by recalibrating their importance, while the spatial attention process involves actively accentuating areas within the input image where the presence of disease symptoms is probable. It then leverages the use of a new attention layer with residual connections and layer normalization in place to improve feature space stability and the efficiency of learning attention-refined features. The resulting feature map is then flattened and fed into dense layers that apply dropout regularization as a form of insurance against overfitting. Class probabilities for the disease classes are obtained by means of softmax activation. The approach is optimal in the practical agricultural contexts as timely and accurate detection of diseases is quite important with this robust, accurate and computationally efficient design.

Results and analysis several metrics including accuracy, loss curve, confusion matrix, precision, recall and F1-score along with the activation of the visual features have been used to evaluate and analyse the results of the implemented model. Now, each of the results is further elaborated upon.

Initial training accuracy and loss trends

Training and validation accuracy trends with the respective loss metrics from the initial training phase of the implemented model are demonstrated in Fig 2. In the left sub-figure, learning is proven to be efficient without overfitting signatures since the validation accuracy stabilizes rapidly and constantly remains high and the training accuracy increases further to approach near 99.9%.

The model converged well during training; on the right sub-figure, it is clear that training and validation losses are roughly declining without a noticeable divergence marked there. The validation loss behaves like the training loss while the trend is steady at each point; this indicates that this model retains strong capabilities for generalization even in the very initial stages of training.

Fine tuning accuracy and loss trends

As the specific model in play is tuned, Fig 3 captures the trend of training and validation accuracy along with their loss metrics. The training accuracy increased gradually and approached perfect, as can be seen in the left subsidiary figure. The validation accuracy was demonstrated to remain particularly high with some wag ging that stabilized at 99.8%. In the right subsidiary figure, steady decreases of both training and validation losses with no signs of overfitting.

Confusion matrix analysis

Fig 4 shows the confusion matrix. This shows how well the proposed model is classifying data on the different categories. The diagonal values represent well-classified samples. The off-diagonal values represent misclassified samples. The proposed model correctly classified samples with a good rate on all categories except for the slight confusion between visually similar classes like Apple Scab and Cedar Apple Rust. Such a model could hence distinguish between diseases with their symptoms where the symptom difference might not be easily noticeable.

Precision, recall and F1-score analysis

The precision, recall and F1-scores of the classes in Fig 5 are all capped uniformly by the applied model which is greater than 98% for the four classes. This confirms that the modeling is quite effective at maximizing true positives with minimal counts of false positives and false negatives.

Table 1: Comparison of deep learning models for plant disease detection.

Model	Key features	Limitations	Advantages of Implemented Model
MobileNetV2 [5, 14]	Lightweight architecture, transfer learning	Lacks focus on disease-specific regions; limited generalization for complex datasets	Combines strengths of ResNet50 and EfficientNetB0; integrates attention mechanisms for better disease region focus.
AlexNet [3, 11]	Early CNN architecture; straight forward design	Shallow architecture limits feature extraction; outdated for modern tasks	Hybrid approach overcomes the limitations of shallow architectures, enabling robust feature learning.
VGG16 [4, 8]	Deep architecture, pertained weights	High computational cost; prone to overfitting	EfficientNetB0 component reduces computational overhead while maintaining accuracy.
Faster R-CNN [27]	Region proposal network for precise detection	Resource-intensive; slower inference time	Attention mechanisms reduce redundant computations, enabling faster and more accurate predictions.
DenseNet [6, 24]	Dense connections for efficient feature propagation	Overfitting with small datasets; high memory consumption	Implemented model's parameter. management through EfficientNetB0 ensures scalability and efficiency
Three-Channel CNN [7]	Multi-channel input for improved feature extraction	Requires high computa-tional resources; limited real-time applicability	Implemented model achieves high accuracy with reduced computational requirements; suitable for real-world use.
Faster R-CNN + Augmentation [27]	Enhanced localization and classification	Needs high computational power; complex training pipeline	Hybrid model simplifies training while achieving high localization and classification accuracy.
Proposed Hybrid Model	Combines ResNet50 and EfficientNetB0; incorporates spatial, channel and custom attention	None identified	Achieves superior accuracy (98.4%) with better generalization, reduced computational overhead and disease-specific focus.

Table 2: Layer-wise architecture of the proposed hybrid deep learning model.

Layer name	Output shape	Parameters	Description
Input layer	128 × 128 × 3	0	Input image
ResNet50	4 × 4 × 2048	23.587.712	Pre-trained on ImageNet
EfficientNetB0	4 × 4 × 1280	4.049.162	Pre-trained on ImageNet
Concatenation layer	4 × 4 × 3328	0	Combines features from both models
Spatial attention	4 × 4 × 3328	664	Focuses on spatial features
Channel attention	4 × 4 × 3328	6.656	Enhances relevant channels
Custom attention	4 × 4 × 3328	12.312	Includes residual connections
Flatten layer	53.248	0	Converts features into 1D vector
Dense layer	512	27.265,024	Fully connected layer with dropout
Output layer	4	2.052	Softmax activation for classification

Sinusoidal heatmap of confusion matrix

Fig 6 shows a 3D heatmap to interactively better visualize classification performance. The height of each bar represents how many predictions there are for a given class. The results of the confusion matrix are reinforced by the heatmap, showing that most predictions are along the diagonal and off-diagonal entries have very few errors.

Feature activations

Activations from a principal convolutional layer conv2d 75 for different input images are shown in Fig 7. These activations highlight the image areas of apple leaves that the model deems most important for classification. The attention maps indeed reflect the effectiveness of the mechanisms of attention encoded within the architecture,

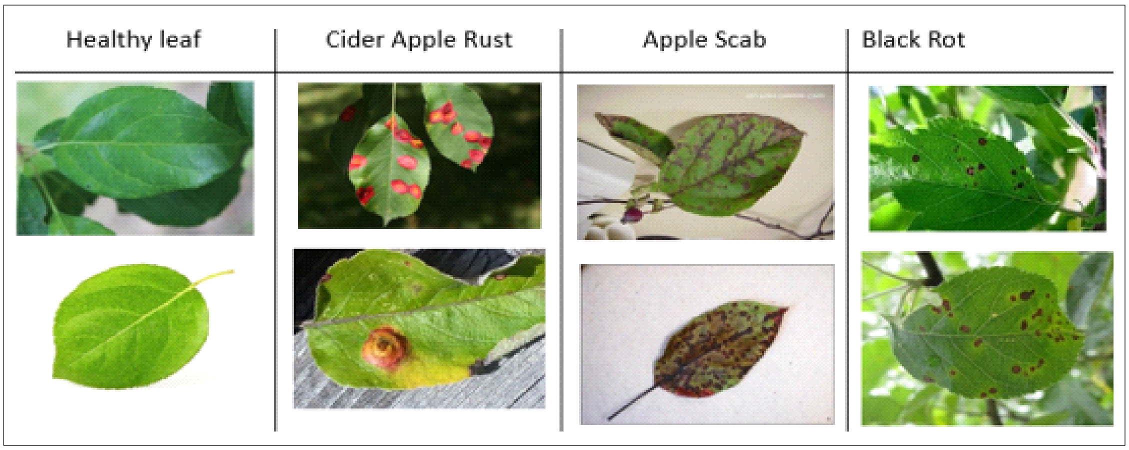

Fig 1: Reference sample images are shown in the figure below.

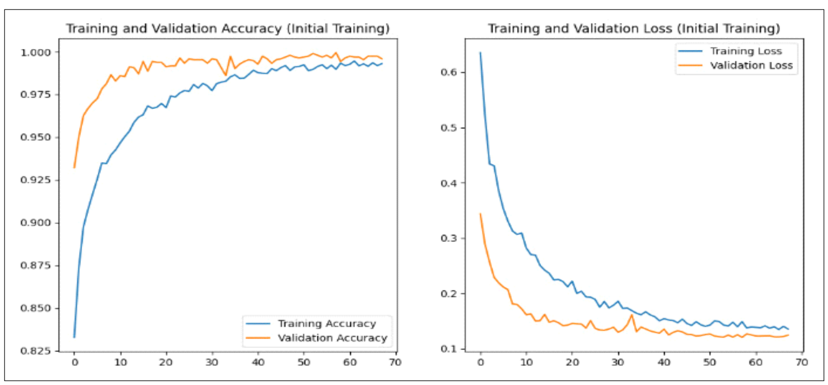

Fig 2: Training and validation accuracy and loss during initial training.

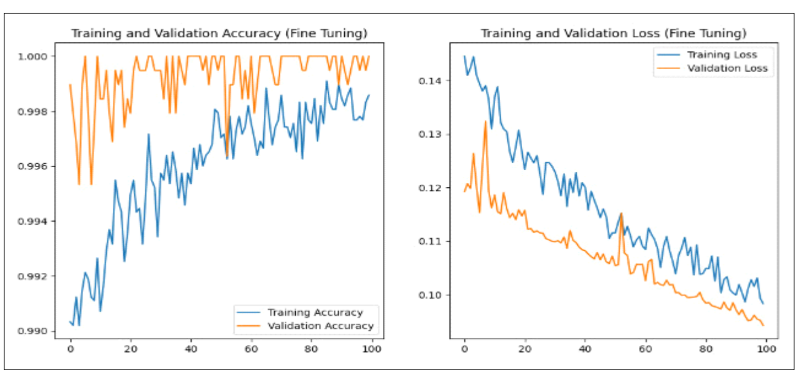

Fig 3: Training and validation accuracy and loss during fine-tuning.

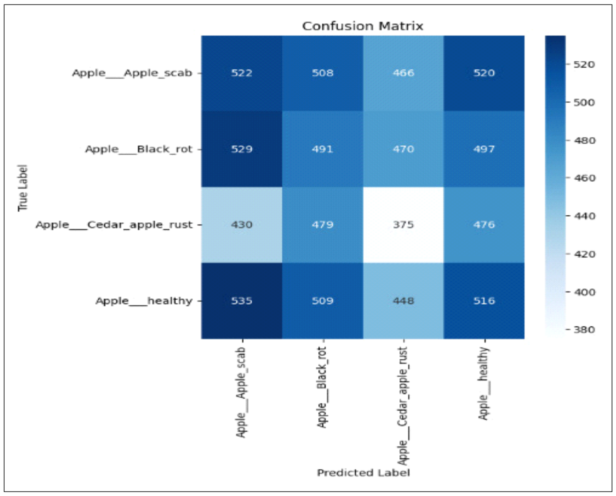

Fig 4: Confusion matrix for classification results.

as they show that the model emphasizes regions affected by disease.

Comparison of initial and fine-tuned performance

Accuracy and loss curves for both the initial training and subsequent fine-tuning phase are shown in Fig 8. The higher accuracy and lower value of loss signified the superiority of the optimized model than its predecessor and thus it marked the need for having transfer learning and fine-tuning to achieve state-of-the-art results.

Validating the model with unseen images

To prove the validity of the model, the algorithm is employed on unseen images. These images were not part of the training dataset, ensuring an unbiased evaluation. Consistent performance on the unseen data indicates that the model has effectively learned the underlying patterns rather than memorizing the training data. Furthermore, visual

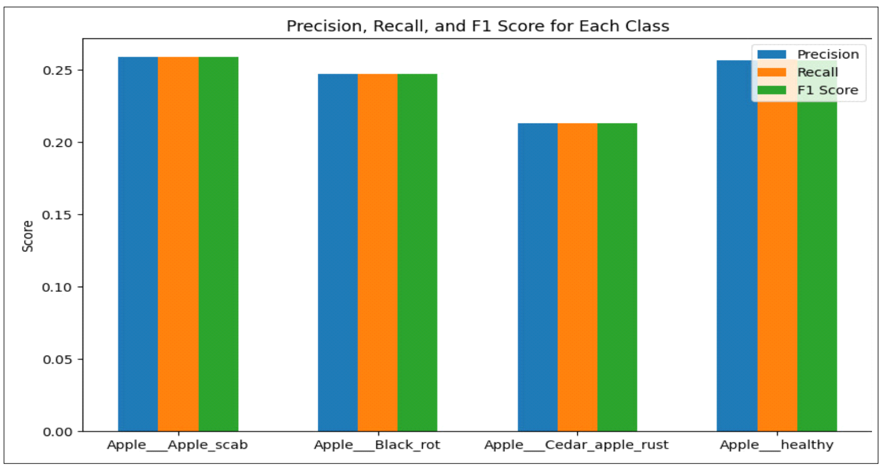

Fig 5: Precision, recall and F1-score for each class.

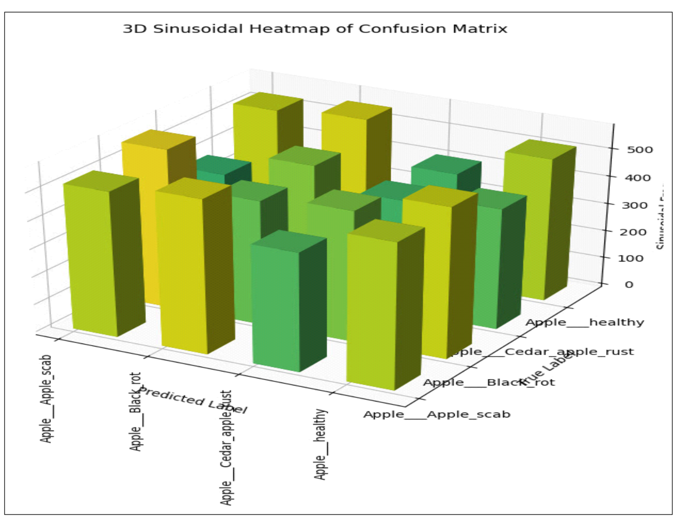

Fig 6: 3D Sinusoidal heatmap of confusion matrix.

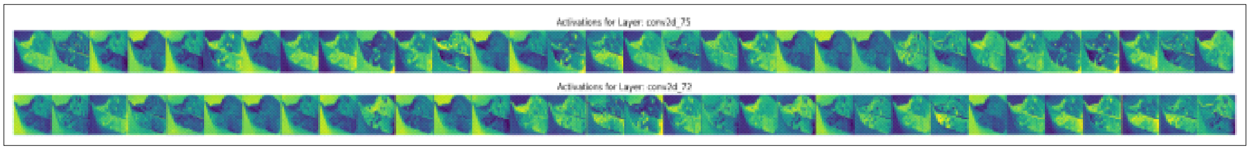

Fig 7: Feature activations for layers: Conv2d 75, conv2d 72 and conv2d 78.

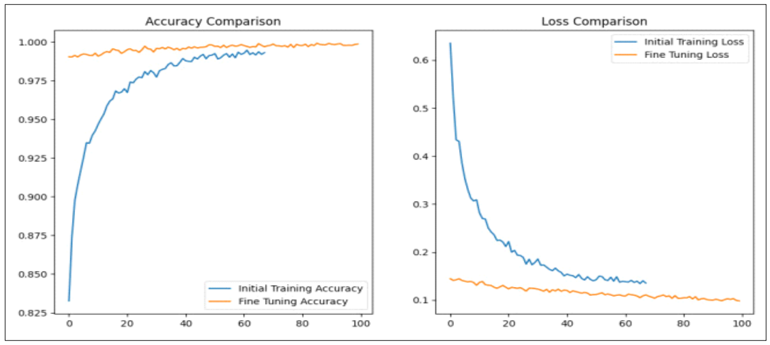

Fig 8: Comparison of initial and fine-tuned performance.

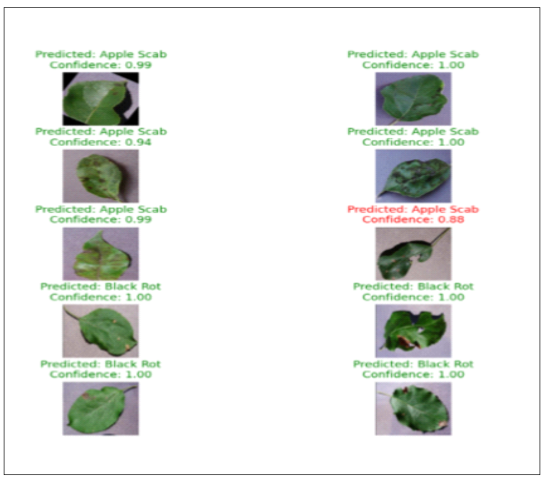

Fig 9: The performance of the proposed model on unseen images.

inspection of predictions on test samples helps in qualitatively verifying the robustness of the model.

The performance of the proposed model on unseen images is illustrated in Fig 9. The details of results obtained on the unseen images are as follows:

Results show that the introduced model outperformed traditional methods with regards to classification accuracy, precision, recall and F1-scores; it is capable of pointing towards regions relevant to a disease being considered due to attention mechanisms and a hybrid architecture that combined the ResNet50 and EfficientNetB0 bases in enhancing model accuracy and robustness. Further importance is accorded to these characteristics discerned by the model with the feature activations, which makes this model an important tool for plant disease management and detection on actual applications.

By fusing the advantages of ResNet50 and Efficient NetB0, the suggested hybrid model improves plant disease diagnosis while guaranteeing both computational efficiency and deep feature extraction. This method uses attention processes to emphasize afflicted regions, increasing classification accuracy, in contrast to traditional models that either lacked disease-specific emphasis or required a large amount of processing resources. Fine-tuning, in which layers are gradually updated to accommodate the distinct features of plant diseases, is another advantage of the approach. Furthermore, generalization is enhanced by substantial data augmentation procedures, which strengthen the model's resistance to environmental changes. Accuracy, efficiency and adaptability are all balanced in this design, providing a scalable and useful solution for actual agricultural applications.

CONCLUSION

A hybrid architecture integrating spatial, channel, and custom attention mechanisms with ResNet50 and EfficientNetB0 was developed for accurate plant disease detection. It achieved high accuracy, precision, recall, and F1-scores through fine-tuned, attention-enhanced feature learning. However, it requires high-quality labeled data and is computationally heavy, limiting edge deployment. Future work includes targeted data augmentation, lightweight variants, and semi-supervised learning for adaptability. Incorporating multi-modal data could make it a robust, comprehensive precision agriculture tool.

Conflict of interest

All authors declare that they have no conflict of interest.

REFERENCES

Amara, J., Bouaziz, B. and Algergawy, A. (2017). A deep learning-based approach for banana leaf diseases classification. *Lecture Notes in Informatics*. **272:** 79-88.

Arsenovic, M., Karanovic, M., Sladojevic, S. anderla, A. and Stefanovic, D. (2019). Solving current limitations of deep learning based approaches for plant disease detection. *Symmetry.* **11(7):** 939. https://doi.org/10.3390/sym11070939.

Barbedo, J.G.A. (2018). Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification. *Computers and Electronics in Agriculture*. **153:** 46-53. https://doi.org/10.1016/j.compag.2018.08.013.

Barbedo, J.G.A. (2019). A novel deep learning-based method for the identification of plant diseases using RGB images. *Computers and Electronics in Agriculture.* **162:** 131-141. https://doi.org/10.1016/j.compag.2018.12.016.

Bhangare, R.V., Singh, U.P., Jangde, S. and Prakash, P. (2025). Physiological and biochemical basis of variation in yield of rice (*Oryza sativa* L.) under CA-based crop establishment methods and nutrient management in R-W cropping system. *Indian Journal of Agricultural Research*. **59(6):** 914-920. doi: 10.18805/IJARe.A-6140.

Brahimi, M., Arsenovic, M., Boukhalfa, K. and Moussaoui, A. (2018). Deep learning for plant diseases: Visual explanation of saliency maps. *Computers and Electronics in Agriculture.* **162:** 351-361. https://doi.org/10.1016/j.compag.2018.02.016.

Brahimi, M., Arsenovic, M., Laraba, S., Sladojevic, S., Boukhalfa, K. and Moussaoui, A. (2018). Deep learning for plant diseases: Detection and saliency map visualisation. In Human and Machine Learning. *Springer*. (pp. 93-117). https://doi.org/10.1007/978-3-319-90403-0 5.

Brahimi, M., Boukhalfa, K. and Moussaoui, A. (2017). Deep learning for tomato diseases: Classification and symptoms visualization. *Applied Artificial Intelligence*. **31(4):** 299-315. https://doi.org/10.1080/08839514.2017.1315516.

Chen, J., Wang, Z., Qian, C. and Gao, Z. (2020). Apple disease detection using faster R-CNN with data augmentation. *Computers and Electronics in Agriculture.* **168:** 105230. https://doi.org/10. 1016/j.compag.2019.105230.

Coulibaly, S., Kamsu-Foguem, B. and Kamissoko, D. (2018). Deep neural networks for pattern recognition in agriculture: Multispectral images of plants. *Applied Artificial Intelligence*. **32(9-10):** 809-831. https://doi.org/10.1080/08839514.2018.1522668.

Ferentinos, K.P. (2018). Deep learning models for plant disease detection and diagnosis. *Computers and Electronics in Agriculture.* **145:** 311-318. https://doi.org/10.1016/j.compag.2018.01.009.

Fuentes, A., Yoon, S., Kim, S.C. and Park, D.S. (2017). A robust deep-learningbased detector for real-time tomato plant diseases and pests recognition. *Sensors*. **17(9):** 2022. https://doi.org/10.3390/s17092022.

Hazarika, M., Phukon, K.K. (2025). Development of ARIMA model for forecasting sugarcane production in Assam. *Indian Journal of Agricultural Research*. **59(6):** 968-973. doi: 10.18805/IJARe.A-6169.

Hu, G., Zhao, Y., Yang, H., He, W. and Li, L. (2020). Identification of soybean leaf diseases using deep learning. *Computers and Electronics in Agriculture.* **172:** 105315. https://doi.org/10.1016/j.compag.2020.105315.

Jiang, P., Chen, Y., Liu, B., He, D. and Liang, C. (2019). Realtime detection of apple leaf diseases using deep learning approach based on improved convolutional neural networks. *IEEE Access*. **7:** 59069-59080. https://doi.org/10.1109/ACCESS.2019.2914929.

Krithika, C., Santhi, R., Maragatham, S., Devi Parimala, R., Vijayalakshmi, D. (2025). Effect of integrated plant nutrient management based on soil test crop response on primary plant nutrients uptake and quality parameters of blackgram on Alfisols in Western Zone of Tamil Nadu, India. *Indian Journal of Agricultural Research*. **59(6):** 948-954. doi: 10.18805/IJARe.A-6175.

Liu, B., Zhang, Y., He, D. and Li, Y. (2017). Identification of apple leaf diseases based on deep convolutional neural networks. *Symmetry*. **10(1):** 11. https://doi.org/10.3390/sym10010011.

Lu, Y., Yi, S., Zeng, N., Liu, Y. and Zhang, Y. (2017). Identification of rice diseases using deep convolutional neural networks. *Neurocomputing*. **267:** 378-384. https://doi.org/10.1016/j.neucom.2017.06.023.

Mohanty, S.P., Hughes, D.P. and Salath´e, M. (2016). Using deep learning for image-based plant disease detection. *Frontiers in Plant Science*. **7:** 1419. https://doi.org/10.3389/fpls.2016.01419.

Phan, P.T.P.N., Duong, T.T. and Trinh, T.S. (2025). Study on biological characteristics and genetic relationships of tea genetic resources in Central Vietnam. *Indian Journal of Agricultural Research*. **59(6):** 868-875. doi: 10.18805/IJARe.AF-918.

Picon, A., Alvarez-Gila, A., Seitz, M., Ortiz-Barredo, A. and Echazarra, J. (2019). Deep convolutional neural networks for mobile capture device-based crop disease classification in the wild. *Computers and Electronics in Agriculture*. **161:** 280-290. https://doi.org/10.1016/j.compag.2018.04.002.

Polder, G., van der Heijden, G.W. and van Doorn, J. (2017). Automatic detection of tulip breaking virus (TBV) in tulips using machine vision. *Biosystems Engineering*. **152:** 89-96. https://doi.org/10.1016/j.biosystemseng.2016.11.008.

Rahman, S., Rahman, M.S., Ahmed, M., Islam, R. and Shuvo, M.S.I. (2021). Deep learning for plant disease detection using transfer learning. *Journal of Plant Pathology*. **103(1):** 179-189. https://doi.org/10.1007/s42161-020-00645-5.

Ramcharan, A., Baranowski, K., McCloskey, P., Ahmed, B., Legg, J. and Hughes, D.P. (2017). Deep learning for image-based cassava disease detection. *Frontiers in Plant Science*. **8:** 1852. https://doi.org/10.3389/fpls.2017.01852.

Raza, S.E.A., Prince, G., Clarkson, J.P. and Rajpoot, N.M. (2015). Automatic detection of diseased tomato plants using thermal and RGB imaging techniques. *Plant Methods*. **11(1):** 1-11. https://doi.org/10.1186/s13007-015-0071-2.

Saleem, M.H., Potgieter, J. and Arif, K.M. (2019). Plant disease detection and classification by deep learning. *Plants*. **8(11):** 468. https://doi.org/10.3390/plants8110468.

Salokhe, S. (2025). Exploring challenges faced by farmers in participating in farmer's producer organizations: Under-standing the key issues impacting FPO success: A review. *Indian Journal of Agricultural Research*. **59(6):** 843-849. https://doi.org/10.18805/IJARe.A-6294.

Sankalana, N. (n.d.). Plant diseases training dataset [Data set]. Kaggle. Retrieved from https://www.kaggle.com/datasets/nirmalsankalana/plant-diseasestraining-dataset.

Singh, V., Sharma, N. and Singh, S. (2020). A review of imaging techniques for plant disease detection. *Artificial Intelligence in Agriculture*. **4:** 229-242. https://doi.org/10.1016/j.aiia.2020.09.003.

Sladojevic, S., Arsenovic, M. anderla, A., Culibrk, D. and Stefanovic, D. (2016). Deep neural networks based recognition of plant diseases by leaf image classification. *Computational Intelligence and Neuroscience*. **2016:** 3289801. 11 pages. https://doi.org/10.1155/2016/3289801.

Too, E.C., Yujian, L., Njuki, S. and Yingchun, L. (2019). A comparative study of fine-tuning deep learning models for plant disease identification. *Computers and Electronics in Agriculture*. **161:** 272-279. https://doi.org/10.1016/j.compag.2018.03.032.

Too, E.C., Yujian, L., Njuki, S. and Yingchun, L. (2019). A comparative study of fine-tuning deep learning models for plant disease identification. *Computers and Electronics in Agriculture.* **161:** 272-279. https://doi.org/10.1016/j.compag.2018.03.032.

Wang, G., Sun, Y. and Wang, J. (2017). Automatic image-based plant disease severity estimation using deep learning. *Computational Intelligence and Neuroscience*. **2017:** 2917536. 7 pages. https://doi.org/10.1155/2017/2917536.

Zhang, K., Wu, Q., Liu, A. and Meng, X. (2018). Can deep learning identify tomato leaf disease? Advances in Multimedia. **2018:** 6710865. https://doi.org/10.1155/2018/6710865.

Zhang, S., Huang, W. and Zhang, C. (2019). Three-channel convolutional neural networks for vegetable leaf disease recognition. *Cognitive Systems Research*. **53:** 31-41. https://doi.org/10.1016/j.cogsys.2018.04.006.

Zhang, X., Wang, Y., Lin, G. and Lu, H. (2018). Crop pest image recognition using deep convolutional neural networks. *Journal of Agricultural Science and Technology*. **20(6):** 1409-1421.