APPLIED ARTIFICIAL INTELLIGENCE AND ENGINEERING JOURNAL
E-ISSN: 0000-0000, DOI: 00.000/XXXX.YYYY.II.NNN
Vol. 2 No. 1 2026 pp. 59-70

An Automated CNN Framework for Multi-Class Apple Leaf Disease

Nilam Cahaya Safitri[1], Raudhatul Husna[2], Rahma Aliya[3], Andini Putri[4*]

[1,2] Jurusan Teknologi Informasi Komputer, Politeknik Negeri Lhokseumawe, Indonesia
[3,4] Fakultas Pertanian, Universitas Syiah Kuala, Indonesia

*Corresponding Author: andiniputri6432@gmail.com

Article info: Received 07/03/2026, Revised 18/03/2026, Accepted 14/04/2026

Abstract

Apple is one of the most economically valuable fruit crops cultivated worldwide; however, its productivity is significantly affected by leaf diseases such as apple scab and cedar apple rust. Conventional disease identification methods are often labor-intensive, time-consuming, and highly dependent on human expertise, resulting in subjective diagnoses. To address these limitations, this study proposes an automated apple leaf disease classification system based on a Convolutional Neural Network (CNN). A subset of the PlantVillage dataset comprising 1,200 images was utilized, including 400 healthy apple leaves, 400 apple scab images, and 400 cedar apple rust images. The dataset was divided into 80% training and 20% testing subsets. Image preprocessing consisted of resizing, pixel normalization, brightness adjustment for contrast enhancement, and data augmentation to improve model generalization. The CNN model was trained using the Adam optimizer with the categorical cross-entropy loss function. Experimental results demonstrated an overall classification accuracy of 98% without evidence of significant overfitting. The classification report showed that the healthy class achieved 0.97 precision, 0.97 recall, and 0.97 F1-score; the cedar apple rust class achieved 0.99 precision, 1.00 recall, and 0.99 F1-score; while the apple scab class obtained 0.97 precision, 0.96 recall, and 0.97 F1-score. Most classification errors occurred in the apple scab class due to its visual similarity to healthy leaves. These findings demonstrate that the proposed CNN architecture effectively learns discriminative visual features for multiclass apple leaf disease classification. The proposed system has strong potential as an early disease detection tool to support precision agriculture and sustainable apple crop management. Future work will focus on employing deeper CNN architectures, transfer learning techniques, and image segmentation methods to further improve the recognition performance of challenging disease classes, particularly apple scab.

Keywords: apple leaf disease classification, convolutional neural network, deep learning, PlantVillage dataset, image preprocessing, computer vision

1. Introduction

Apple (Malus domestica) is one of the world's most economically important fruit crops, contributing significantly to global agricultural production and food security. However, apple productivity is frequently threatened by various foliar diseases that reduce both fruit quality and yield. Among the most destructive diseases are apple scab, caused by Venturia inaequalis, and cedar apple rust, caused by Gymnosporangium juniperi-virginianae. These fungal infections not only diminish the aesthetic quality and market value of apples but also lead to severe defoliation, reduced photosynthetic activity, and substantial yield losses, particularly in susceptible cultivars. Consequently, early and accurate disease identification is essential for implementing timely disease management strategies and minimizing economic losses. Despite its importance, disease diagnosis in commercial apple orchards is still largely dependent on manual field inspection performed by agricultural experts. This conventional approach is labor-intensive, time-consuming, and highly dependent on individual expertise, making it impractical for large-scale plantations. Furthermore, the limited availability of plant pathologists and agricultural consultants often delays disease diagnosis, reducing the effectiveness of control measures [1]. These limitations have motivated the development of automated disease recognition systems capable of providing rapid, objective, and reliable diagnoses.

Recent advances in computer vision and deep learning have significantly improved the automatic detection and classification of plant diseases from digital images. Among deep learning techniques, Convolutional Neural Networks (CNNs) have demonstrated outstanding capability in automatically extracting hierarchical visual features without requiring handcrafted feature engineering. Their ability to learn discriminative representations from raw image data has made CNNs the dominant approach for plant disease recognition. Khan et al. (2022) reported that CNN-based models have become one of the most widely adopted deep learning techniques for plant disease identification because of their high classification accuracy and robustness across various agricultural applications [2]. Likewise, Bhatt et al. (2021) proposed a transfer learning framework based on the GoogLeNet Inception architecture with Rainbow Concatenation, achieving promising performance on a dataset containing more than 26,000 apple leaf images representing multiple disease categories [3]. In addition, Hughes and Salathé (2015) introduced the PlantVillage dataset, which has become the most widely used open-access benchmark for plant disease classification by providing over 50,000 annotated images of healthy and diseased plant leaves collected under controlled conditions [4].

Numerous studies have investigated different CNN architectures to improve plant disease classification performance while maintaining computational efficiency. Ferentinos (2018) compared several deep CNN models, including AlexNet, VGG, and GoogLeNet, on a publicly available dataset comprising 58 crop–disease categories and reported that the VGG architecture achieved the highest classification accuracy of 99.53% [5]. More recently, Rizwan et al. (2024) highlighted that many existing approaches achieve excellent predictive performance at the expense of high computational complexity, indicating a trade-off between classification accuracy and computational efficiency [6]. Similarly, Too et al. (2019) demonstrated that DenseNet consistently improved classification accuracy with increasing training epochs while requiring fewer trainable parameters and maintaining computational efficiency comparable to state-of-the-art CNN architectures [7]. Although these studies have reported impressive results, further investigation is still required to develop CNN-based models specifically optimized for apple leaf disease classification while achieving a balanced trade-off between predictive performance and computational cost.

Motivated by these challenges, this study proposes a CNN-based automatic classification system for apple leaf diseases using the PlantVillage dataset. The study utilizes the apple subset of PlantVillage, consisting of 3,171 leaf images collected under controlled environmental conditions [8]. The dataset includes four classes: Healthy, Apple Scab, Black Rot, and Cedar Apple Rust. The proposed framework consists of image preprocessing, including normalization and data augmentation to improve model generalization, followed by CNN-based feature extraction and classification. The performance of the proposed model is evaluated using accuracy, precision, recall, and F1-score. The proposed system is expected to provide a reliable and efficient tool for early apple leaf disease detection while supporting intelligent crop monitoring and sustainable precision agriculture.

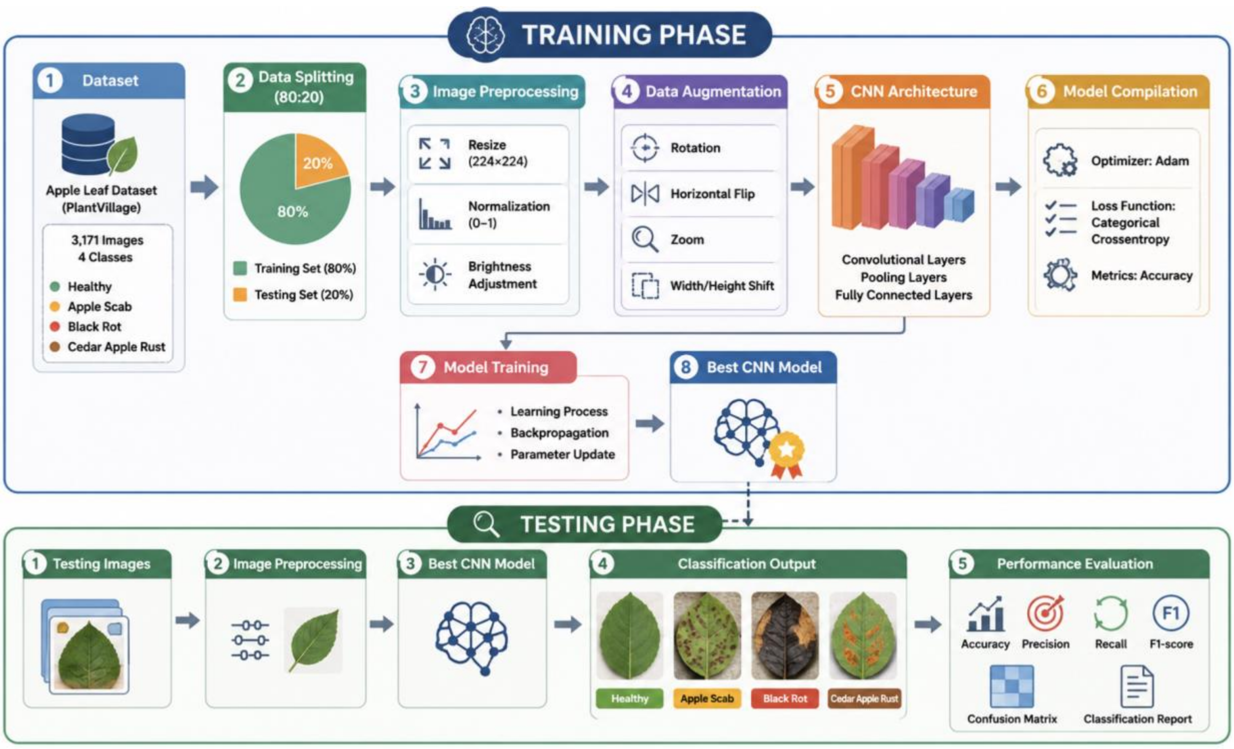


Figure 1. Research framework for the proposed CNN-based apple leaf disease classification method

2. Methods

This study employs a deep learning approach based on a Convolutional Neural Network (CNN) to automatically classify apple leaf diseases from digital images. The proposed methodology consists of four main stages: dataset acquisition and preparation, image preprocessing, CNN model development, and performance evaluation. The dataset is first prepared and partitioned into training and testing subsets before undergoing preprocessing, including image resizing, normalization, and data augmentation to improve model generalization. Subsequently, a CNN model is designed and trained to learn discriminative visual features associated with different apple leaf disease categories. Finally, the trained model is evaluated using standard classification metrics to assess its effectiveness and robustness. The model development process is divided into two primary phases: training and testing. During the training phase, the CNN learns representative feature patterns from the training dataset by optimizing network parameters through backpropagation. In the testing phase, the trained model is evaluated on previously unseen images to measure its classification performance and generalization capability. The overall workflow of the proposed research methodology is illustrated in Fig. 1.

2.1. Dataset

This study utilized a subset of the PlantVillage dataset specifically containing apple leaf images. The dataset was obtained from the publicly available PlantVillage repository, which provides high-quality images of healthy and diseased plant leaves captured under controlled laboratory conditions with a uniform background [9]. Such controlled acquisition conditions ensure consistent image quality while minimizing variations caused by environmental factors, such as illumination, shadows, and complex backgrounds, thereby facilitating the development and evaluation of computer vision models. The dataset consists of 1,200 apple leaf images evenly distributed across three classes to prevent class imbalance during model training. Each class contains 400 images, representing Healthy Apple Leaves (Healthy), Apple Scab, and Cedar Apple Rust. The Healthy class includes leaves without visible disease symptoms, whereas the Apple Scab and Cedar Apple Rust classes represent leaves infected by Venturia inaequalis and Gymnosporangium juniperi-virginianae, respectively. Maintaining an equal number of samples in each category helps reduce classification bias and enables a more objective evaluation of the proposed model.

Table 1. Representative samples of healthy and diseased apple leaf images from the PlantVillage dataset

No	Images	Class
1		Healthy
2		Scab
3		Cedar Apple Rust

Prior to model development, the dataset was randomly divided into training and testing subsets using an 80:20 ratio. Consequently, 960 images were allocated for model training, while the remaining 240 images were reserved for performance evaluation. Random stratified sampling was applied to preserve an equal class distribution in both subsets, ensuring a fair and reliable assessment of the model's generalization capability. All images are stored in JPEG (JPG) format with a uniform spatial resolution of 256 × 256 pixels [10]. The consistent image dimensions simplify the preprocessing stage and eliminate the need for extensive image normalization before feature extraction. Representative examples of apple leaf images from each disease category, along with their corresponding visual characteristics, are presented in Table 1.

2.2. Image Preprocessing

Image preprocessing is a crucial stage in developing a robust deep learning model, as it standardizes the input data and improves image quality before feature extraction by the CNN. Proper preprocessing reduces unwanted variations in the dataset while enhancing the discriminative visual characteristics of the target objects, thereby improving model convergence and classification performance [11]. In this study, the preprocessing pipeline consists of four sequential stages: image resizing, pixel normalization, brightness adjustment, and data augmentation.

2.2.1. Image Resizing

The first preprocessing step involves resizing all images to a fixed resolution of 224 × 224 pixels, ensuring compatibility with the input requirements of the proposed CNN architecture. Standardizing image dimensions is essential because CNN models require uniform input sizes for efficient mini-batch processing and optimized memory utilization during both training and inference. Moreover, an input resolution of 224 × 224 pixels has become the de facto standard for many deep learning architectures, providing a practical balance between computational efficiency and the preservation of important visual information [12]. The resizing process was performed using bilinear interpolation while maintaining the original aspect ratio as much as possible. This operation reduces computational complexity without significantly affecting the visual characteristics required for disease recognition, such as lesion shape, color distribution, and texture patterns. Examples of the original and resized apple leaf images are presented in Fig. 2(a) and Fig. 2(b), respectively.

(a) Before Resizing (b) After Resizing

Figure 2. Image resizing process: (a) original apple leaf image and (b) resized image (224 × 224 pixels)

2.2.2. Image Rescaling (Normalization)

Following the resizing stage, all input images were normalized by rescaling their pixel intensity values from the original range of **[0, 255]** to **[0, 1]** through division by **255**. Pixel normalization is a standard preprocessing technique in deep learning that ensures numerical consistency across input images and facilitates more stable network optimization during training [13]. Normalizing pixel values reduces the magnitude of the input data, enabling faster convergence of the optimization algorithm and improving gradient stability during backpropagation. Without normalization, large pixel values may lead to unstable gradient updates, resulting in slower convergence and reduced learning efficiency. By transforming the input data into a smaller and more uniform numerical range, the model can learn representative visual features more effectively while improving overall training stability. Mathematically, the normalization process is expressed as:

$$I_{\text{normalized}} = \frac{I_{\text{original}}}{255}$$

where I_{original} denotes the original pixel intensity and $I_{\text{normalized}}$ represents the normalized pixel value within the range of 0 to 1. Examples of the apple leaf images before and after the normalization process are illustrated in Fig. 3(a) and Fig. 3(b), respectively.

(a) Before Rescaling (b) After Rescaling

Figure 3. Image normalization process: (a) original image with pixel values in the range of 0–255 and (b) normalized image with pixel values rescaled to the range of 0–1.

2.2.3. Contrast Adjustment

To improve the visual quality of the input images, a contrast adjustment technique was applied after the normalization stage. This preprocessing step enhances the visibility of important leaf structures while reducing the influence of illumination variations introduced during image acquisition. Improving image contrast enables disease symptoms to become more distinguishable, thereby facilitating the extraction of discriminative visual features by the CNN [14]. Contrast enhancement emphasizes subtle differences in color intensity, lesion boundaries, and surface texture, which are essential characteristics for distinguishing healthy leaves from diseased ones. In particular, symptoms associated with apple scab and cedar apple rust, such as discoloration, necrotic spots, and rust-colored lesions, become more prominent after contrast enhancement. Consequently, the CNN can learn more representative feature patterns, leading to improved classification performance and model robustness. Examples of apple leaf images before and after the contrast adjustment process are presented in Fig. 4(a) and Fig. 4(b), respectively.

(a) Original Image (b) Contrast-enhanced Image

Gambar 4 . Contrast adjustment process: (a) original apple leaf image and (b) contrast-enhanced image

2.2.4. Image Augmentation

To improve the diversity of the training dataset and enhance the generalization capability of the proposed CNN model, image augmentation was applied exclusively to the training images. Data augmentation is a widely adopted strategy in deep learning to artificially increase the number of training samples by generating transformed versions of existing images. This approach effectively reduces the risk of overfitting and enables the model to learn more robust and invariant visual representations [15]. In this study, six augmentation techniques were employed to simulate variations commonly encountered during real-world image acquisition, including differences in camera orientation, object position, scale, and illumination. The applied transformations consist of:

- Vertical Flip: Reflects the image along the vertical axis to generate inverted leaf orientations.
- Rotation (+30°): Rotates the image by 30° clockwise, introducing viewpoint variations.
- Rotation (−30°): Rotates the image by 30° counterclockwise to improve rotational invariance.
- Zoom: Randomly enlarges image regions to simulate variations in camera distance.
- Shift: Translates the image horizontally and vertically, producing different object positions within the image frame.
- Brightness Adjustment: Modifies image brightness to simulate different illumination conditions during image acquisition.

These transformations increase the variability of the training data without altering the semantic characteristics of the disease classes. Consequently, the CNN becomes more robust to changes in orientation, scale, position, and lighting, allowing it to learn more discriminative visual features and improving its ability to generalize to previously unseen images. Previous studies have shown that combining augmentation techniques, such as flipping, rotation, translation, scaling, and brightness adjustment, can significantly improve classification performance and reduce overfitting in plant disease recognition tasks [15], [16]. Representative examples of the augmented apple leaf images generated using the applied transformations are presented in Fig. 5.

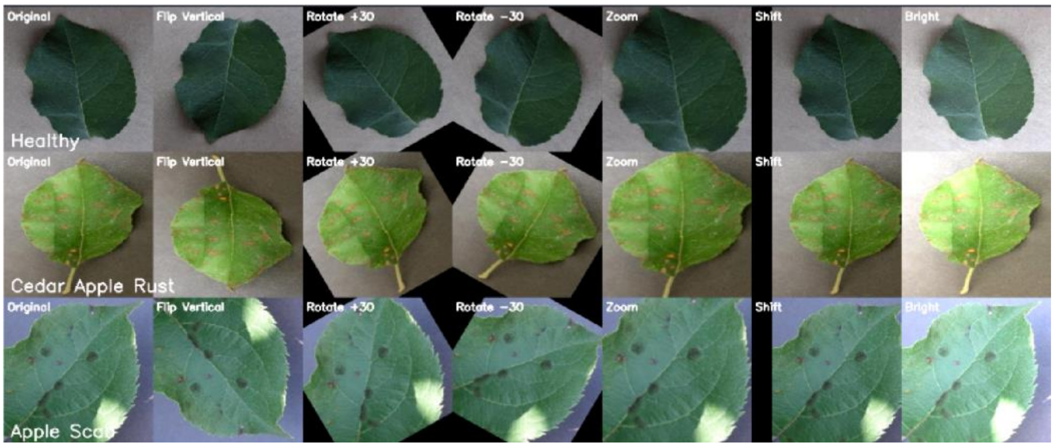


Figure 5. Representative examples of image augmentation techniques applied to the apple leaf training dataset

2.3. Convolutional Neural Network (CNN)

Following the image preprocessing stage, a Convolutional Neural Network (CNN) was employed to automatically classify apple leaf images into their corresponding disease categories.

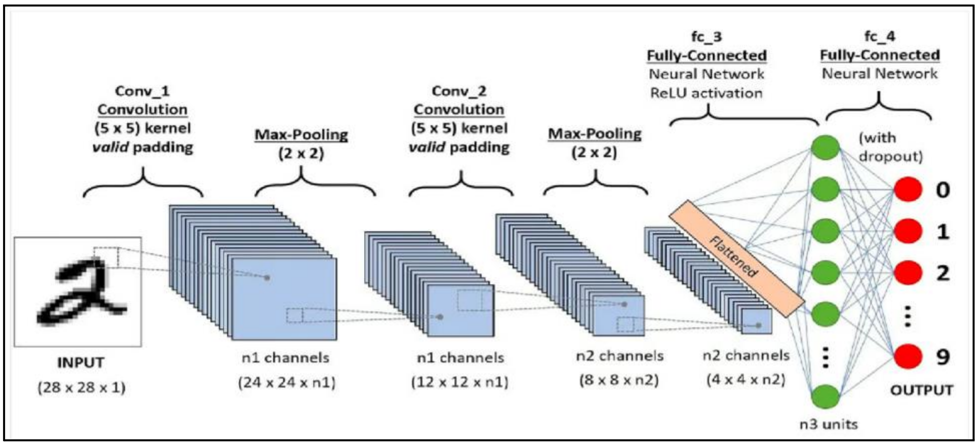


Figure 6. Architecture of the proposed Convolutional Neural Network (CNN)

CNN is a deep learning architecture specifically designed for image analysis, owing to its capability to learn hierarchical feature representations directly from raw image data. Unlike conventional machine learning methods that rely on manually engineered features, CNN performs end-to-end feature learning, enabling the automatic extraction of discriminative visual characteristics from the input images [17]. The network progressively learns both low-level and high-level visual features through multiple convolutional layers. Early layers capture fundamental image characteristics, such as edges, color gradients, and simple textures, whereas deeper layers learn more complex and disease-specific patterns, including lesion morphology, discoloration, and irregular surface textures. These hierarchical representations enable the model to effectively distinguish between Healthy, Apple Scab, and Cedar Apple Rust leaf images.

The proposed CNN architecture consists of several convolutional blocks for feature extraction, followed by pooling layers for spatial dimensionality reduction while preserving essential visual information. The extracted feature maps are subsequently transformed into one-dimensional feature vectors and passed through fully connected layers to perform the final classification. A Softmax activation function is employed in the output layer to estimate the probability distribution across the three disease classes, allowing the model to assign each input image to the class with the highest predicted probability. The overall architecture of the proposed CNN model is illustrated in Fig. 6.

2.3.1. CNN Architecture

The proposed CNN architecture consists of several sequential layers designed to perform automatic feature extraction and image classification. Each component plays a specific role in transforming raw image data into discriminative feature representations for disease recognition.

- Convolutional Layer
  The convolutional layer is the core component of the CNN architecture, responsible for extracting local visual features from the input images through convolution operations. During this process, multiple learnable filters (kernels) slide across the image to detect fundamental visual patterns, including edges, textures, and lesion boundaries. Multiple convolutional layers with progressively increasing numbers of filters are employed to enable hierarchical feature learning, where shallow layers capture simple visual characteristics and deeper layers learn more complex disease-specific representations.
- Pooling Layer
  Pooling layers are incorporated after the convolutional layers to reduce the spatial dimensions of the extracted feature maps while preserving the most informative features. This dimensionality reduction decreases computational complexity, minimizes the number of trainable parameters, and helps alleviate overfitting. In this study, Max Pooling is adopted because it effectively retains the most prominent features while providing robustness against small spatial translations.
- Activation Function
  Each convolutional layer employs the Rectified Linear Unit (ReLU) activation function to introduce nonlinearity into the network. ReLU is defined as

  $$f(x)=\max(0,x)$$

  where negative values are mapped to zero while positive values remain unchanged. This activation function accelerates network convergence, alleviates the vanishing gradient problem, and improves the learning capability of deep neural networks.
- Fully Connected Layer
  After feature extraction, the resulting feature maps are transformed into a one-dimensional feature vector through a Flatten operation before being passed to one or more Fully Connected (Dense) layers. These layers integrate the learned features and model the complex relationships among them, enabling the network to generate highly discriminative representations for classification.
- Output Layer
  The final layer employs the Softmax activation function to produce a probability distribution over the three disease classes. Softmax converts the network outputs into normalized probabilities whose sum equals one, allowing the input image to be assigned to the class with the highest predicted probability [18].

2.3.2. Model Compilation

After constructing the CNN architecture, the model was compiled prior to training. The Adam optimizer was selected because of its adaptive learning rate mechanism, which combines the advantages of AdaGrad and RMSProp, resulting in faster and more stable convergence during optimization. The network was optimized using the categorical cross-entropy loss function, which is widely used for multi-class classification problems. This loss function quantifies the discrepancy between the predicted probability distribution and the corresponding ground-truth labels encoded using one-hot representation. It is mathematically expressed as :

$$L = -\sum_{i=1}^{C} y_i \log(\hat{y}_i)$$

where y_i represents the true class label and $\hat{y}_i$ denotes the predicted probability for class i. Model performance during training was monitored using classification accuracy, which measures the proportion of correctly classified samples throughout the optimization process [19].

2.3.3. Model Training

The CNN model was trained using a batch size of 32 for 50 epochs, allowing sufficient iterations for the network to learn representative visual patterns from the apple leaf images. During each training iteration, input images were propagated through the network via forward propagation to generate prediction probabilities. The prediction errors were subsequently propagated backward using the backpropagation algorithm to update the network weights through gradient-based optimization. Throughout the training process, both training accuracy and training loss were monitored at every epoch to evaluate the learning progress and convergence behavior of the proposed model. These learning curves provide valuable insights into model stability and help identify potential overfitting or underfitting. After completing the training phase, the optimized CNN model had learned discriminative features related to leaf color, texture, lesion patterns, and morphological characteristics, enabling accurate classification of healthy and diseased apple leaves. The trained model was subsequently evaluated using an independent testing dataset to assess its generalization capability on previously unseen images [20].

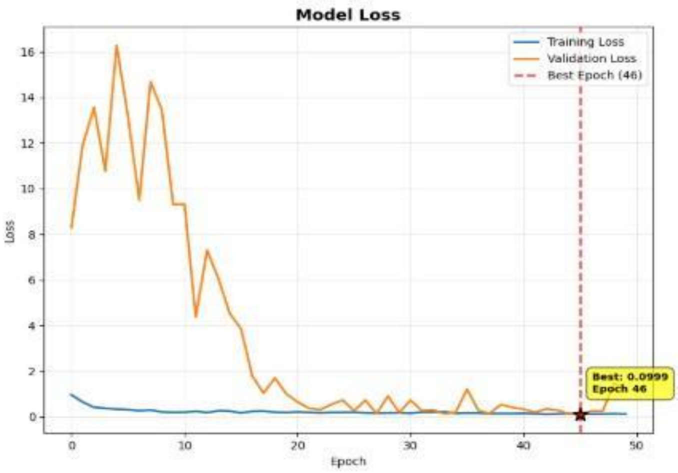
2.3.4. Model Evaluation

The performance of the proposed CNN model was evaluated using an independent testing dataset that was not involved during the training phase. This evaluation aimed to assess the model's ability to generalize to unseen apple leaf images and accurately classify different disease categories. Four widely used classification metrics were employed, namely accuracy, precision, recall, and F1-score. Accuracy measures the overall proportion of correctly classified samples, while precision quantifies the reliability of positive predictions for each disease class. Recall evaluates the model's capability to correctly identify all samples belonging to a particular class, whereas the F1-score provides a balanced assessment by combining precision and recall into a single metric. In addition to these quantitative metrics, the model was further analyzed using a confusion matrix, which provides detailed information regarding correctly and incorrectly classified samples for each disease category. The confusion matrix facilitates the identification of class-specific misclassification patterns and reveals which disease classes exhibit similar visual characteristics. Such analysis is particularly useful for understanding the limitations of the proposed model and identifying opportunities for future improvements. Overall, the combination of classification metrics and confusion matrix analysis provides a comprehensive evaluation of the CNN model. This evaluation protocol has been widely adopted in image-based plant disease recognition studies and offers a reliable assessment of model effectiveness, robustness, and generalization performance.

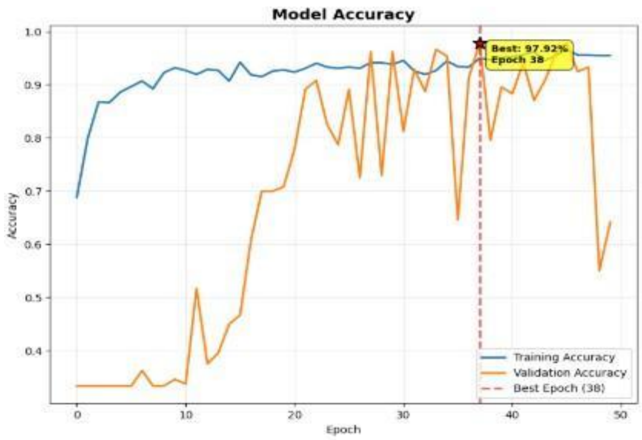
3. Result and Discussions

3.1. Performance Evaluation on the PlantVillage Dataset

The proposed Convolutional Neural Network (CNN) model was evaluated using the apple leaf subset of the PlantVillage dataset. Prior to training, the dataset was randomly divided into training (80%) and testing (20%) subsets, while a portion of the training data was further reserved for validation to monitor the learning process and prevent overfitting. The model was trained for 50 epochs using the Adam optimizer and the categorical cross-entropy loss function. The learning performance of the CNN was assessed by monitoring the training and validation accuracy, as well as the training and validation loss throughout the training process. These learning curves provide valuable insights into the convergence behavior, learning stability, and generalization capability of the proposed model. As illustrated in Fig. 7, the CNN exhibited a steady increase in both training and validation accuracy accompanied by a gradual decrease in the corresponding loss values, indicating that the network successfully learned discriminative visual features from the apple leaf images. The relatively small gap between the training and validation curves further suggests that the proposed model generalized well to unseen data without exhibiting significant overfitting.



(a) Training dan Validation Loss



(b) Training dan Validation Accuracy

Figure 7. Loss dan Accuracy on Dataset Plantvillage

Based on Fig. 7(a), the training loss exhibits a consistent downward trend throughout the 50 training epochs, indicating that the proposed CNN progressively learned meaningful feature representations from the apple leaf images. The validation loss also decreased steadily and reached its minimum value of 0.10 at epoch 46. Although minor fluctuations were observed in the validation loss, particularly between epochs 20 and 30, these variations remained relatively small and did not indicate unstable learning behavior. Overall, the convergence of both training and validation loss curves demonstrates that the model achieved stable optimization without showing evidence of significant overfitting. The accuracy curves presented in Fig. 7(b) further confirm the effectiveness of the proposed model. The training accuracy increased rapidly during the initial learning stage and gradually stabilized at approximately 1.00 after epoch 20, indicating that the network successfully captured discriminative visual features from the training data. A similar trend was observed for the validation accuracy, although with slightly larger fluctuations that are commonly encountered during deep learning optimization. The highest validation accuracy of 98% was achieved at epoch 38, demonstrating the strong generalization capability of the proposed CNN on unseen validation samples. To further evaluate the classification performance, a confusion matrix was generated using the independent testing dataset. The confusion matrix provides detailed information regarding the prediction results for the three disease categories: Healthy, Apple Scab, and Cedar Apple Rust. As illustrated in Fig. 8, the proposed CNN achieved excellent classification performance across all classes. For the Healthy class, the model correctly classified 78 of 80 testing images, while only two samples were misclassified as Apple Scab. In the Apple Scab class, 77 of 80 images were correctly recognized, with two samples incorrectly classified as Healthy and one sample misclassified as Cedar Apple Rust. The Cedar Apple Rust class achieved perfect classification performance, with all 80 testing images correctly identified and no misclassification observed.

The dominance of values along the principal diagonal of the confusion matrix indicates that the proposed CNN effectively discriminates among the three apple leaf categories. The limited number of misclassified samples suggests that most prediction errors occurred between visually similar classes, particularly Healthy and Apple Scab, whose early disease symptoms often exhibit subtle color and texture differences. Overall, these results demonstrate the robustness and high classification capability of the proposed CNN for automatic apple leaf disease recognition.

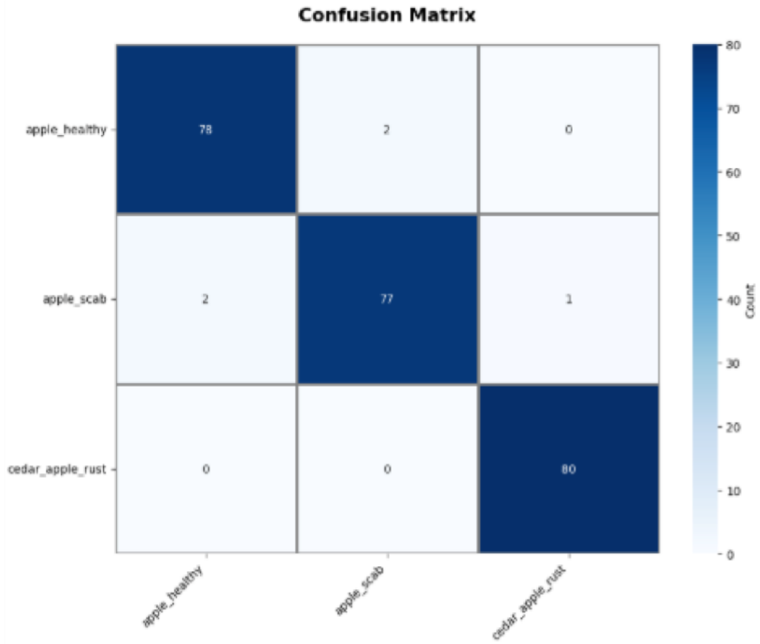


Figure 8. Confusion matrix for the classification of Healthy, Apple Scab, and Cedar Apple Rust leaf images

Based on the evaluation performed on the 240 testing images, the proposed CNN model achieved an overall classification accuracy of 98%, demonstrating its effectiveness in distinguishing healthy and diseased apple leaves. The model performance was further assessed using precision, recall, and F1-score for each disease category to provide a comprehensive evaluation of its classification capability. As summarized in Table 2, the Healthy class achieved both a precision and recall of 0.97, indicating consistent performance in correctly identifying healthy apple leaves while maintaining a low false-positive rate. For the Apple Scab class, the model obtained a precision of 0.97 and a recall of 0.96, suggesting that the majority of infected leaves were successfully recognized, although a small number of samples were misclassified due to visual similarities with other classes.

Tabel 2. Clasification Report

Class	Precision	Recall	F1-Score	Support
Leaf Healthy	0.97	0.97	0.97	88
Leaf Apple Scab	0.97	0.96	0.97	88
Leaf Cedar Apple Rust	0.99	1.00	0.99	88
Accuracy			**0.98**	**240**
Macro avg	0.98	0.98	0.98	240
Weighted avg	0.98	0.98	0.98	240

The Cedar Apple Rust class exhibited the highest classification performance, achieving a precision of 0.99, a recall of 1.00, and an F1-score of 0.99. These results indicate that the proposed CNN was able to identify cedar apple rust symptoms with exceptional reliability and without any false-negative predictions in the testing dataset. Overall, the proposed model achieved an accuracy of 0.98, while the macro-averaged and weighted-averaged values for precision, recall, and F1-score were all approximately 0.97. The close agreement between these aggregate metrics indicates that the classifier maintained balanced performance across all disease categories without exhibiting significant bias toward any particular class. These findings demonstrate the robustness and strong generalization capability of the proposed CNN for automatic apple leaf disease classification. The complete quantitative evaluation is presented in Table 2 in the form of a classification report.

4. Conclusion

This study successfully developed a Convolutional Neural Network (CNN)-based model for the automatic classification of apple leaf diseases using the PlantVillage dataset. The experimental results demonstrate that the proposed model achieved an overall testing accuracy of 98%, indicating its strong capability to distinguish healthy and diseased apple leaves without exhibiting significant overfitting. The training and validation curves further confirmed stable convergence and good generalization throughout the learning process. According to the classification report, the Healthy class achieved a precision, recall, and F1-score of 0.97, reflecting consistent classification performance. The Cedar Apple Rust class exhibited the highest performance, with a precision of 0.99, recall of 1.00, and F1-score of 0.99, demonstrating the model's excellent ability to recognize this disease. Meanwhile, the Apple Scab class achieved a precision of 0.97, recall of 0.96, and F1-score of 0.97. The slightly lower recall for this class suggests that a small number of samples were misclassified, primarily due to visual similarities between early scab symptoms and healthy leaf tissues. The experimental findings also indicate that the proposed preprocessing pipeline, including image resizing, normalization, contrast enhancement, and data augmentation, contributed substantially to improving the model's robustness and generalization capability. Overall, the proposed CNN-based framework demonstrates strong potential as an automated tool for the early detection of apple leaf diseases, supporting more efficient and reliable crop monitoring in precision agriculture.

REFERENCES

[1] R. Thapa, “The Plant Pathology Challenge 2020 data set to classify foliar disease of apples,” vol. 8, no. 9, pp. 1–8, 2020, doi: 10.1002/aps3.11390.

[2] A. I. Khan, S. M. K. Quadri, S. Banday, and J. Latief Shah, “Deep diagnosis: A real-time apple leaf disease detection system based on deep learning,” Comput. Electron. Agric., vol. 198, p. 117093, 2022, doi: https://doi.org/10.1016/j.compag.2022.117093.

[3] P. Bansal, R. Kumar, and S. Kumar, “Disease Detection in Apple Leaves Using Deep Convolutional Neural Network,” 2021, doi: 10.3390/agriculture11070617.

[4] D. P. Hughes and M. Salathé, “An open access repository of images on plant health to enable the development of mobile disease diagnostics,” arXiv preprint arXiv:1511.08060, 2015.

[5] K. P. Ferentinos, “Deep learning models for plant disease detection and diagnosis,” Comput. Electron. Agric., vol. 145, pp. 311–318, 2018, doi: https://doi.org/10.1016/j.compag.2018.01.009.

[6] M. Rizwan, S. Bibi, S. Ul, T. Jan, and M. Haseeb, “Automatic plant disease detection using computationally efficient convolutional neural network,” no. February, pp. 1–14, 2024, doi: 10.1002/eng2.12944.

[7] E. Chebet, L. Yujian, S. Njuki, and L. Yingchun, “A comparative study of fine-tuning deep learning models for plant disease identification,” in Proc. Int. Conf. on Intelligent Agriculture (ICIA), Mar. 2018.

[8] A. Islam, T. Tahsin, Z. Anjum, M. B. Hasan, and M. H. Kabir, “A domain-adapted lightweight ensemble for resource-efficient few-shot plant disease classification,” arXiv preprint, 2023.

[9] S. P. Mohanty, D. P. Hughes, and M. Salathé, “Using Deep Learning for Image-Based Plant Disease Detection,” vol. 7, no. September, pp. 1–10, 2016, doi: 10.3389/fpls.2016.01419.

[10] E. Aronés, J. Espinal, and C. Salas, “App2 : software solution for apple leaf disease detection based on deep learning (CNN + SVM),” no. October, pp. 1–12, 2025, doi: 10.3389/frai.2025.1648867.

[11] H. J. D. Koresh, “Impact of the Preprocessing Steps in Deep Learning-Based Image Classifications,” Natl. Acad. Sci. Lett., vol. 47, no. 6, pp. 645–647, 2024, doi: 10.1007/s40009-023-01372-2.

[12] H. Talebi and P. Milanfar, “Learning to resize images for computer vision tasks,” in Proc. IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR), pp. 497–506, 2021.

[13] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” Nature, vol. 521, no. 7553, pp. 436–444, 2015, doi: 10.1038/nature14539.

[14] K. Hong and Y. Zhou, “The pipelines of deep learning-based plant image processing,” Computers and Electronics in Agriculture, vol. 192, p. 106607, Jan. 2022. doi: 10.1016/j.compag.2021.106607.

[15] M. K. Elfouly, A. M. Abdelaziz, W. H. Gomaa, and M. Abdalla, A Deep Learning-Based Framework for Large-Scale Plant Disease Detection Using Big Data Analytics in Precision Agriculture. Cham, Switzerland: Springer International Publishing, 2025.

[16] X. Liu, F. Guo, X. Lu, W. Peng, and A. Jolfaei, “Data augmentation in classification and segmentation: A survey and new strategies,” Journal of Imaging, vol. 9, no. 2, p. 46, Feb. 2023. doi: 10.3390/jimaging9020046.

[17] A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, “A survey of the recent architectures of deep convolutional neural networks,” Artif. Intell. Rev., vol. 53, no. 8, pp. 5455–5516, 2020, doi: 10.1007/s10462-020-09825-6.

[18] A. Penelitian, A. F. Fuady, D. O. Amsyah, M. Farhan, R. Riansyah, and M. D. Dhiyaul, “Jurnal Publikasi Ilmu

Komputer dan Implementasi Algoritma Convolutional Neural Network (CNN) untuk Pengenalan dan Klasifikasi Buah Berdasarkan Citra Digital".

[19] K. E. N. Ratri, R. Wardani, and L. Leonardi, "Klasifikasi Penyakit pada Daun Anggur menggunakan Metode Convolutional Neural Network," vol. 17, no. 2, pp. 112–126, 2023.

[20] K. Azmi and S. Defit, "Implementasi Convolutional Neural Network (CNN) Untuk Klasifikasi Batik Tanah Liat Sumatera Barat," vol. 16, no. 1, pp. 2580–2582, 2023.

[21] J. Gu et al., "Recent advances in convolutional neural networks," Pattern Recognit., vol. 77, pp. 354–377, 2018, doi: https://doi.org/10.1016/j.patcog.2017.10.013.