



Attention-Based Efficient Net for Grape Leaf Disease Recognition

Rahul¹, Gurpreet Singh², Amandeep³

Department of Computer Science & Engineering, Guru Kashi University, Bathinda, India

Abstract: Early identification and treatment of diseases of plants is of great significance in enhancing agri-cultural productivity and reducing economic loss. Conventional techniques of disease identification depend upon expert knowledge and inspection by hand, which can be lengthy and prone to mistakes made by hu-mans. At the present study, that we propose an attention-based deep learning based method for grape leaves classification of diseases using an Efficient-NetB3 with a Convolutional Block Attention Module (CBAM). The PlantVillage grape dataset is used in this study. Data augmen-tation techniques were used to improve the generalisation performance of the model and reduce overfitting. Transfer learning approach was used with ImageNet pretrained EfficientNetB3 and two stage training strategy (feature extraction and fine-tuning). Our model achieved a test accuracy of 99.67% which is better than the MobileNetV2 (72.18%), DenseNet121 (77.74%) and ResNet50 (98.53%). We have also used Grad-CAM visualisation to interpret the model pre-dictions and highlight the disease affected regions. Experiments show that incorporating the use of CBAM significantly improves the feature representation and classification performance, so that our framework can be applied to intelligent agricultural disease diagnosis systems.

Keywords: Grape Disease Detection, EfficientNetB3, CBAM, Transfer Learning, Deep Learning, PlantVil-lage Dataset.

I. INTRODUCTION

Farming is an essential element for global food security and economic development. Grapes are among the world's most commercially crucial and widely grown crops. Grapes are extensively used in the fresh market, as well as in the production of juice, sultanas and wine. However grape growing is highly susceptible to many diseases that greatly affect the quality of the crop, yield and market value . Grape diseases involving leaves such as Esca (Black Measles), Black Rot and Leaf Blight are common and can easily spread throughout vine-yards causing considerable economic losses if not detected and treated at an early stage. The conventional methods of disease diagnosis predominantly depend on the manual visual inspection by agricultural experts. Though manual inspection may be effective in some contexts, it is time-consuming, labor-intensive, subjec-tive, and often un-available to farmers in remote regions. Furthermore, visual symptoms of many diseases are often similar, making it difficult even for experienced experts to correctly diagnose them. The impor-tance of the auto-mated disease detection system is increasing which can diagnose accurately, quickly and at low cost. Recent advances in techniques changed the paradigms of agricultural monitoring and disease man-agement. CNNs have been demonstrated to be very effective for image classification and object recognition tasks. Several re-searchers have used deep CNN architecture such as ResNet, VGGNet, MobileNet, DenseNet and EfficientNet for plant disease classification. These methods have shown promising results, but may not be effective when disease symptoms are subtle, isolated, or masked by background alterations. To address these limitations, attention mechanism s have recently been incorporated into deep learning models to en-hance feature repre-sentation and focus on disease-

relevant regions. The attention modules in neural networks can enable neural networks to focus on informative features and ignore irrelevant information, which can not only improve the classification accuracy but also enhance the interpretability of the model. CBAM, is a very popular attention module which is light weight and sequentially applies channel and spatial attention. CBAM is concerned with the important feature channels and the disease-affected regions so that it can enhance the discriminative capability of convolutional neural networks. In our project, we developed a grape plant diseases classification framework based on the EfficientNetB3 model combined with the CBAM. The backbone system is selected as the EfficientNetB3 that achieves a good balance between the classification accuracy and computational efficiency via compound scaling. Transfer learning is used to transfer the knowledge learned from large-scale image datasets and fine-tuning is used to adapt the model to the task of grape disease classification. In addition, Mapping (Grad-CAM) is also used to provide visual explanations of model predictions, which enhances the interpretability. The proposed framework is tested on PlantVillage grape leaf disease dataset which has four classes namely Black Rot, Esca (Black Measles), Leaf Blight and Healthy leaves. Test results demonstrate that the designed EfficientNetB3-CBAM model outperforms several well-known deep learning architectures, such as MobileNetV2, DenseNet121 and ResNet50 in terms of classification performance. The primary contributions of this work can be summarised in the following:

- Development of an attention-enhanced grape leaf disease classification framework based on EfficientNetB3 and the Convolutional Block Attention Module (CBAM).
- Application and fine-tuning strategies to enhance the feature extraction capability and classification results.
- Comparison DenseNet121, MobileNetV2 and ResNet50.
- Integration Activation Mapping (Grad-CAM) to improve model interpretability and visual explanations of the classification decisions.
- Classification accuracy of 99.67% shows the effectiveness and robustness of the proposed framework for intelligent grape leaf disease diagnosis in precision agriculture.

II. LITERATURE REVIEW

The latest development in artificial intelligence and deep learning techniques has greatly improved auto-mated grape leaf disease detection. Several studies have demonstrated the use of convolutional neural network systems (CNNs), residual networks, transfer based learning and hybrid and standard models for accurately detecting disease symptoms from leaf images. Comparative studies have shown that hybrid deep learning models and ResNet-based architectures provide higher classification accuracy than traditional CNNs by extracting more discriminative features from diseased leaves [1], [2]. Furthermore, many extensive reviews have reported the growing use of deep transfer learning approaches for plant disease classification, due to their ability to achieve high accuracy with less training data [3]. Research has also been extended to real-time disease monitoring using autonomous robotic systems that can perform vineyard scouting and disease surveillance, thereby facilitating precision agriculture practices [4]. Moreover, image segmentation techniques have been identified as a vital pre-processing step to define infected areas and improve disease recognition performance [5]. To address deployment issues in resource-limited environments, compact deep learning models and optimized network architectures have been proposed to achieve efficient and real-time disease detection on edge devices with less computation [6], [7]. Also, the study results suggest that transfer learning and deep neural networks are more effective than traditional machine learning algorithms in detecting diseases in grape leaves [8]. Recent research efforts to improve the performance of the disease detection have been directed to exploit advanced architectures, image enhancement techniques, and object detection models. The use of data augmentation in conjunction with YOLO-based frameworks has helped improve the robustness and the generalization capability of the disease detection systems for different plant species [9]. As super-resolution techniques, residual skip networks have been used to improve image quality before classification, thus improving disease

identification accuracy [10]. Attention mechanisms have been added to deep learning models to find affected areas and improve their feature extraction capability [11]. Recently, state-of-the-art CNN architectures and deep convolutional networks [12], [13] have exceeded the detection of grape leaf disease. Inception-ResNet feature fusion networks and transformer-based modules have recently achieved impressive classification results by effectively capturing local and global features of images [14]. The importance of early diagnosis for crop loss reduction and vineyard productivity increase has been highlighted in early dis-ease detection frameworks [15]. Moreover, some other reviews have confirmed the efficacy of deep learning techniques for grapevine disease diagnosis and reported their ongoing development towards more accurate and robust systems [16]. In addition, field trials have also confirmed the practical application of deep learning-based approaches for disease detection of grapevine downy mildew, under real vineyard conditions, in spite of illumination changes and complex backgrounds [17]. Overall, the literature suggests that the combination of deep learning, transfer learning, attention mechanisms, transformer architectures and intelligent monitoring systems has significantly improved the grape leaf disease detection. However, chal-lenges concerning model generalization, field deployment and computational efficiency remain as future research directions.

III. RESEARCH METHODOLOGY

This section describes the dataset, preprocessing procedures, data augmentation techniques, EfficientNetB3 backbone architecture, Convolutional Block Attention Module (CBAM), proposed classification framework, and training strategy employed for grape leaf disease classification.



Figure 1: workflow for research

3.1 Dataset

The experiments were carried via the public dataset of grape leaf diseases PlantVillage. The dataset includes RGB images of grape leaves taken according to controlled conditions with good quality and classified in four classes representing healthy and diseased leaves. Disease classes are Leaf Blight (Isariopsis Leaf Spot), Black Rot, Esca (Black Measles) and the fourth class is healthy leaves. In this study, the PlantVillage dataset consisting of 4,062 images of grape leaves classified into four classes is used. Esca (Black Measles) is the class with the highest number of images, with a total of 1,383 images, followed by Black Rot with 1,180 images and Leaf Blight with 1,076. Class Healthy has 423

images. This dataset includes different types of grape leaf samples (diseased and healthy) to train and test deep learning models. The dataset was randomly split into validation(10%), training(75%), and test(15%) sets for solid model evaluation.

3.2 Data Preprocessing

All of the images were scaled down to 300×300 pixels to be adapted to EfficientNetB3 architecture input. The image normalisation was performed using the EfficientNetB3 pre-processing function, which corresponds to ImageNet pre-training statistics. Also, the preprocessing phase includes automatic label encoding and directory-based image loading to simplify the data management process. These processes help to improve the computational performance and to maintain a consistent image representation during the training stage.

3.3 Data Augmentation

Data augmentation was applied for improving the generalisation of the model and the diversity of the dataset. Agricultural datasets are often characterized by changes of lighting conditions, viewing angles, leaf orientations and environmental conditions. Therefore, several augmentation techniques were applied during training. The augmentation pipeline was:

- Random rotation
- Horizontal flipping
- Width shifting
- Height shifting
- Zoom transformation
- Brightness adjustment

These transformations generate more image variations without changing the disease characteristics, so that overfitting is prevented and the model robustness is increased.

3.4 Perposed EfficientNetB3 Architecture

EfficientNetB3 model was selected as the primary model for feature extraction as it offers the best balance between complexity and classification accuracy. The compound scaling method coupled with MBConv blocks with squeeze-and-excitation mechanisms allows efficient feature extraction with a smaller number of parameters. Transfer learning was applied by utilizing a pretrained EfficientNetB3 model on ImageNet and its original classification layers were replaced with custom layers for grape leaf disease classification.

3.5 Convolutional Block Attention Module

We used CBAM in the EfficientNetB3 framework to further improve the learning of discriminative characteristics. CBAM is an elementary attention module which sequentially infers attention maps along two separate axes, channel and spatial, after which the learned feature maps are enhanced.

3.5.1 Channel Attention

The channel attention mechanism uses global average pooling and global max pooling operations to aggregate spatial information and to identify the most informative feature channels. The generated attention weights highlight disease-related channels and suppress irrelevant features. The channel attention map can be calculated as follows:

$$M_c(F) = \sigma(\text{MLP}(\text{AvgPool}(F)) + \text{MLP}(\text{MaxPool}(F))) \quad (1)$$

where σ denotes the sigmoid activation function and F represents the input feature map.

3.5.2 Spatial Attention

After channel refinement, spatial attention is used to highlight regions of interest for disease in the image. Average pooling and max pooling are done along the channel axis and then convolution operations are used to generate a spatial attention map. The spatial attention map is obtained as:

$$M_s(F) = \sigma f^{7 \times 7} [\text{AvgPool}(F); \text{MaxPool}(F)] \quad (2)$$

where σ denotes the sigmoid activation function and $f^{7 \times 7}$ represents a convolution operation with a kernel size of 7×7 .

3.6 Proposed EfficientNetB3-CBAM Framework

The proposed framework is a combination of EfficientNetB3 and CBAM to enhance disease-specific feature extraction and classification accuracy. The overall working procedure of the work is as follows:

1. Image acquisition
2. Image pre-processing and normalization
3. Feature extraction using EfficientNetB3
4. Feature refinement using the Convolutional Block Attention Module (CBAM)
5. Global Average Pooling (GAP)
6. Dropout regularization
7. Fully connected dense layer
8. Softmax classification layer

The architecture enables a model to concentrate on disease affected leaf regions while being computationally efficient. The proposed framework combines attention mechanisms and transfer learning to efficiently detect the subtle visual symptoms of grape diseases.

3.7 Training Strategy

To improve the model performance, a dual-step transfer learning method was employed. Initial, we froze the backbone of EfficientNetB3 and trained only the new classification layers. The next phase is to unfreeze the last layers of EfficientNetB3 and fine-tune them with a smaller learning rate to better fit the features of grape leaf disease. The model was trained using categorical cross entropy loss, early stopping, the adam optimiser and model checkpoint to avoid overfitting. This resulted in a system that was able to reliably and accurately diagnose grape leaf diseases.

IV. EXPERIMENTAL SETUP

The following section summarises the computational setup, training parameters and evaluation metrics used to measure the effectiveness proposed EfficientNetB3-CBAM framework for grape leaf diseases standard classification.

4.1 Hardware and Software Environment

The model proposed was developed in Python as well as trained in TensorFlow-Keras deep learning framework. Experimental analyses were performed in a cloud-based computational environment with Graphics Processing Unit (GPU) acceleration to facilitate efficient model training and inference. The software environment consists of Numpy, a Pandas, TensorFlow, Matplotlib, Keras, Seaborn and Scikit-learn based libraries.

4.2 Training Hyperparameters

The EfficientNetB3-CBAM model was trained with a two-step transfer learning strategy. The first step was to freeze the EfficientNetB3 backbone and train only the added classification layers. The second step was fine-tuning, which unfreezes some of the higher layers of the backbone network and re-trains the model with a smaller learning rate. We opt for the Adam optimiser because of its adaptable learning support and good convergence standard property. As the problem is multi class classification problem we used the categorised cross entropy function for loss. The images were resized to 300x300 pixels and the model was trained with batch size 32. For classification with multiple classes, the loss function was categorical cross entropy loss function with Adam optimiser. Transfer learning was performed with an EfficientNetB3 model trained over the ImageNet database with a learning rate of (1×10^{-4}) and (1×10^{-5}) in initial and fine tuning, respectively. To prevent over-fitting, dropout rate was set to 0.4. The output layer used Softmax activation function for classes images classification. During training, In order to prevent over-fitting and to improve generalisation, we applied early stopping and model checkpoint techniques. The best performing model was saved based on the validation accuracy and early stopping was used to stop the training process if no improvement was observed in the validation accuracy.

4.3 Evaluation Metrics

The proposed framework is tested using some popular classification metrics like Accuracy, Precision, Re-call and F1-score.

Accuracy -Accuracy is the ratio between the number of samples correctly classified and the total number of samples and is calculated as follows:

$$Accuracy = (TP + TN) / (TP + TN + FP + FN) \quad (3)$$

where TP , TN , FP , and FN are true positives, true negatives, false positives and false negatives respect.

Precision - Precision is the ratio of correctly predicted positive samples to the total predicted positive samples and is computed as follows:

$$Precision = TP / (TP + FP) \quad (4)$$

Recall measures the ability of the model to identify the positive samples properly, and is calculate

$$Recall = TP / (TP + FN) \quad (5)$$

The F1-score is the harmonic mean of precision and recall and is calculated as:

$$F1 = (2 * Precision * Recall) / (Precision + Recall) \quad (6)$$

These evaluation metrics together provide a comprehensive assessment of the classification performance, especially in multi-class disease identification tasks.

V. RESULTS AND DISCUSSION

Results from experiments of the proposed EfficientNetB3-CBAM framework have shown. The model performance was evaluated by standard classification metrics, confusion matrix analysis, comparison studies with existing architectures and explainability analysis with Grad-CAM visualisation.

5.1 Training Performance Analysis

In the initial phase of the training stage, the model rapidly learned the discriminative features from the grape leaf images and the validation performance was stable. Further improvement in classification accuracy was achieved by fine-tuning the feature representations of pre-trained EfficientNetB3 to disease-specific characteristics. Figure 2 shows the accuracy curves. The two accuracy curves are smooth with small fluctuations, which shows that the model has good learning and generalisation ability.

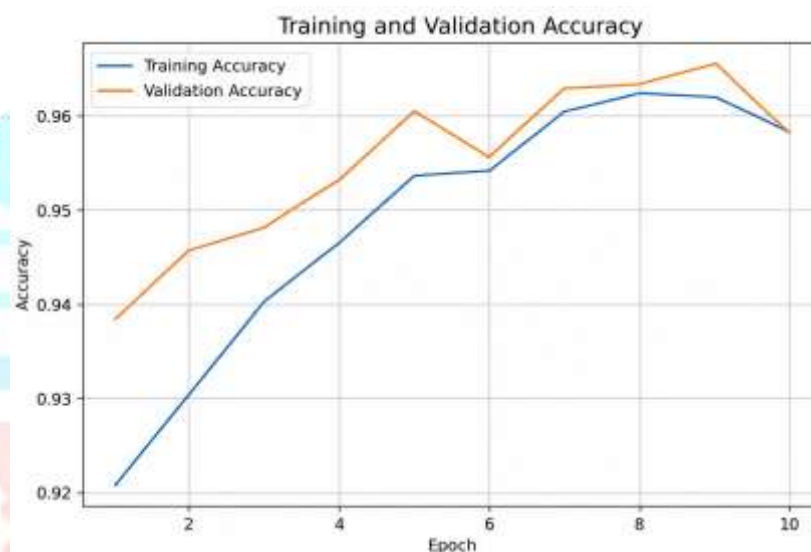


Figure 2: Accuracy curve

Loss curves are also shown in Figure 3. The small gap between the training and validation losses indicates the proposed model is not overfitted.

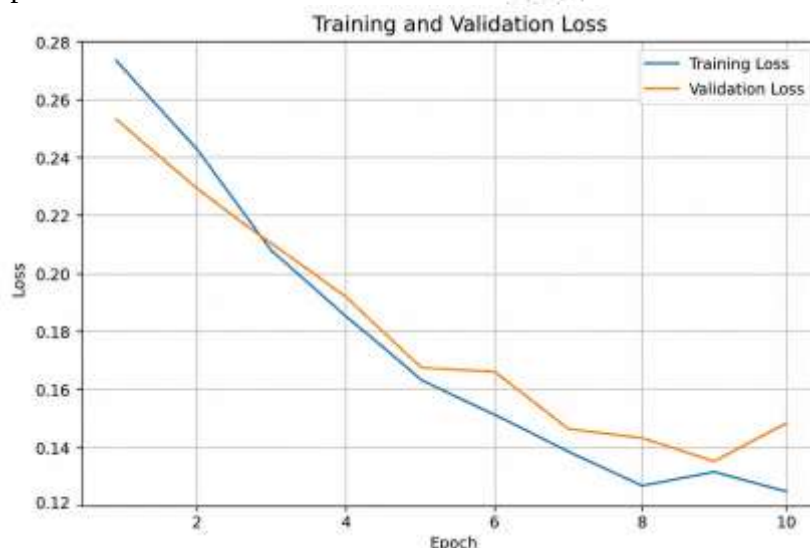


Figure 3: Loss curve

5.2 Overall Classification Performance

Proposed EfficientNetB3-CBAM model has achieved good classification performance on PlantVillage dataset. The final evaluation's accuracy rate was 99.67% which showed that the model capable to accurately detects the diseased and normal healthy leaves of grapes. The effectiveness of the transfer learning as well as the attention-guided feature refinement is proved by the high classification accuracy. The CBAM attention allows the model to pay attention to the disease-specific regions and suppress the background information.

5.3 Classification Report

A detailed classification report was generated to verify the performance at the class level.

Table 1: Classification Performance

Class	Precision (%)	Recall (%)	F1-Score (%)
Black Rot	99.44	99.44	99.44
Esca (Black Measles)	99.52	99.52	99.52
Healthy	100.00	100.00	100.00
Leaf Blight	100.00	100.00	100.00

The results show good recognition for all the disease classes. The model achieved perfect classification of Leaf Blight and Healthy leaf classes and the Black Rot and Esca classes also presented very high recognition rates. The good and balanced performance of the model for all classes confirms the strength as well as reliability of the proposed framework.

5.4 Confusion Matrix Analysis

The confusion matrix further reveals the prediction behavior of each class. The confusion matrix of the testing dataset is shown in Fig.3. The confusion matrix shows that most of the images of grape leaves were classified correctly with a few misclassification cases, indicating the model's high discriminative power among the classes of diseases. The low confusion between the classes Black Rot and Esca also indicates that the attention mechanism was able to successfully capture subtle features of the diseases.

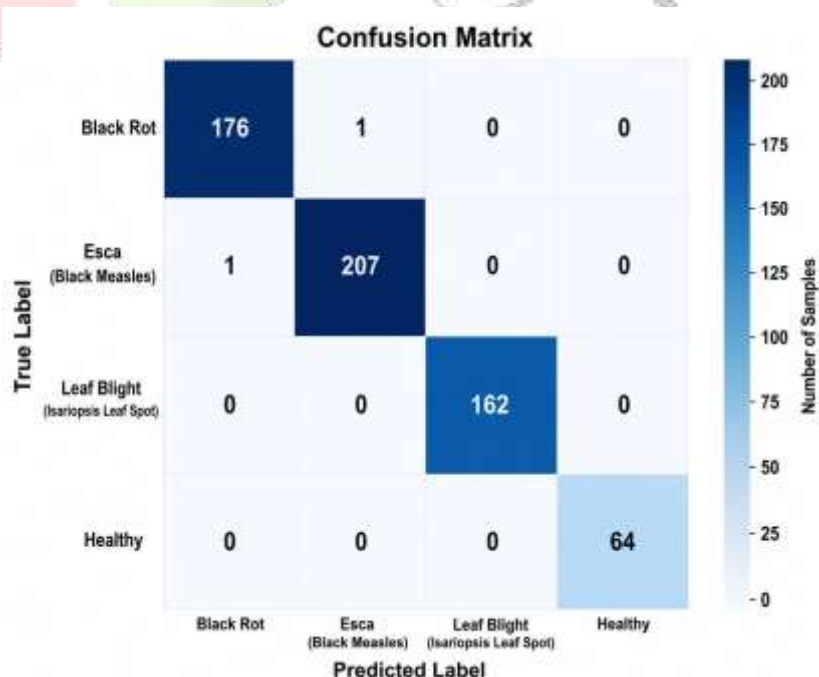


Figure 4: confusion matrix

5.5 Comparative Evaluation with Existing Architectures

We compare the tests with three well-known deep learning architectures, MobileNetV2, DenseNet121 and ResNet50. Each of the models undergo testing and training under similar experimental settings.

Table 2: Comparison of Classification Accuracies Obtained by Different Models

Model	Accuracy (%)
MobileNetV2	72.18
DenseNet121	77.74
ResNet50	98.53
Proposed EfficientNetB3-CBAM	99.67

The proposed model achieved the best classification accuracy among all the tested architectures. The proposed framework demonstrated significant performance improvements compared to MobileNetV2 and DenseNet121. Although ResNet50 showed competitive results, EfficientNetB3-CBAM still outperformed it by more than one percentage point. The performance gains achieved by the proposed framework are mainly due to three factors:

- Efficient feature extraction capability of EfficientNetB3 achieved through compound scaling.
- Enhanced discriminative representation learning provided by the Convolutional Block Attention Module (CBAM), which focuses on disease-relevant regions.
- Improved adaptation to grape leaf disease characteristics through transfer learning and fine-tuning strategies.

5.6 Grad-CAM Explainability Analysis

Interpretability is an important requirement for practical agricultural decision-support systems. Therefore, Grad-CAM visualization was employed to investigate the regions influencing model predictions. Figure 5 presents Grad-CAM heatmaps generated for different disease categories.

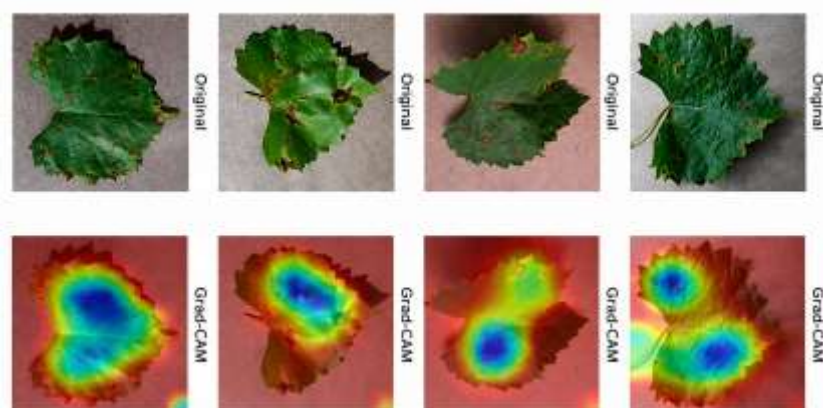


Figure 5: Grad-CAM

The generated heatmaps show that the proposed model consistently attends the disease affected regions, e.g., lesions, discolorations, necrotic spots and infection boundaries. The healthy leaf samples showed distributed activation patterns over intact leaf structures, while the diseased samples focused on the symptomatic areas. These observations confirm that the model learns biologically meaningful visual features instead of irrelevant background information.

5.7 Discussion

Experimental results shown that the EfficientNetB3-CBAM framework is quite suitable for grape leaf dis-ease classification. The model obtained an accuracy of 99.67% by combining transfer learning, attention mechanism and fine-tuning, which demonstrates the model's excellent ability of disease recognition and good generalization ability. The incorporation of attention mechanisms has improved the network's ability to extract features and improve classification

VI. CONCLUSION

In the present work, we outline an automated grape leaves disease identification structure based on Effi-cientNetB3 with Convolutional Block Attention Module (CBAM). The evaluation on the PlantVillage dataset with four classes, Leaf Blight, a Black Rot, Esca Species (Black Measles) and healthy leaves was conducted using the proposed model with transfer learning technique, attention-guided extraction of features, and fine-tuning methods. Model contain accuracy of 99.67% that better than the popular architec-tures namely Mo-bileNetV2, DenseNet121 and ResNet50. The CBAM module improved feature representa-tion by focusing on disease-related regions, and Grad-CAM provided visual explanations for enhancing the model's interpretabil-ity and reliability. The results show that the EfficientNetB3-CBAM framework is an ac-curate, robust and com-putationally efficient solution for grape leaf disease diagnosis. Although trained with images taken under controlled conditions, it has great potential for precision agriculture applications. Future work will include testing the framework on real-world vineyard datasets, exploring ad-vanced attention and transformer-based architectures, enabling mobile deployment, and extend-ing the system to support multiple crops and plant diseases.

References

- [1] J. D. Palkar and A. S. Deshpande, "Performance Evaluation of CNN, ResNet50, and Hybrid Architectures for Automated Grape Leaf Disease Detection," *Applied Fruit Science*, vol. 68, no. 2, p. 98, 2026.
- [2] C. Kosaraju, M. Sridevi, K. Kalyani, C. Yasaswini, S. Sameera, and B. Katyayani, "Grape Plant Disease Detection Using Residual Neural Network (ResNet50V2)," in *Advances in Science and Technology*, CRC Press, pp. 46–52, 2026.
- [3] N. Hajraoui, M. Azrour, Y. Farhaoui, and A. El Allaoui, "A Review of AI and Deep Transfer Learning Methods for Plant Disease Detection and Classification," 2026.
- [4] E. Liu, K. M. Gold, L. Cadle-Davidson, K. Kanaley, D. Combs, and Y. Jiang, "PhytoPatholoBot: Autonomous Ground Robot for Near-Real-Time Disease Scouting in the Vineyard," *Journal of Field Robotics*, vol. 43, no. 1, pp. 442–453, 2026.
- [5] S. Duhan, P. Gulia, N. S. Gill, and C. Gupta, "A Comprehensive Review of Image Segmentation Techniques for Plant Disease Detection in Precision Agriculture: Models, Trends, and Challenges," *Archives of Computational Methods in Engineering*, pp. 1–36, 2026.
- [6] R. Deepa, M. P. Mathew, and S. Baskar, "L-Net: A Lightweight CNN Framework for Sustainable Multicrop Leaf Disease Detection and Classification on Edge Devices," *Sustainable Food Technology*, 2026.
- [7] J. Rahaman, P. Paul, A. Chowdhury, and M. Quamruzzaman, "A Customized MobileNetV3Large-Based Deep Learning Framework for Plant Disease Detection," *Discover Artificial Intelligence*, 2026.
- [8] P. G. Dhavane, T. G. Mangave, T. S. Deshmukh, and P. P. Kajale, "A Review on Deep Learning and Transfer Learning Models: Grape Leaf Disease Detection," in *Proc. 3rd Int. Conf. Inventive Computing and Informatics (ICICI)*, IEEE, pp. 1013–1018, 2025.
- [9] R. Kaur, U. Mittal, A. Wadhawan, A. Almogren, J. Singla, S. Bharany *et al.*, "YOLO-LeafNet: A Robust Deep Learning Framework for Multispecies Plant Disease Detection with Data Augmentation," *Scientific Reports*, vol. 15, no. 1, p. 28513, 2025.

- [10] P. V. Yeswanth, S. Deivalakshmi, S. George, and S. B. Ko, “Residual Skip Network-Based Super-Resolution for Leaf Disease Detection of Grape Plant,” *Circuits, Systems, and Signal Processing*, vol. 42, no. 11, pp. 6871–6899, 2023.
- [11] W. Guo, Q. Feng, X. Li, S. Yang, and J. Yang, “Grape Leaf Disease Detection Based on Attention Mechanisms,” *International Journal of Agricultural and Biological Engineering*, vol. 15, no. 5, pp. 205–212, 2022.
- [12] B. Liu, Z. Ding, L. Tian, D. He, S. Li, and H. Wang, “Grape Leaf Disease Identification Using Improved Deep Convolutional Neural Networks,” *Frontiers in Plant Science*, vol. 11, p. 1082, 2020.
- [13] R. Nagi and S. S. Tripathy, “Deep Convolutional Neural Network Based Disease Identification in Grapevine Leaf Images,” *Multimedia Tools and Applications*, vol. 81, no. 18, pp. 24995–25006, 2022.
- [14] R. Karthik, R. Menaka, S. Ompirakash, P. B. Murugan, M. Meenakshi, S. Lingaswamy, and D. Won, “GrapeLeafNet: A Dual-Track Feature Fusion Network with Inception-ResNet and Shuffle-Transformer for Accurate Grape Leaf Disease Identification,” *IEEE Access*, vol. 12, pp. 19612–19624, 2024.
- [15] R. M. Math and N. V. Dharwadkar, “Early Detection and Identification of Grape Diseases Using Convolutional Neural Networks,” *Journal of Plant Diseases and Protection*, vol. 129, no. 3, pp. 521–532, 2022.
- [16] S. A. S. Aldaya and S. S. Abu-Naser, “Deep Learning for Grapevine Disease Detection,” 2025.
- [17] J. Boulent, M. Beaulieu, P. L. St-Charles, J. Théau, and S. Foucher, “Deep Learning for In-Field Image-Based Grapevine Downy Mildew Identification,” in *Precision Agriculture '19*, pp. 141–148, Wageningen Academic Publishers, 2019.

