An AI-Powered Diagnose of Plant Disease Using CNN And Attention Mechanism

Parveen Kaur[1]
{dhaliwalz.noor@gmail.com[1]}

School of Computer Science & Engineering, Lovely Professional University, Phagwara, India[1]

Abstract.

The inclusion of Artificial Intelligence (AI) in precision agriculture is growing rapidly, particularly in the automated detection of plant diseases. Conventional CNN-based models have shown success. However, they often fail to focus attention on the most significant parts of images, where vital information about the disease is found. This work proposes an extended CNN architecture that is integrated with Convolutional Block Attention Module (CBAM) for enhancing channel-wise and spatial-wise feature focus in plant disease detection in tomato, potato, and bell pepper leaves. The proposed pipeline consists of preprocessing, segmentation, data augmentation, feature extraction, and classification, and is tested on the PlantVillage dataset with six classes of diseases/health. The CNN-CBAM model attains an overall accuracy of around 98.0% with increased precision, recall, and F1-score compared to a baseline CNN and a MobileNet-based approach, which makes it robust and capable of generalization. The key contributions are multi-crop attention-driven architecture, end-to-end pipeline, and comparative analysis of the proposed method with the state-of-the-art deep learning methods for plant disease diagnosis."

Keywords: Artificial Intelligence (AI), Precision Agriculture, Convolutional Neural Network (CNN), Convolutional Block Attention Module (CBAM), Early Diagnosis, Plant Diseases

1 Introduction

In contemporary agriculture, two key needs must be addressed: improving food security and achieving sustainability. This is possible by controlling plant diseases that can halve yields and threaten global food security. Plant diseases are vicious and require rapid, practical actions to detect and control them before situations worsen [1][2][3]. Agriculture has always been vital to the economy and food security, but it is susceptible to diseases that cause abrupt yield declines [4]. The serious impact of such illnesses on food and lives makes it necessary to develop policies for managing parasitic plants.

Previously, the diagnosis of disease in plants was accomplished through laboratory and large-scale manual surveys. These techniques, although precise, require a significant amount of human

IICTDS 2025, December 19-20, Chandigarh, India
Copyright © 2026 EAI
DOI 10.4108/eai.19-12-2025.2362579

time and long hours, making them infeasible for a large-scale farming business. Moreover, the small-scale physiology of plant pathogens may contribute to difficulties in location-based diagnosis, and the limited diagnostic skills available in farming communities may compromise the accuracy of diagnosis. These restrictions are also key to utilizing incredibly advanced tools that can automatically identify and address the disease. It has thus become apparent that artificial intelligence is a powerful tool for enhancing the detection of plant diseases [5]. The latest developments in deep learning (DL) and computer vision have significantly improved the industry's precision and speed; however, the accuracy and speed of the industry have improved substantially lately, in fact, nearly twofold, as stated in the article [6]. The current techniques rely on convolutional neural networks (CNNs) and are likely to be highly automated, minimizing the need for human intervention. Disease management has also become quicker and more reliable with the use of these technological advancements. Mohanty et al.[7]and Barbedo[8] are among the first to conduct a ground-breaking work in this field.

1.1 Motivation

Convolutional neural network (CNN) models have demonstrated exceptional performance in image classification over the past few years. Their strength lies in identifying hierarchical feature representations. However, these models cannot show relationships between pixels that are far apart in space because their receptive fields are limited to the width of a filter. As a result, CNNs require large data sets and high computational resources. This can be inefficient and often unfeasible in resource-limited settings. Modern studies have addressed this by analyzing architectures with attention-wise structures. Such structures improve attention to distant pixel connections. This motivates the development of the hybrid framework based on an attention-enhanced mechanism. It combines channel-and-spatial attention modules, CBAM, into traditional convolutional blocks, which will selectively identify salient components in a visual scene, thereby enhancing overall accuracy and interpretability while still being appropriate for real-world agricultural scenarios.

1.2 Main Contribution of this work as follows :

1. CNN-CBAM-based deep learning architecture is proposed for plant leaf disease detection, to which channel and spatial attention modules are applied to select the disease-related features and suppress the background noise.

2. An end-to-end pipeline is created consisting of preprocessing, segmentation, augmentation, feature extraction, and classification for automated diagnosis from images of leaf samples of tomato, potato, and bell pepper.

3. The proposed model is trained and evaluated on the PlantVillage dataset for six classes (healthy and multiple types of disease) presenting better accuracy, precision, recall and F1-score results when compared to a baseline CNN model.

4. A comparative analysis is given in contrast to a method based on MobileNet and other recent deep learning methods, indicating the performance and applicable value of the proposed CNN-CBAM framework for plant disease detection."

2 Literature Review

Deep learning has proven effective in disease detection, particularly when enhanced with attention modules such as CBAM. CBAM improves feature extraction by highlighting the most relevant image regions. In a 2022 study by Pang and Bo [9], integrating CBAM with Spatial Group-wise Enhancement (SGE) in an EfficientNet-based architecture improved channel and spatial attention, while SGE reduced noise and managed intra-spatial feature significance. This approach achieved 98.53% accuracy on the Flowers dataset, demonstrating its value for disease identification. Similarly, Al-Gaashani et al.[10] designed SANet, a variant of ResNet-50 with self-attention, which achieved an accuracy of 98.71% in rice disease identification. Additionally, their work on residual convolutional networks with attention showed improved feature extraction and accuracy for tomato leaf disease detection [11].

Transformers have been getting a lot of attention lately for their ability to model global dependencies. A hybrid CNN-ViT model (Vision Transformer), which was proposed in 2023 by Thakur et al. [12], produced 98.86% accuracy on PlantVillage and 89.24% accuracy on Embrapa while outperforming nine state-of-the-art methods on five datasets while maintaining a light-weight parameter count of 0.85 million. In 2025, an end-to-end transformer architecture with Collaborative Hybrid Assignment (Co-DETR) for fruit disease detection was trained further, leading to fruit disease detection with a mean average precision of 0.89 on six classes and its robustness under controlled adversarial conditions of low-lumination and image blurring [13].

Multimodal approaches have been developed to overcome the limitations of unimodal image-only systems. HVAF-XAI-Net framework proposed in 2025 [14], combined features of ViT-based spatial features from UAV imagery and temporal patterns from IoT-based environmental data using Temporal Convolutional Networks and cross attention fusion. The model achieved 97.0% accuracy (95% confidence), saved 22% of false positives, and outperformed ResNet-50 (88.2%), CNN-LSTM hybrid (89.5%) and Swin Transformer (90.1%). Likewise, a Multi-modal framework to detect Tomato diseases (2024) was EfficientNet-B0 coupled with recurrent neural networks to classify the diseases and estimate severity with 96.40% classification accuracy and 99.20% severity accuracy with SHAP and Lime explainability [15].

Modern leaf disease detection methods actively compare hybrid and pre-trained architectures to determine the most effective model for this purpose. Patil et al. [16] utilized ResNet-50 and ResNet-101 for feature extraction, then employed Support Vector Machines for classification, achieving an F1-score of 99.82%. Yadav et al.[17] compared VGG-16, VGG-19, and ResNet-50 for grape leaf analysis and reported that ResNet-50 attained an accuracy of 95%. Unal [18] investigated and compared the performance of five deep-learning models—VGG-16, VGG-19, DenseNet-121, Xception, and InceptionV3—and demonstrated that DenseNet-121 produced the highest accuracy. To further improve performance, Unal combined DenseNet-121 with CBAM and Squeeze-and-Excitation Networks, resulting in an increased classification accuracy to 96.36%, up from the original 92.73

Febriantono et al. [19] used EfficientNet-B0 as a baseline for apple leaf disease detection, with the addition of CBAM improving both feature extraction and accuracy, achieving an F1 score of 0.92 and an accuracy of 0.9367. Incorporating CBAM-based attention raised the robustness, speed, and accuracy. Rahman et al.[20] employed six architectures for cotton leaf disease identification,

with CBAM models focusing on essential features. DenseNet-169 achieved 96.26% accuracy, and EfficientNet-B1 with CBAM achieved 99.21% on the augmented dataset.

3 Deep Learning Algorithm for Plant Disease Classification

Classification is a supervised learning process where input is assigned to certain categories. Initially, labeled data is used to train the classifier and thus to build patterns and characteristics that represent different classes of diseases. The model that has been trained is then used to classify the unlabeled data. Standard measures of performance are incorporated to validate the outcome.

In the current research, two deep neural network models are taken into consideration, i.e., Convolutional Neural Networks (CNNs) and a variant model with Convolutional Block Attention Modules (CBAM).These models are made to automatically detect complex, discriminative features leading to high accuracy in identifying various plant diseases.

3.1 Convolutional Neural Network

A Convolutional Neural Network (CNN) is a powerful type of deep learning model; especially useful for classifying plants and diseases from images, and this area is the primary domain where visual evaluation is important. By operating convolution filters on input data, CNNs automatically extract descriptive features from the input that capture localized features, i.e. edges, textures, and color variations in the input data. This process enables the model to identify pathological symptoms, such as spots, lesions, and discoloration, without requiring manual feature engineering. When trained on annotated datasets, CNNs demonstrate strong generalization across diverse conditions, image types, lighting, plant postures, and disease appearances. During refinement, the model minimizes classification errors and maps specific disease types to their visual signatures through repeated training.

3.2 CNNs Integrated with Convolutional Block Attention Module (CBAM)

This architecture augments a typical CNN with CBAM as an attention mechanism, which works sequentially on the channels and spatial locations. Channel attention is used to aggregate global information to give higher weights to channels that react strongly to discriminative phenomena (e.g. lesion texture, edge patterns, characteristic discolorations) and suppress less informative channels. Spatial attention then generates a map over the spatial dimensions, emphasizing critical regions on the leaf surface (e.g. spots, blighted areas or mottling) and de-emphasizing homogeneous or background regions. By combining these two stages, the CNN-CBAM model concentrates its capacity in areas and patterns associated with diseases, increasing the accuracy of classification and allowing the representations to be interpreted."

4 Proposed Methodology

This work aims to develop an accurate and interpretable model for automating plant disease diagnosis from leaf images. The approach involves using CNNs for recognizing the image patterns

and CBAM attention mechanism for focusing on important regions. The architecture, as shown in **Figure 1**, is made of two main phases, namely training and testing. Each phase consists of image acquisition, image preprocessing, segmentation (leaf isolation from the image), feature extraction, classification and performance evaluation. The model learns features from segmented images when it is being trained. In testing, when a new picture of the leaf is present, it is passed through the same procedure and the features are analyzed for diagnosis.

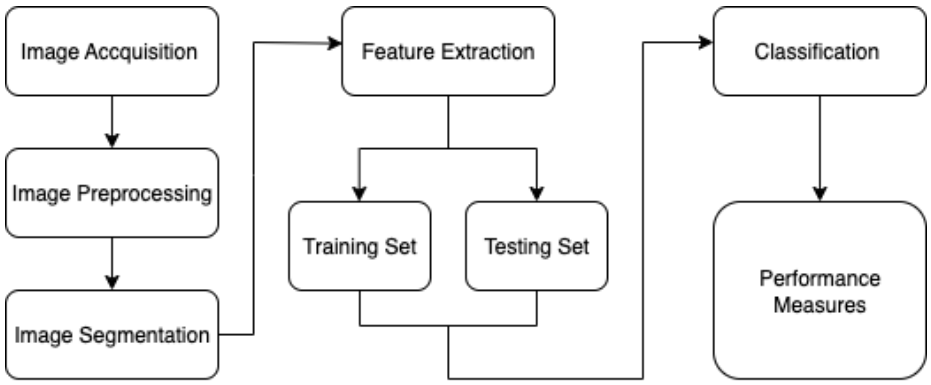


Fig. 1. Block Diagram of proposed methodology

4.1 Image Acquisition

This research is based on images from the PlantVillage database [21], which is a comprehensive database of labeled plant-leaf images from different plant species and diseases. The crops used for the study include tomatoes, potato, and bell peppers and the leaves of both healthy and diseased plants of these crops, as shown in **Figure 2**. The samples of tomato include bacterial spot, septoria leaf spot, early and late blight, leaf mold, yellow leaf curl virus, and healthy leaves. Potato samples include early blight, late blight, and healthy leaves. Bell pepper samples feature bacterial spots and healthy leaves. Details of the sample, such as the count and categories of crop leaves, are given in Table 1. To simplify the analysis, the entire dataset is separated into training (90%) and testing (10 %) sets.

4.2 Image Preprocessing

To ensure reliable inference, images were preprocessed before being input to the neural network. The dataset included dew droplets, dust particles, and insect excretions, which are considered noise in leaf images. The following preprocessing methods were applied:

1. **RGB to Grayscale:** Color images were converted to grayscale to emphasize texture and structure while reducing computational complexity.

Table 1: Details of the dataset used for classification.

S. no	Type of leaf disease	Count of images
1	Healthy	395
2	Blight (Early + Late)	590
3	Bacterial Spot	200
4	Yellow Leaf Curl Virus	355
5	Mosaic Virus	165
6	Septoria Leaf Spot	180

2. **Resizing:** The resizing of all images to 224x224 pixels is performed to ensure that a common input dimension will be present across labels to the whole network, a condition that is supported by the majority of the CNN structures that have this stated input dimension.

3. **Normalization:** During training, the pixel values are also re-scaled to lie in the range of 0 to 1, therefore center-weighting the optimization process by equalizing relative magnitudes and also compensating for vanishing and exploding gradients.

4. **Data Augmentation:** Augmentation strategies are applied to impose real-world variability where there are simulations of random rotations (±20 degrees), horizontal and vertical flips, zoom-in cropping, and shifts in brightness and contrast. This augmentation introduced the model to varying visual representations of each disease, thereby increasing robustness and preventing overfitting.

5. **Label Encoding:** Class labels can be a class name like '*Tomato_Late_Blight*', and we recoded the values into categorical values that can be used in the final output layer, which is the softmax activation.

4.3 Segmenting of images

Image segmentation is crucial and difficult for plant disease identification. It segments images into components that are meaningful in order to aid in further analysis. Segmentation is based on two major principles: similarity and discontinuity. Similarity groups pixels with similar properties, such as color or texture, while edge detection detects boundaries between regions. Discontinuity is focussed on sharp changes in intensity, on the boundaries between structures, for instance between leaves and lesions. Effective segmentation improves the feature extraction and classification accuracy so that the CNN can pay attention to the relevant areas.

4.4 Feature Extraction

1. **CNN Approach:** After preprocessing, the segmented image is fed to a CNN. The network automatically detects important features in multiple convolutional layers. This approach cap-

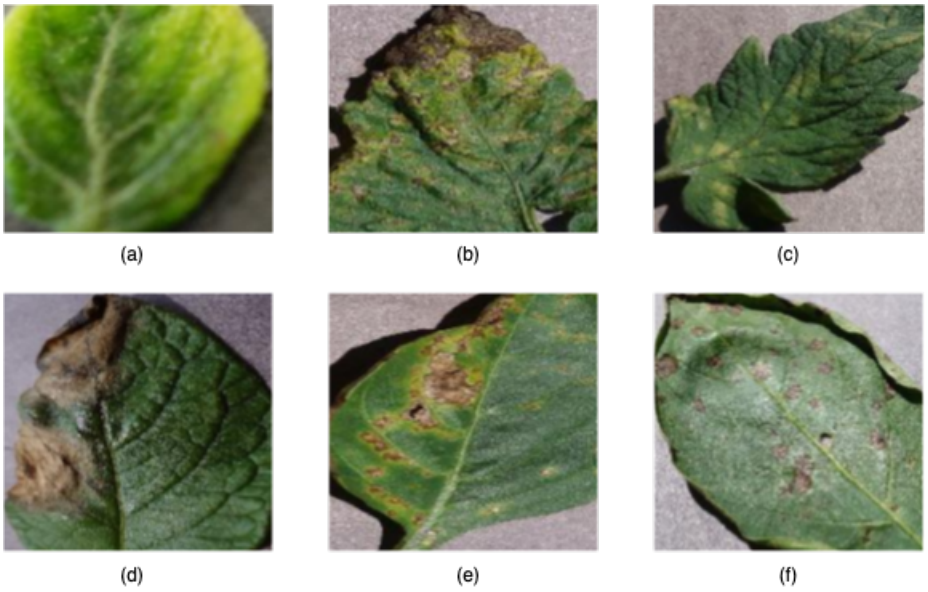


Fig. 2. Plant leaf sample images: (a) Yellow Leaf Curl Virus (b) Septoria Leaf Spot (c) Leaf Mold (d) Late Blight (e) Early Blight (f) Bacterial Spot

tures the spatial hierarchies, texture, color, geometric profiles and disease patterns.It eliminates the need for manual feature engineering.

2. **CNN+CBAM Approach:** Convolutional Block Attention Module (CBAM) is further an extension of this approach where attention is paid to the most informative channels and spatial regions. This allows the model to better tell the difference between very subtle differences between healthy and diseased tissue, even if there are very few visual differences.

4.5 Classification

These characteristics are extracted and relayed to the final levels of the network. These dense layers or rather, other layers of classification extract features and label each input image with a predetermined number of class labels (e.g. healthy or a specific type of disease). When applied to an employment process, the model predictions are compared to the actual labels and various performance measures described below are estimated to get an overall performance estimate. The process usually involves setting aside a certain percentage of the images in order to validate and test the system, thereby assuring a detailed analysis of the system's capacity to classify unknown and unseen locations correctly

4.6 Performance Measures

A classifier is also tested and trained to evaluate the performance. In this regard, four performance metrics—accuracy, precision, recall, and the F1-score, represented by A, P, R, and F1, respectively —have been integrated to assess both the CNN and CNN+CBAM classifiers using the feature set on the testing set. The ROC curve is also used to determine the performance of the trained classifier. This information is obtained by computing the number of true positives (P1), false positives (P2), true negatives (N1), and false negatives (N2) that emerge from the confusion matrix.

$$A = \frac{P1 + N1}{P1 + N1 + P2 + N2} \tag{1}$$

$$P = \frac{P1}{P1 + P2} \tag{2}$$

$$R = \frac{P1}{P1 + N2} \tag{3}$$

$$F1 = 2 * \frac{P * R}{P + R} \tag{4}$$

5 Experimental Result Analysis

The efficiency of the plant disease classification performed with CNN and CNN combined with CBAM is discussed, and it is found that the CNN + CBAM classification technique is more qualitative than the CNN classification technique. There are 1,885 images of leaves from potatoes, peppers, and tomatoes, with six label classes, out of which 395 are healthy leaves. This data is split into two subsets, with 90% of the data used for training and 10% used for testing. These classifiers are being installed in Python 3.13, running on the Windows 11 operating system. It is powered by a 64-bit operating system, 16.00 GB of DDR4 RAM, and features an Intel Core i5 processor with a speed of 4.5 GHz. The Python language, the machine learning library in Python, and Pandas are used. The average outcomes of the classifications given by CNN and CNN+CBAM, respectively, for all diseases are tabulated in Table 2. As can be seen, there is a substantial improvement in all figures, particularly in Precision and Recall statistics, after the incorporation of CBAM.

Table 2: Comparison of Performance of CNN and CNN+CBAM

S.no	Performance metric	CNN	CNN+CBAM
1	A	92.8%	98.0%
2	P	91.7%	96.0%
3	R	91.3%	95.6%
4	F1	91.5%	95.8%

This implies that the model containing CBAM has a higher detection rate for any crop disease symptoms and can distinguish among parallel types of diseases. Whereas, the Cropwise accuracy

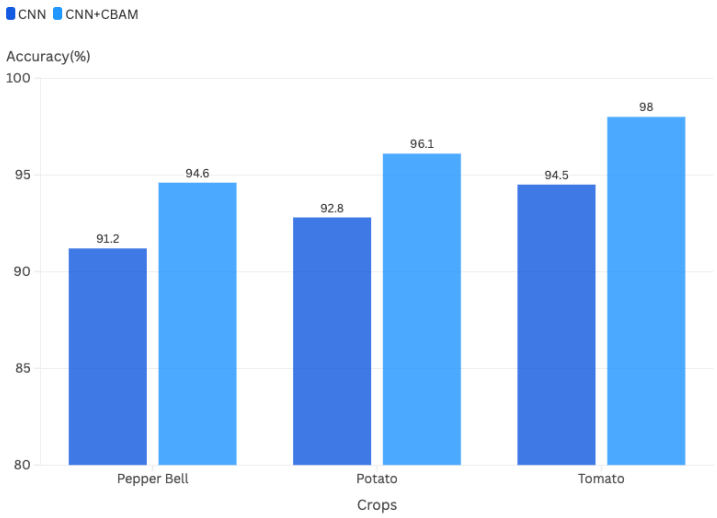


Fig. 3. Cropwise Disease Detection Accuracy Comparison

obtained by CNN and CNN+CBAM is demonstrated in Fig. 3. We simply tabulate them into healthy and diseased categories. Our CNN-based architecture proved to be faster in inference and relatively accurate compared to other architectures. In particular for tomato disease detection, a comparison with existing state-of-the-art models like MobileNet[21], CNN[22], Faster R-CNN[23] and YOLOv8[24], as shown in table3.

Table 3: Comparsion of achieved accuracy by proposed model with existing models for tomato

S.no	Model	Year	Accuracy
1	MobileNet	2024	96%
2	CNN	2024	86.21%
3	Faster R-CNN	2023	0.58 mAP
4	YOLOv8	2024	0.93 mAP
5	Proposed Model	2025	98%

6 Conculsion

The rapid and accurate detection of plant diseases has gained increased significance in the face of growing demands for food security and environmental sustainability. To address this long-standing agricultural issue, we utilize deep learning, specifically the CNN model and CBAM. The

outcomes of our findings provide us with additional data regarding the usefulness of such models. The CNNs+CBAM performed quite well, with 98% of the responses being correct. Nevertheless, it may be inefficient when it comes to low-resolution images or images that were not included in the training set and are located in large fields. Despite all these facts, its simplicity makes it a great option to use in real life and for fast diagnostics of the disease. especially real-field data. controlled environment.

7 Future Scope

Future work will include:

1. Incorporate real-field leaf images, as most of the models are trained and tested on the PlantVillage dataset, which mainly contains single leaves captured under controlled conditions. Such images have simple backgrounds and stable lighting, whereas real fields involve clutter, overlapping leaves, changing illumination, occlusions and motion blur, all of which can reduce accuracy when a model trained only on laboratory images is used in practice. To narrow this gap, future work will include field images and apply domain adaptation methods that align features between controlled and real-world data.

2. employ domain adaptation techniques such as fine-tuning on field data, adversarial domain adaptation, or self-supervised pretraining on unlabeled field images, which will match features between controlled and real-world data, to improve robustness and scalability.

3. for deployment of the model, work will be done on compression techniques—such as pruning, quantization, or knowledge distillation—to make the proposed approach fit for microcontroller-class platforms and low-power devices, as the current architecture is trained and evaluated on a desktop setup only.

4. To utilize multimodal fusion approaches that combine leaf images with environmental or sensor data (e.g., temperature, humidity, soil parameters, or phenological stage) will be utilized to improve robustness under varying seasons and management practices.

References

[1] Chen J, Chen J, Zhang D, Sun Y, Nanehkaran YA. Using deep transfer learning for image-based plant disease identification. Computers and electronics in agriculture. 2020;173:105393.

[2] Bouguettaya A, Zarzour H, Kechida A, Taberkit AM. A survey on deep learning-based identification of plant and crop diseases from UAV-based aerial images. Cluster Computing. 2023;26(2):1297-317.

[3] Prakash C, Singh LP, Gupta A, Lohan SK. Advancements in smart farming: A comprehensive review of IoT, wireless communication, sensors, and hardware for agricultural automation. Sensors and Actuators A: Physical. 2023;362:114605.

[4] Ferehan N, Haqiq A, Ahmad MW. Research Article Smart Farming System Based on Intelligent Internet of Things and Predictive Analytics. 2022.

[5] Borhani Y, Khoramdel J, Najafi E. A deep learning based approach for automated plant disease classification using vision transformer. Scientific Reports. 2022;12(1):11554.

[6] Li L, Zhang S, Wang B. Plant disease detection and classification by deep learning—a review. IEEE access. 2021;9:56683-98.

[7] Mohanty SP, Hughes DP, Salathé M. Using deep learning for image-based plant disease detection. Frontiers in plant science. 2016;7:215232.

[8] Barbedo JGA. Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification. Computers and electronics in agriculture. 2018;153:46-53.

[9] Pang B. Classification of images using EfficientNet CNN model with convolutional block attention module (CBAM) and spatial group-wise enhance module (SGE). In: International conference on image, signal processing, and pattern recognition (ISPP 2022). vol. 12247. SPIE; 2022. p. 34-41.

[10] Al-Gaashani MS, Samee NA, Alnashwan R, Khayyat M, Muthanna MSA. Using a Resnet50 with a kernel attention mechanism for rice disease diagnosis. Life. 2023;13(6):1277.

[11] Al-gaashani MS, Shang F, Muthanna MS, Khayyat M, Abd El-Latif AA. Tomato leaf disease classification by exploiting transfer learning and feature concatenation. IET Image Processing. 2022;16(3):913-25.

[12] Thakur PS, Chaturvedi S, Khanna P, Sheorey T, Ojha A. Vision transformer meets convolutional neural network for plant disease classification. Ecological Informatics. 2023;77:102245.

[13] Li Y, Fayaz M, Danish S, Tightiz L, Wang H, Nguyen TN, et al. End-to-end plant disease detection using transformers with collaborative hybrid assignment trainings. Applied Soft Computing. 2025:114137.

[14] Devi K, Benila S, Lekshmi K, Omana J. HVAF-XAI-Net: A Multimodal Explainable AI Framework for Accurate Mulberry Leaf Disease Detection. Smart Agricultural Technology. 2025:101377.

[15] Nasir N, Ramzan S, Ali B, Jabbar S, Raza A, Kim C, et al. Interpretable deep multimodal-based tomato disease diagnosis and severity estimation. Scientific Reports. 2025;15(1):37919.

[16] Patil RG, More A. Grape leaf disease diagnosis system using fused deep learning features based system. Procedia Computer Science. 2024;235:372-82.

[17] Yadav APS, Thapliyal N, Aeri M, Kukreja V, Sharma R. Advanced Deep Learning Approaches: Utilizing VGG16, VGG19, and ResNet Architectures for Enhanced Grapevine Disease Detection. In: 2024 11th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO); 2024. p. 1-4.

[18] Unal Y. Integrating CBAM and Squeeze-and-Excitation Networks for Accurate Grapevine Leaf Disease Diagnosis. Food Science & Nutrition. 2025;13(6):e70377.

[19] Febriantono A, Soeparno H, Gaol FL, Arifin Y. Deep Learning Approach with EfficientNet-CBAM for Apple Disease Detection. In: 2025 International Conference on Computer Sciences, Engineering, and Technology Innovation (ICoCSETI). IEEE; 2025. p. 170-5.

[20] Rahman M, Ullah M, Devnath R, Chowdhury T, Rahman G, Rahman M. Cotton leaf disease detection: an integration of CBAM with deep learning approaches. Int J Comput Appl. 2025;975:8887.

[21] Khalid MM, Karan O. Deep learning for plant disease detection. International Journal of Mathematics, Statistics, and Computer Science. 2024;2:75-84.

[22] Rastogi P, Dua S, Dagar V, et al. Early Disease Detection in Plants using CNN. Procedia Computer Science. 2024;235:3468-78.

[23] Giakoumoglou N, Pechlivani EM, Frangakis N, Tzovaras D. Enhancing tuta absoluta detection on tomato plants: ensemble techniques and deep learning. AI. 2023;4(4):996-1009.

[24] Uygun T, Ozguven MM. Determination of tomato leafminer: Tuta absoluta (Meyrick)(Lepidoptera: Gelechiidae) damage on tomato using deep learning instance segmentation method. European Food Research and Technology. 2024;250(6):1837-52.