Current Applied Science and Technology e-ISSN: 2586-9396

Curr. Appl. Sci. Technol. ….., Vol... (No...), e0266103

Research article

Diagnosis of Mango Leaf Diseases Using Deep Learning Techniques

Chonnakarn Wongnim[1], Benyapha Saardmuang[1] and Sanya Khruahong[2*]

[1]*Naresuan University Secondary Demonstration School, Thailand*
[2]*Naresuan University, Faculty of Science, Thailand*

Received: 27 January 2025, Revised: 29 October 2025, Accepted: 2 November 2025, Published: 5 February 2026

Abstract

Mango cultivation is a cornerstone of Thailand's agriculture and economy, but diseases such as anthracnose, algal leaf spot, and gall midge present significant challenges as they can reduce crop yield and quality. In this study, we developed a machine learning-based system to diagnose mango leaf diseases using a dataset of 1,900 images collected from mango orchards in Phitsanulok province. The data underwent preprocessing and augmentation to optimize model training. Five deep learning models—Convolutional Neural Network (CNN), VGG16, DenseNet121, ResNet50, and InceptionV3—were trained and evaluated. Among these, ResNet50 demonstrated the best performance, with an accuracy of 99.8%, a precision of 0.998, a recall of 0.998, and an F1-score of 0.998. Leveraging its superior performance, the ResNet50 model was integrated into a mobile application designed for real-time disease diagnosis. This user-friendly application enables mango farmers to upload images of affected leaves and receive instant disease identification and treatment recommendations. The findings highlight the potential of deep learning models in agricultural applications, offering a reliable and efficient tool for early disease detection and management. By enabling timely intervention, this innovation enhances crop health, reduces losses, and boosts productivity, contributing significantly to sustainable farming practices and improving farmers' livelihoods.

Keywords: machine learning; mango diseases; classification; neural network; deep learning

1. Introduction

Mango is a key agricultural product in Thailand, celebrated for its sweet flavor, vibrant color, and versatility in both local cuisine and global markets. As one of the world's top producers and exporters of mangoes, Thailand relies heavily on this crop for both economic value and farmers' livelihoods (Misra, 1992). However, mango cultivation is increasingly threatened by plant diseases—particularly those caused by fungi and pathogens—which significantly reduce yield and quality, thus affecting both domestic consumption and export potential (Ferentinos, 2018).

*Corresponding author: E-mail: sanyak@nu.ac.th
https://doi.org/10.55003/cast.2026.266103

Among the most common mango leaf diseases in Thailand are anthracnose (*Colletotrichum gloeosporioides*), algal leaf spot (*Cephaleuros virescens*), and gall midge (*Procontarinia* spp.). These diseases often present overlapping symptoms, making early and accurate diagnosis difficult (Kumar et al., 2021). Misdiagnosis can result in ineffective treatment, further aggravating crop losses.

To address these challenges, recent advancements in artificial intelligence (AI) and machine learning (ML) have shown strong potential in agricultural applications, particularly in plant disease detection (Saleem et al., 2019; Kattenborn et al., 2021). Convolutional neural networks (CNNs), a subset of deep learning, are particularly effective in image-based classification tasks due to their ability to automatically extract and learn hierarchical visual features (O'Shea, 2015; Ferentinos, 2018).

Several studies have validated the use of CNNs in plant disease classification. Ferentinos (2018) achieved over 99% accuracy across multiple plant diseases using CNNs but emphasized the importance of dataset diversity for generalizability. Kumar et al. (2021) applied CNNs to mango leaf diseases, reaching 96.16% accuracy, yet faced difficulties in distinguishing visually similar symptoms, highlighting the need for advanced data augmentation. Saleem et al. (2019) leveraged transfer learning with ResNet and DenseNet on small datasets, achieving promising results but encountering challenges in real-time applications due to computational demands. Pandian et al. (2022) confirmed CNN effectiveness in fungal disease detection, though their limited dataset restricted the model’s broader applicability. More recently, Rao et al. (2021) demonstrated the power of AlexNet for mango and grape leaf disease detection, reinforcing the feasibility of deep learning deployment in practical tools like mobile applications.

While these studies demonstrate the promise of deep learning, key limitations remain, including the handling of overlapping symptoms, reliance on simple or small datasets, and difficulty in deploying high-accuracy models on mobile platforms. These gaps form the foundation of this study’s contribution.

In this study, we aimed to address the challenges of mango leaf disease diagnosis by applying and comparing multiple deep learning models. Our key contributions included the construction of a high-quality, expert-labeled dataset consisting of 1,900 mango leaf images representing four categories—anthracnose, algal leaf spot, gall midge, and healthy leaves—captured under controlled conditions to ensure consistency. We performed a comprehensive evaluation of five deep learning architectures (CNN, VGG16, DenseNet121, ResNet50, and InceptionV3) using a standardized training framework with advanced preprocessing and augmentation techniques to ensure fair and robust comparisons. The results identified ResNet50 as the most accurate and reliable model, achieving 99.8% accuracy across disease classes. To bridge the gap between research and practice, we deployed the ResNet50 model into a real-time Android mobile application designed for ease of use by mango farmers. This tool allows for rapid, on-site disease diagnosis, supporting early intervention and contributing to more efficient and sustainable mango cultivation in Thailand.

2. Materials and Methods

2.1 Convolutional neural network (CNN)

Convolutional neural networks (CNNs) have emerged as a powerful tool for image classification, particularly in agricultural applications, due to their ability to extract hierarchical features from images automatically (Yamashita et al., 2018; Li et al., 2021;

Sujitranan, M., & Anongporn, 2024; Brahmi et al., 2024). Studies such as those by Ferentinos (2018) have demonstrated the effectiveness of CNNs in plant disease detection, achieving accuracy levels exceeding 99% when trained on diverse datasets. CNN architectures are designed to process spatially structured data by applying convolutional layers to identify local patterns, pooling layers to downsample features, and fully connected layers to classify inputs. These characteristics make CNNs particularly effective for identifying diseases with subtle visual distinctions, such as anthracnose, algal leaf spot, and gall midge in mango leaves. However, challenges remain, including the need for large, diverse datasets to ensure robust model performance and generalizability.

Preprocessing and data augmentation techniques play a crucial role in enhancing the performance of CNNs. Kumar et al. (2021) highlighted the importance of these techniques in mango leaf disease classification, achieving a 96.16% accuracy by employing methods such as image resizing and brightness adjustments. Saleem et al. (2019) further emphasized the role of augmentation in addressing data scarcity, demonstrating that generating synthetic variations can improve generalization and reduce overfitting. While these approaches significantly enhance CNN performance, studies such as Pandian et al. (2022) pointed out the limitations of simple datasets, which may restrict the applicability of trained models in real-world scenarios. These findings underscore the importance of employing comprehensive preprocessing pipelines and robust augmentation strategies to optimize CNNs for agricultural applications.

In addition to the previously mentioned models, other notable deep learning architectures have been widely utilized in various applications. VGG16, for instance, is renowned for its straightforward design, comprising 16 layers with small 3 x 3 convolutional filters (Yang et al., 2023; Mukesh et al., 2024). Despite its simplicity, VGG16 has demonstrated robust performance across diverse datasets. However, its depth and fully connected layers result in a high parameter count, increasing computational demands.

ResNet50 tackles the vanishing gradient problem in deep networks through residual learning and skip connections (Theckedath & Sedamkar, 2020; Koonce, 2021; Mascarenhas & Agarwal, 2021; Hossain et al., 2022; Khalifa et al., 2025). This innovative design enables the training of deeper networks without performance degradation, making it highly effective for complex image recognition tasks. Additionally, DenseNet121 features dense connectivity, where each layer integrates outputs from all preceding layers (Rochmawanti &Utaminingrum, 2021; Nandhini & Ashokkumar, 2022; Arulananth et al., 2024; Simangunsong et al., 2024). This approach enhances feature reuse, reduces the overall parameter count, and achieves impressive accuracy with computational efficiency.

Finally, InceptionV3 employs inception modules that utilize parallel convolutional filters of varying sizes to capture multi-scale features effectively (Xia et al., 2017; Chowdary et al., 2020; Ahmed et al., 2023; Huang et al., 2025). This architecture achieves an optimal balance between depth and width, delivering high performance while maintaining computational efficiency. Collectively, these models showcase the versatility and strength of deep learning architectures in addressing diverse image classification challenges.

To evaluate and compare the effectiveness of different deep learning architectures, five models were selected: CNN, VGG16, ResNet50, DenseNet121, and InceptionV3. These models represent a diverse range of architectural styles, from basic custom CNNs to complex, pre-trained deep networks.

1) CNN (Custom model): A simple convolutional architecture designed specifically for mango leaf disease classification, consisting of convolutional layers, max-pooling, and fully connected layers. It serves as a baseline for comparison due to its lightweight design and fast training time.

2) VGG16: Known for its simplicity and uniform architecture, VGG16 consists of 13 convolutional layers with small 3x3 filters followed by 3 fully connected layers. It provides high classification performance but has a high parameter count.
3) ResNet50: Incorporates residual learning with skip connections to address the vanishing gradient problem in deep networks. Its 50-layer architecture allows deep feature learning without degradation, making it ideal for complex tasks.
4) DenseNet121: Features dense connectivity, where each layer receives inputs from all preceding layers. This enhancement improves feature reuse and gradient flow, resulting in enhanced accuracy and efficiency with fewer parameters.
5) InceptionV3: Employs inception modules that combine multiple convolutional filter sizes in parallel, allowing the model to capture features at different scales. It offers a good balance between accuracy and computational efficiency.

These models were selected to investigate how various architectural principles (depth, connectivity, modularity, and parameter size) impact the performance of mango disease classification. Comparing these diverse CNN models provides insights into trade-offs between model complexity, accuracy, and practical deployment for mobile applications. Figure 1 presents an overview of the overall process framework used in this study.

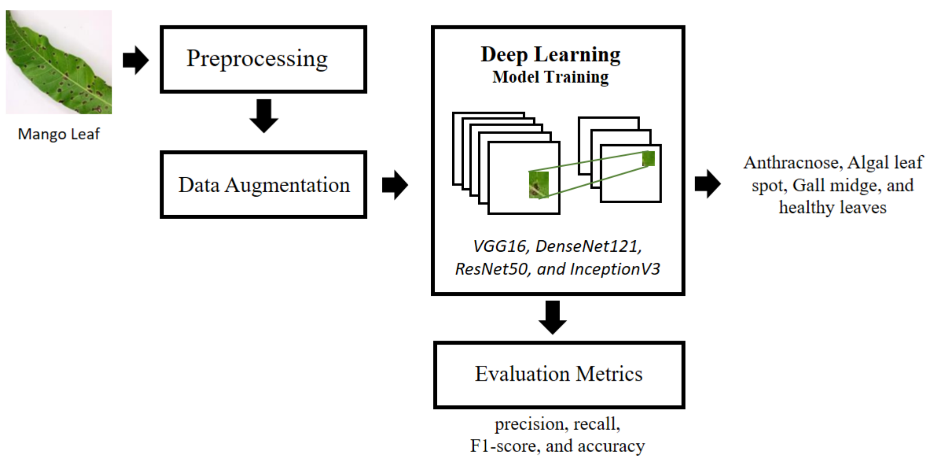

Figure 1. Overview of process

2.2 Data collection

The dataset used in this study was gathered from mango orchards located in the Noen Maprang subdistrict of Phitsanulok province, Thailand. Images were captured for mango leaves exhibiting symptoms of three major diseases—anthracnose (*Colletotrichum gloeosporioides*), algal leaf spot (*Cephaleuros virescens*), and gall midge (*Procontarinia* spp.)—as well as healthy leaves. A total of 1,900 images were collected using an iPad Gen 9 paired with a PhotoBox to maintain consistent lighting conditions and a clean white background, ensuring high-quality image capture as shown in Figures 2 and 3. To ensure accurate classification, each image was carefully labeled by agricultural experts based on the observed disease type. This systematic approach to data collection ensured a reliable and well-structured dataset for model training and evaluation.

(a) Anthracnose (b) Algal leaf spot (c) Gall midge

Figure 2. Three major mango leaf diseases

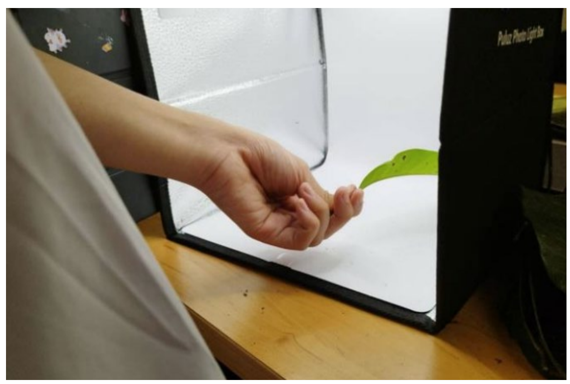

Figure 3. Capturing mango leaf images using a PhotoBox

2.3 Preprocessing

Preprocessing plays a crucial role in preparing the dataset for effective machine-learning model training in mango leaf disease classification. The comprehensive preprocessing pipeline included multiple essential steps: all images were standardized to 224 × 224 pixels to meet deep learning model input requirements and normalized to a 0-1 pixel value range for consistent processing. A thorough quality control process involved manual inspection to eliminate noisy, blurred, or misaligned images, ensuring dataset integrity. To minimize environmental variations and enhance feature visibility, we applied standardization techniques, including histogram equalization and adaptive thresholding, to address lighting inconsistencies and color profile variations. The dataset was further enriched through data augmentation techniques, implementing random rotations (0-360 degrees), horizontal and vertical flips, brightness adjustments (±20%), contrast variations, and random cropping with padding. These augmentation methods artificially expanded the dataset diversity and helped prevent model overfitting. Additionally, color jittering and shear transformations were applied to improve model robustness against real-world variations. The preprocessing workflow was automated through a systematic pipeline with error handling and logging capabilities, ensuring reproducibility and maintaining preprocessing quality across the entire dataset. This comprehensive preprocessing approach significantly enhanced the dataset's quality and contributed to the development of more accurate and reliable machine-learning models for mango leaf disease diagnosis.

2.4 Data augmentation

Data augmentation techniques were employed to prevent overfitting and improve the generalization of the models. Methods such as random rotations, flips, and brightness adjustments were applied to generate synthetic variations of the original images, significantly increasing the dataset size (Shorten & Khoshgoftaar, 2019). This approach created a more diverse and representative training set, closely mimicking real-world scenarios and enhancing the model's ability to handle varied data, as shown in Figure 4.

Figure 4. Data augmentation

In this study, data augmentation was applied dynamically during training using TensorFlow's ImageDataGenerator rather than by creating a static, expanded dataset. Therefore, the total dataset size remained at 1,900 images as shown in Table 1, but the training data was virtually diversified in real-time through random transformations such as rotation, flipping, and brightness adjustment during each training epoch.

Table 1. Class distribution across data splits

Class	Training (70%)	Validation (15%)	Test (15%)	Total
Anthracnose	333	71	71	475
Algal Leaf Spot	333	71	71	475
Gall Midge	333	71	71	475
Healthy	333	71	71	475
Total	**1,332**	**284**	**284**	**1,900**

2.5 Model selection and training

This study evaluated five deep learning models: Convolutional neural network (CNN), VGG16, DenseNet121, ResNet50, and InceptionV3. The models were implemented using TensorFlow and trained on the augmented dataset. The training was conducted over 30 epochs with a batch size of 32 and a learning rate of 0.001, utilizing the Adam optimizer to minimize the loss function. The dataset was divided into three subsets: 70% for training, 15% for validation, and 15% for testing, ensuring a balanced evaluation process (Ferentinos, 2018; Taylor et al., 2018).

Although k-fold cross-validation was considered, this study adopted a stratified 70%-15%-15% train-validation-test split to maintain consistency and comparability across all models. This approach allowed for efficient experimentation while controlling evaluation conditions. Due to the computational demands of training deep models such as ResNet50 and InceptionV3, full cross-validation was not performed, but it remains a promising avenue for future enhancement of model reliability analysis.

2.6 Model architectures

We employed four pretrained CNN backbones—VGG16, ResNet50, DenseNet121, and InceptionV3—together with a custom baseline CNN. To adapt them for our 4-class mango leaf disease dataset (anthracnose, algal leaf spot, gall midge, and healthy), the original fully connected classification layers were replaced with a custom head consisting of a Global Average Pooling layer, a dropout layer (p = 0.3), a dense layer with 256 units and ReLU activation, and a final dense layer with 4 units and Softmax activation. During fine-tuning, convolutional blocks up to the mid-depth were frozen to preserve pretrained features, while deeper layers were unfrozen to allow task-specific adaptation. All convolutional layers used ReLU activations, and classification was performed with argmax over softmax probabilities.

The backbone specifications—including the number of convolutional and pooling layers, filter sizes, activation functions, and the total number of trainable parameters (after head replacement)—are summarized in Table 2. These details, together with the custom head design, enable full reproducibility of our models.

2.7 Application development

Based on the experimental results, the model with the best performance was used to develop a mobile application for the Android platform. This application was designed with a user-friendly interface to ensure ease of use, allowing mango farmers to upload images of leaves and instantly receive disease diagnoses and treatment recommendations.

2.8 Evaluation metrics

The performance of each model was evaluated using the accuracy, precision, recall, and F1-score metrics. Additionally, confusion matrices were used to visualize classification performance and identify misclassifications (O'Shea, 2015; Amiruddin & Kadir, 2020). The evaluation aimed to identify the most effective model for accurately diagnosing mango leaf diseases.

Table 2. CNN backbone specifications and customizations

Backbone	Layers (conv/pool)	Filter Sizes (summary)	Activation	Trainable Params (after 4-class head)	Custom Additions
Baseline CNN	3 conv + 2 FC (~5 total)	3×3 conv, 2×2 max-pool	ReLU	~1.0 M	GAP → Dropout(0.3) → Dense(256, ReLU) → Dense(4, Softmax)
VGG16	13 conv + 3 FC (16 total)	All 3×3 conv, 2×2 max-pool	ReLU	~134.3 M	Same custom head
ResNet50	49 conv + 1 FC (50 total)	7×7 conv + 3×3 bottleneck	ReLU	~23.6 M	Same custom head
DenseNet121	120 conv + 4 pool (121 total)	7×7 conv + 1×1, 3×3 dense blocks	ReLU	~8.0 M	Same custom head
InceptionV3	~48 conv + inception modules	Parallel 1×1, 3×3, 5×5 filters	ReLU	~21.8 M	Same custom head

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (2)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3)$$

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

3. Results and Discussion

3.1 Results

3.1.1 Model performance

Five deep-learning models were trained and evaluated on the mango leaf disease dataset. Tables 3-6 present a comprehensive summary of their performance metrics, including precision, recall, F1-score, and accuracy, highlighting the effectiveness of each model.

Table 3 summarizes the performance metrics of various machine learning models for classifying anthracnose disease. ResNet50 demonstrated the best performance, achieving a precision, recall, and F1-score of 0.998, indicating exceptional accuracy in disease identification. InceptionV3 closely followed with an F1-score of 0.994, reflecting its reliability. DenseNet121 and VGG16 also delivered strong results, while CNN showed considerably lower scores, underscoring its limitations in this task. These findings identify ResNet50 as the most suitable model for practical applications in anthracnose classification.

Table 4 highlights the performance metrics of machine learning models for classifying Gall Midge disease. ResNet50 demonstrated the best performance, achieving the highest precision, recall, and F1-score of 0.998, indicating exceptional accuracy and reliability. InceptionV3 followed closely with an F1-score of 0.994, while VGG16 also performed well with 0.974. DenseNet121 achieved moderate results, and CNN recorded the lowest metrics, particularly for recall (0.632), reflecting its limited effectiveness. These results confirmed ResNet50 as the most suitable model for this classification task.

Table 5 compares the performance of machine learning models in classifying Algal Leaf Spot. InceptionV3 outperformed other models with the highest F1-score of 0.994 and strong precision and recall values, making it the most reliable. VGG16 followed closely with an F1-score of 0.982, showcasing consistent performance. ResNet50 and DenseNet121 also performed well, while CNN demonstrated significantly lower metrics, particularly for recall (0.531) and F1-score (0.601), highlighting its limited effectiveness for this task. These results confirmed InceptionV3's superiority in this classification.

Table 6 presents the performance metrics of machine learning models for classifying healthy leaves. Both ResNet50 and VGG16 achieved perfect scores across all metrics, indicating their exceptional accuracy and reliability. InceptionV3 also performed well with an F1-score of 0.994. DenseNet121 achieved slightly lower precision (0.980) but maintained high recall and F1-score. CNN lagged behind with the lowest precision (0.693) and F1-score (0.781), reflecting its limited effectiveness. These results confirmed ResNet50 and VGG16 as the most effective models for healthy leaves classification.

Table 3. Model performance metrics of anthracnose

Deep Learning Models	Precision	Recall	F1-score
CNN	0.741	0.861	0.791
VGG16	0.982	0.982	0.982
ResNet50	0.998	0.998	0.998
DenseNet121	0.991	0.952	0.972
InceptionV3	0.994	0.994	0.994

Table 4. Model performance metrics of gall midge

Deep Learning Models	Precision	Recall	F1-score
CNN	0.941	0.632	0.754
VGG16	0.974	0.982	0.974
ResNet50	0.998	0.998	0.998
DenseNet121	0.952	0.923	0.942
InceptionV3	0.994	0.994	0.994

Table 5. Model performance metrics of algal leaf spot

Deep Learning Models	Precision	Recall	F1-score
CNN	0.700	0.531	0.601
VGG16	0.981	0.972	0.982
ResNet50	0.952	1.000	0.973
DenseNet121	0.975	0.924	0.972
InceptionV3	0.994	0.994	0.994

Table 6. Model performance metrics of healthy leaves

Deep Learning Models	Precision	Recall	F1-score
CNN	0.693	0.922	0.781
VGG16	1.000	1.000	1.000
ResNet50	1.000	1.000	1.000
DenseNet121	0.980	1.000	0.991
InceptionV3	0.994	0.994	0.994

Table 7 summarizes the average performance metrics across all classes for the evaluated deep learning models. InceptionV3 achieved the highest scores with a near-perfect average precision, recall, and F1-score of 0.994, demonstrating outstanding overall performance. ResNet50 followed closely with exceptional results (0.987/0.999/0.992), particularly notable for its perfect recall across classes. VGG16 also delivered strong and balanced performance at 0.984/0.988/0.985. DenseNet121 performed respectably but ranks lower among the pre-trained models, with an average F1-score of 0.969. In contrast, the custom CNN significantly underperformed, recording the lowest values of 0.769 precision and 0.732 F1-score. These results clearly establish InceptionV3 and ResNet50 as the most effective models for the multi-class image classification task, reaffirming the superiority of transfer learning approaches over conventionally trained architectures.

Table 7. Average performance metrics across all classes

Deep Learning Models	Average Precision	Average Recall	Average F1-score
CNN	0.769	0.737	0.732
VGG16	0.984	0.988	0.985
ResNet50	0.987	0.999	0.992
DenseNet121	0.974	0.950	0.969
InceptionV3	0.994	0.994	0.994

3.1.2 Training and validation performance

To evaluate the effectiveness of the deep learning models, the training and validation performance metrics were analyzed. These metrics provide insights into how well each model learns from the training data and generalizes to unseen validation data. By examining trends in accuracy and loss during the training process, we can assess the models' ability to converge and avoid issues such as overfitting or underfitting. This section highlights the comparative performance of the models, emphasizing their strengths and limitations in diagnosing mango leaf diseases. Figures 5-9 present the training and validation performance curves for each model, highlighting differences in convergence speed and overfitting behavior. These visualizations are included to support evaluation of generalization and training stability, which are critical considerations for real-world deployment.

Figure 5 represents the training and validation performance of the CNN model. The accuracy curve (left) shows rapid growth in early epochs, stabilizing around 98% for training and 97% for validation, indicating effective learning. The loss curve (right) exhibits consistent decreases, with a narrow gap between training and validation loss, suggesting good generalization. However, further optimization could address the slightly lower validation accuracy compared to other models.

Figure 6 displays the training and validation performance of the VGG16 model. Training and validation accuracy (left) rapidly improve and stabilize at approximately 99% and 98%, respectively, demonstrating strong learning and good generalization. The loss curves (right) show consistent decreases, with validation loss slightly higher than training loss, indicating effective optimization with minimal overfitting. These results suggest VGG16 is a highly effective model for the task, with balanced performance on both training and unseen data.

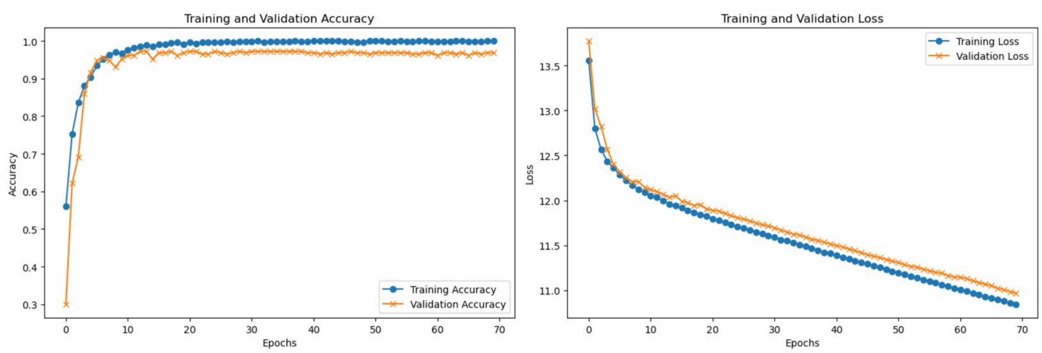

Figure 5. CNN model training

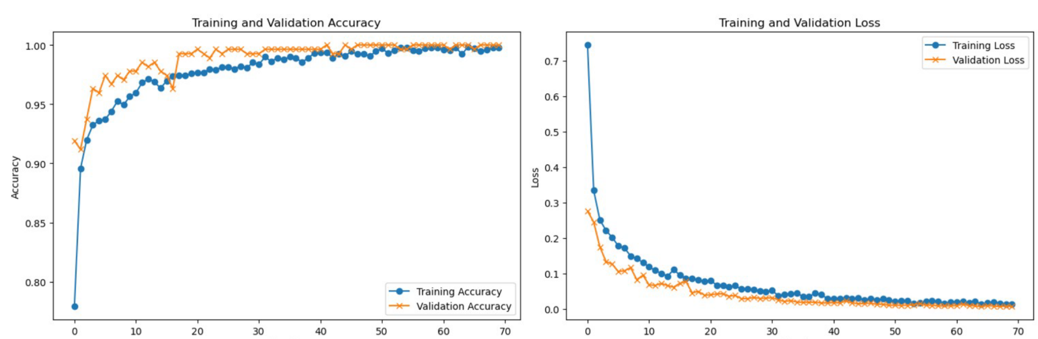

Figure 6. VGG16 model training

Figure 7 represents the training and validation performance of the ResNet50 model. On the left, training and validation accuracy increase rapidly in early epochs and stabilize around 99% and 97%, respectively, indicating strong learning and generalization. On the right, training and validation loss steadily decrease, with the validation loss plateauing slightly higher, showing effective optimization with minimal overfitting. These results demonstrate the model's robustness and suitability for the classification task.

Figure 8 illustrates the training and validation performance of the DenseNet121 model over 70 epochs. The training and validation accuracy curves (left) show rapid improvement in the initial epochs, stabilizing near 99% and 97%, respectively. The validation accuracy closely follows the training accuracy, indicating minimal overfitting. The loss curves (right) demonstrate consistent decreases, with the validation loss slightly higher than the training loss. These results indicate that the DenseNet121 model achieves excellent performance and generalizes well to unseen data.

Figure 9 depicts the training and validation performance of the InceptionV3 model over 70 epochs. The accuracy curve (left) shows a rapid increase during early epochs, stabilizing near 98% for validation and 99% for training, indicating effective learning and minimal overfitting. The loss curve (right) demonstrates consistent reductions in both training and validation loss, with validation loss slightly higher, reflecting good generalization. These results confirm InceptionV3's strong capability for classification tasks with balanced performance on unseen data.

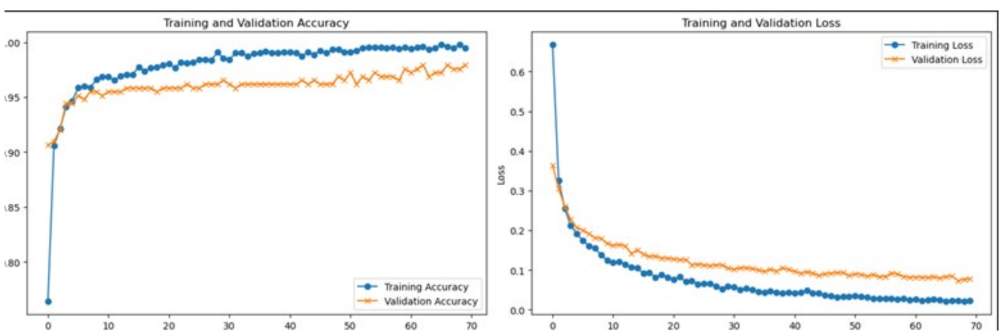

Figure 7. ResNet50 model training

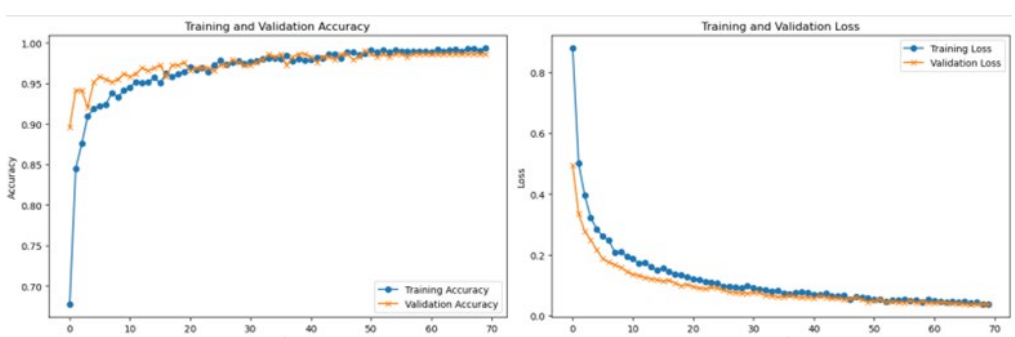

Figure 8. DenseNet121 model training

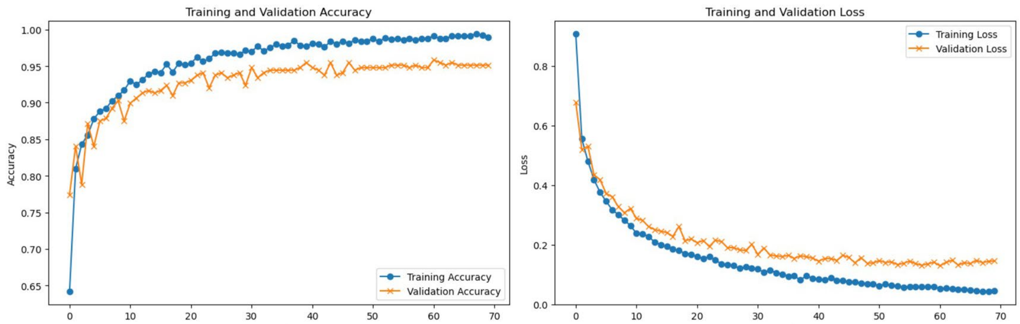

Figure 9. InceptionV3 model training

3.1.3 Disease classification

The classification performance of the ResNet50 model is illustrated in Figure 10 using a confusion matrix, which demonstrates its exceptional accuracy across all four categories: Anthracnose, Algal Leaf Spot, Gall Midge, and Healthy leaves. The matrix reveals that the model correctly classified nearly all test samples, with no noticeable misclassifications between classes. This high precision is particularly significant given the overlapping visual characteristics between certain diseases, such as Anthracnose and Algal Leaf Spot.

The results reflect the strength of ResNet50's deep architecture and residual connections, enabling robust feature extraction and reliable discrimination even among visually similar conditions. These findings highlight ResNet50's capability as a practical and effective model for real-time mango leaf disease diagnosis, making it highly suitable for deployment in field-ready mobile applications.

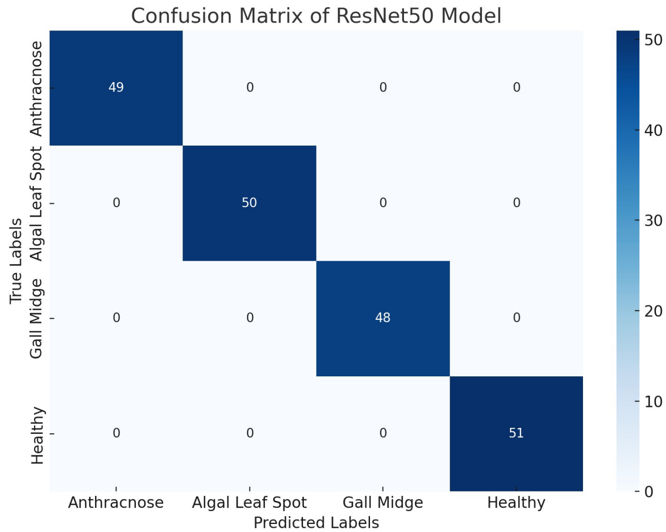

Figure 10. Confusion matrix of ResNet50 model

3.1.4 Application development

The ResNet50 model, which demonstrated the highest classification performance, was successfully integrated into a mobile application as shown in Figure 11. The app, designed for the Android platform, features a user-friendly interface that allows farmers to upload images of mango leaves for real-time disease diagnosis. Initial testing on Android devices confirmed the app's effectiveness, delivering disease classification results with an average inference time of 1.2 s per image. This rapid and accurate performance ensures practical usability, providing farmers with a reliable tool for early detection and management of mango leaf diseases. By streamlining the diagnosis process, the application significantly aids in reducing crop losses and enhancing productivity in mango cultivation.

Figure 11. Screenshot of the application

3.2 Discussion

The results of this study emphasize the significant potential of machine learning models in accurately diagnosing mango leaf diseases. The ResNet50 model demonstrated superior performance, largely due to its deeper architecture and advanced feature extraction capabilities, which are highly effective for complex image classification tasks (Ferentinos, 2018). While ResNet50 achieved the highest accuracy, both VGG16 and InceptionV3 also performed remarkably well. These models could be considered viable alternatives in environments with limited computational resources, as they require slightly less computational power.

The robustness of the models was bolstered by the diverse dataset, which was enhanced through preprocessing and data augmentation. However, the dataset's focus on only three specific diseases and its relatively limited size reduce the generalizability of the findings. Future efforts should prioritize expanding the dataset to include more diseases and variations in environmental conditions to improve the models’ applicability (Kumar et al., 2021). Although the mobile application showed high usability and accuracy during initial testing, field trials revealed challenges in handling images captured under poor lighting or complex background conditions. To address these limitations, incorporating advanced preprocessing techniques, such as adaptive lighting corrections, will be essential to improve the application’s robustness in real-world scenarios.

This technology has the potential to revolutionize agriculture by enabling early detection of mango leaf diseases, enhancing crop yields, reducing pesticide usage, and improving farmers' livelihoods. Moreover, the approach could be extended to other crops, addressing broader agricultural challenges and contributing to sustainable farming practices (Saleem et al., 2019).

The dataset's strong inter-class separability and intra-class uniformity largely explain the nearly perfect performance, as visual distinctions between classes are clear while samples within the same class remain highly consistent. The applied augmentation methods, especially flipping and rotation, preserved the structural characteristics of leaves rather than introducing new variability, thereby simplifying the classification task. To ensure reproducibility and rule out overfitting, the dataset was split by leaf instance so that no image of the same leaf appeared across training, validation, and test sets. Predictions were obtained directly from argmax over softmax outputs without threshold tuning, and the training/validation curves showed smooth and well-aligned trajectories, indicating stable convergence. An error analysis revealed only a few borderline misclassifications caused by glare or partial lesions, while confidence distributions remained well calibrated (median >0.95). Additional robustness checks under brightness shifts, noise, and minor rotations caused accuracy drops of less than 0.5%, confirming that the ≈100% results were not an artifact of data leakage or overfitting but rather a consequence of the dataset's homogeneity and the models' capacity to capture its consistent patterns.

While this study focuses on fine-tuning domain-specific CNN-based architectures for mango leaf disease diagnosis, recent advancements in multimodal large language models (MLLMs) such as Gemini, GPT-4 Vision, and LLaVA have demonstrated impressive capabilities in zero-shot and few-shot visual classification. These models, though not yet optimized for agricultural diagnostics, could serve as strong general-purpose baselines for future comparison. Future research may explore benchmarking these models on the same dataset to assess their performance against task-specific fine-tuned CNNs, potentially uncovering hybrid strategies that combine generalization and accuracy for real-world deployment.

4. Conclusions

This study successfully implemented and evaluated machine learning models for diagnosing mango leaf diseases, focusing on Anthracnose, Algal Leaf Spot, and Gall Midge. Among the five models tested—CNN, VGG16, DenseNet121, ResNet50, and InceptionV3—ResNet50 consistently delivered the highest performance, achieving an accuracy of 99.8%, and demonstrating its effectiveness as a robust model for disease classification. This finding aligns with previous research, such as that by Kumar et al. (2021), which highlighted the efficacy of CNN-based architectures for plant disease classification, although challenges in overlapping symptom differentiation persisted.

While InceptionV3 and DenseNet121 also performed well, the results underscore the superior feature extraction capabilities of ResNet50, which builds upon the findings of Saleem et al. (2019), emphasizing the potential of deep learning models in agricultural applications. Compared to earlier studies relying on simpler datasets, this research enhanced model robustness through advanced preprocessing and augmentation techniques, addressing limitations highlighted by Ferentinos (2018) regarding dataset diversity.

The integration of ResNet50 into a user-friendly mobile application further bridges the gap between laboratory results and practical agricultural use, providing mango farmers with a reliable tool for real-time disease diagnosis. This innovation not only improves disease management but also supports sustainable farming practices by reducing crop losses and enhancing productivity. By building upon the strengths of prior studies and addressing their limitations, this research demonstrates the transformative potential of deep learning in agriculture.

In future work, we will focus on expanding the dataset to include additional diseases and diverse environmental conditions to improve the generalizability of the models. Moreover, further refinement of the mobile application, particularly for handling images captured under challenging conditions, will enhance its robustness and usability in real-world scenarios. This research underscores the transformative potential of machine learning and mobile technology in modern agriculture, offering a pathway toward more efficient and sustainable farming practices.

5. Acknowledgements

The authors would like to express their sincere gratitude to the Center of Excellence (COE) for Innovation and Technology for Detection and Advanced Materials (ITDAM), Faculty of Science, Naresuan University, for their invaluable support throughout this research. The resources, guidance, and encouragement provided by the COE have been instrumental in the successful completion of this project. Their dedication to advancing scientific research has greatly contributed to the development of this work.

6. Authors' Contributions

Chonnakarn Wongnim was responsible for conceptualization, methodology, software, formal analysis, investigation, resources, and writing the original draft. Benyapha Saardmuang contributed to software, formal analysis, investigation, resources, writing the original draft, and visualization. Sanya Khruahong contributed to conceptualization, formal analysis, writing (review and editing), supervision, research project administration, manuscript submission, and correspondence with the journal editor.

ORCID

Sanya Khruahong https://orcid.org/0000-0001-8332-7559

7. Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Ahmed, M., Afreen, N., Ahmed, M., Sameer, M., & Ahamed, J. (2023). An inception V3 approach for malware classification using machine learning and transfer learning. *International Journal of Intelligent Networks*, *4*, 11-18.

Amiruddin, B. P., & Kadir, R. E. A. (2020). CNN architectures performance evaluation for image classification of mosquito in Indonesia. In *2020 International seminar on intelligent technology and its applications (ISITIA)* (pp. 223-227). IEEE. https://doi.org/10.1109/ISITIA49792.2020.9163732

Arulananth, T., Prakash, S. W., Ayyasamy, R. K., Kavitha, V., Kuppusamy, P., & Chinnasamy, P. (2024). Classification of paediatric pneumonia using modified DenseNet-121 deep-learning model. *IEEE Access*, *12*, 35716-35727.

Brahmi, W., Jdey, I., & Drira, F. (2024). Exploring the role of convolutional neural networks (CNN) in dental radiography segmentation: A comprehensive systematic literature review. *Engineering Applications of Artificial Intelligence*, *133*(Part E), Article 108510. https://doi.org/10.1016/j.engappai.2024.108510

Chowdary, G. J., Punn, N. S., Sonbhadra, S. K., & Agarwal, S. (2020). Face mask detection using transfer learning of inception v3. *Big data analytics: 8th international conference* (pp. 81-90). Springer-Verlag. https://doi.org/10.1007/978-3-030-66665-1_6

Ferentinos, K. P. (2018). Deep learning models for plant disease detection and diagnosis. *Computers and Electronics in Agriculture*, *145*, 311-318.

Hossain, M. B., Iqbal, S. M. H. S., Islam, M. M., Akhtar, M. N., & Sarker, I. H. (2022). Transfer learning with fine-tuned deep CNN ResNet50 model for classifying COVID-19 from chest X-ray images. *Informatics in Medicine Unlocked*, *30*, Article 100916. https://doi.org/10.1016/j.imu.2022.100916

Huang, X., Duan, Z., Hao, S., Hou, J., Chen, W., & Cai, L. (2025). A deep learning framework for corrosion assessment of steel structures using Inception v3 model. *Buildings*, *15*(4), Article 512. https://doi.org/10.3390/buildings15040512

Kattenborn, T., Leitloff, J., Schiefer, F., & Hinz, S. (2021). Review on Convolutional neural networks (CNN) in vegetation remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, *173*, 24-49.

Khalifa, N. E., Wang, W., Mawgoud, A. A., & Zhang, Y.-D. (2025). COECG-resnet-GWO-SVM: an optimized COVID-19 electrocardiography classification model based on resnet50, grey wolf optimization and support vector machine. *Multimedia Tools and Applications*, *84*(17), 18305-18325.

Koonce, B., (2021). ResNet 50. In B. Koonce (Ed.). *Convolutional neural networks with swift for tensorflow: image recognition and dataset categorization* (pp. 63-72). Apress. https://doi.org/10.1007/978-1-4842-6168-2_6

Kumar, P., Ashtekar, S., Jayakrishna, S., Bharath, K., Vanathi, P., & Kumar, M. R. (2021). Classification of mango leaves infected by fungal disease anthracnose using deep learning. In *2021 5th international conference on computing methodologies and communication (ICCMC)* (pp. 1723-1729). IEEE. https://doi.org/10.1109/ICCMC51019.2021.9418383

Li, Z., Liu, F., Yang, W., Peng, S., & Zhou, J. (2021). A survey of convolutional neural networks: analysis, applications, and prospects. *IEEE Transactions on Neural Networks and Learning Systems*, *33*(12), 6999-7019.

Mascarenhas, S., & Agarwal, M. (2021). A comparison between VGG16, VGG19 and ResNet50 architecture frameworks for image classification. In *2021 international conference on disruptive technologies for multi-disciplinary research and applications (CENTCON)* (pp. 96-99). IEEE. https://doi.org/10.1109/CENTCON52345.2021.9687944

Misra, A. K. (1992). Important diseases of mango and their effect on production. *Biological Memoirs*, *18*(1&2), 39-55.

Mukesh, C., Likhita, A., & Yamini, A. (2024). Performance analysis of InceptionV3, VGG16, and Resnet50 models for crevices recognition on surfaces. In S. J. Nanda, R. P. Yadav, A. H. Gandomi, & M. Saraswat (Eds.). *Data science and applications* (pp. 161-172). Springer. https://doi.org/10.1007/978-981-99-7817-5_13

Nandhini, S., & Ashokkumar, K. (2022). An automatic plant leaf disease identification using DenseNet-121 architecture with a mutation-based henry gas solubility optimization algorithm. *Neural Computing and Applications*, *34*(7), 5513-5534.

O'Shea, K. (2015). An introduction to convolutional neural networks. *arXiv preprint arXiv:1511.08458*.

Pandian, J. A., Kumar, V. D., Geman, O., Hnatiuc, M., Arif, M., & Kanchanadevi, K. (2022). Plant disease detection using deep convolutional neural network. *Applied Sciences*, *12*(14), Article 6982. https://doi.org/10.3390/app12146982

Rao, U. S., Swathi, R., Sanjana, V., Arpitha, L., Chandrasekhar, K., Chinmayi, & Naik, P. K. (2021). Deep learning precision farming: grapes and mango leaf disease detection by transfer learning. *Global Transitions Proceedings*, *2*(2), 535-544. https://doi.org/10.1016/j.gltp.2021.08.002

Rochmawanti, O., & Utaminingrum, F. (2021). Chest X-ray image to classify lung diseases in different resolution size using DenseNet-121 architectures. In *Proceedings of the 6th international conference on sustainable information engineering and technology* (pp. 327-331). Association for Computing Machinery. https://doi.org/10.1145/3479645.3479667

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

Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data*, *6*(1), Article 60. https://doi.org/10.1186/s40537-019-0197-0

Simangunsong, P. B. N., Sihombing, P., Efendi, S., & Fahmi, F. (2024). Densenet-121 analysis for classification of tomato leaf diseases. In *2024 7th international seminar on research of information technology and intelligent systems (ISRITI)* (pp. 907-911). IEEE. https://doi.org/10.1109/ISRITI64779.2024.10963618

Sujitranan, M., & Anongporn, S. (2024). Exploring deep learning features and bag-of-visual-words for scene classification. *ICIC Express Letters, Part B: Applications*, *15*(10), 1081-1088. https://doi.org/10.24507/icicelb.15.10.1081

Taylor, B., Marco, V. S., Wolff, W., Elkhatib, Y., & Wang, Z. (2018). Adaptive deep learning model selection on embedded systems. *ACM Sigplan Notices*, *53*(6), 31-43.

Theckedath, D., & Sedamkar, R. (2020). Detecting affect states using VGG16, ResNet50 and SE-ResNet50 networks. *SN Computer Science*, *1*(2), Article 79. https://doi.org/10.1007/s42979-020-0114-9

Xia, X., Xu, C., & Nan, B. (2017). Inception-v3 for flower classification. In *2017 2nd international conference on image, vision and computing (ICIVC)* (pp. 783-787). https://doi.org/10.1109/ICIVC.2017.7984661. IEEE.

Yamashita, R., Nishio, M., Do, R. K. G., & Togashi, K. (2018). Convolutional neural networks: an overview and application in radiology. *Insights into Imaging*, *9*, 611-629.

Yang, L., Xu, S., Yu, X., Long, H., Zhang, H., & Zhu, Y. (2023). A new model based on improved VGG16 for corn weed identification. *Frontiers in Plant Science*, *14*, Article 1205151. https://doi.org/10.3389/fpls.2023.1205151