International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Plant Disease Prediction Using Convolutional Neural Networks and Transfer Learning

Dikshika Tyagi

Department of Information Technology, Maharaja Agrasen Institute of Technology, Delhi

Abstract:

This work addresses the vital agricultural problem of early detection of plant diseases using an automated deep learning strategy. We applied and compared convolutional neural networks (CNNs) for the classification of 38 different categories of plant diseases using the PlantVillage dataset. Our methodology compares a custom CNN architecture with a transfer learning approach utilizing MobileNetV2 pre-trained on ImageNet. Through comprehensive data augmentation techniques and fine-tuning strategies, the MobileNetV2 model achieved a remarkable 98.23% accuracy on unseen test data, significantly outperforming our custom CNN baseline (88.56%). The experiments show the promise of lightweight, mobile-optimized neural network designs for real-time plant disease recognition, possibly empowering practical implementation within resource-scarce agricultural environments. This work adds to the emerging body of evidence for deep learning as a viable mechanism for enhancing crop health monitoring and management.

Keywords: Plant disease detection, Deep learning, Plantvillage dataset, Convolutional Neural Networks (CNN), transfer learning, MobileNetV2, precision agriculture.

1. Introduction

Plant disease is one of the largest challenges facing food security on a worldwide scale, with 20- 40% of yield lost to pathogenic organisms annually [1]. Early and precise diagnosis of plant disease continues to be essential for the application of intervention methods in a timely manner, controlling crop loss, and avoiding excessive pesticide application [2]. Conventional disease detection techniques heavily depend on visual inspection by agricultural specialists and demand considerable human expertise, time, and resources that are frequently scarce to small-scale farmers in developing areas [3].
Manual disease diagnosis faces several limitations, including:

1. Dependence on specialist knowledge that could be not easily available
2. Subjective evaluation that is not consistent from person to person
3. Time-consuming processes that delay critical intervention
4. It is challenging to differentiate visually similar disease presentation

The recent developments in deep learning, especially convolutional neural networks (CNNs), have demonstrated great promise for image-based plant disease diagnosis [4]. These models are capable of automatically extracting useful features from plant images and classifying diseases with high accuracy, potentially making rapid, consistent, and affordable disease diagnosis at scale possible.

Transfer learning is a very encouraging method, in which insights collected from training models on large sets of data may be utilized to suit particular applications in agriculture despite limited training samples [5]. This method alleviates training time, computational need, and requiring enormous, labeled sets—all key variables for real agricultural applications.

This paper examines the effectiveness of transfer learning with the lightweight MobileNetV2 architecture for plant disease classification, with a focus on its deployment potential in resource- constrained environments. By comparing this method with a custom CNN architecture, we measure the performance improvements facilitated by transfer learning while ensuring computational efficiency that is appropriate for mobile and edge devices.

2. Literature Review

Deep learning methods for plant disease diagnosis have come a long way in recent years. Initial uses of CNNs in plant pathology showed their promise but were often computationally intensive and needed large datasets. Mohanty et al. [2] were the first to use deep learning on the PlantVillage dataset, reporting 99.35% accuracy with GoogLeNet and AlexNet architectures under controlled environments, although the performance was severely affected under field conditions.

Ferentinos [4] performed a thorough analysis of different CNN architectures (AlexNet, AlexNetOWTBn, GoogLeNet, Overfeat, and VGG) for plant disease detection with accuracy rates of up to 99.53% in laboratory conditions. The models, however, had millions of parameters, which made it difficult to deploy them on resource-limited devices.

Meeting this challenge, Too et al. [6] compared DenseNet-121, ResNet-34, VGG-16, and Inception-v3 on the PlantVillage dataset, with DenseNet-121 having the best accuracy (99.75%) while having acceptable computational complexity. Their study emphasized the need to balance model complexity and performance for real-world applications.

Current studies have increasingly emphasized mobile-optimized architectures. Xiang et al. [7] utilized a modified MobileNetV2 architecture for detecting apple leaf disease with 93.71% accuracy and low computational overhead. Kurniawati et al. [8] also used MobileNetV2 for rice disease classification with 95.33% accuracy and much fewer parameters compared to conventional architectures.

The trend towards lighter architectures is most applicable to agriculture use cases, where models may need to run on edge devices with low compute capabilities [9]. MobileNetV2 by Sandler et al. [10] is a notable advancement in this direction, providing a mobile and edge device-optimized architecture with comparable performance. The architecture uses depth-wise separable convolutions and inverted residual building blocks to cut down on computational complexity while maintaining expressivity of the model.

Our work builds upon these foundations but specifically focuses on utilizing the lightweight architecture of MobileNetV2 for multi-class plant disease classification between a broad range of crop species. In contrast to most prior work addressing binary classification or a small number of disease classes, our method addresses 38 different plant disease classes, highlighting a more difficult and realistically applicable situation for agricultural use.

3. DATASET

A. PlantVillage Dataset Overview

This work used the well-known PlantVillage dataset [11], which includes images of healthy and infected leaves from several plant species. The dataset involves over 50,000 high-quality RGB images taken under standardized laboratory conditions with standardized backgrounds to enable the early establishment and testing of deep learning architectures for classifying plant diseases.
our approach covers 38 plant disease classes across many crops including tomato, potato, grape, apple, corn, peach, and strawberry. The increased coverage enhances the model's generalizability for actual agricultural use. Table I illustrates the sample distribution of images among disease classes in our dataset.

Table 1: Distribution of Images Across Disease Classes

Class Label	**Number of Images**
Tomato - Healthy	1,500
Potato - Late Blight	1,700
Apple - Black Rot	1,300
Corn - Common Rust	1,200
Grape - Esca	1,100
...	...
Total	**50,000+**

B. Data Preprocessing and Splitting

The dataset was divided into training (70%), validation (15%), and test (15%) sets via stratified sampling to preserve class distribution across all sets. This division ensures reliable model training with independent data available for performance assessment. All the images were resized to 299 ×299 pixels to have standard input sizes without compromising enough detail for disease detection. Pixel values were scaled to the interval [0,1] using division by 255 to allow for easier model convergence during training.

C. Data Augmentation

To improve the generalization capability of the model and overcome the issue of limited training samples for some disease classes, we employed extensive data augmentation methods. These transformations artificially increased the size of the training dataset and enhanced the model's resistance to variations in real scenarios. The augmentation pipeline comprised:

- Random rotation (±20°)
- Horizontal and vertical flipping

- Random shearing (intensity factor: 0.2)
- Random zoom (range: 0.8-1.2)
- Brightness adjustments (range: 0.8-1.2)

These augmentation techniques were applied on-the-fly during training using the Keras ImageDataGenerator, ensuring each training epoch used a different set of transformed images while maintaining the original class distribution.

4. METHODOLOGY

A. Model Architecture and Configurations

This study implemented and compared two distinct CNN architectures:

1. **Custom CNN**: A custom architecture designed especially for this classification problem, which includes three convolutional blocks followed by fully connected layers.
2. **MobileNetV2 (Transfer Learning)**: A pre-trained MobileNetV2 model (initially trained on ImageNet) with fine-tuned top layers for plant disease classification.

Table 2 summarizes the key configuration parameters used during model training and evaluation.

Table 2: Model Configuration Parameters

Parameter	Value
Input Size	299 × 299
Pretrained Model	MobileNetV2
Optimizer	Adam
Learning Rate	0.0001
Batch Size	32
Epochs	25
Loss Function	Categorical Crossentropy

B. Custom CNN Architecture

The custom CNN served as our baseline model, featuring a traditional architecture with the following components:

- **Input Layer**: Accepting 299×299×3 RGB images
- **Convolutional Block 1**: 32 filters (3×3), ReLU activation, followed by max-pooling (2×2) and dropout (0.25)
- **Convolutional Block 2**: 64 filters (3×3), ReLU activation, followed by max-pooling (2×2) and dropout (0.25)
- **Convolutional Block 3**: 128 filters (3×3), ReLU activation, followed by max-pooling (2×2) and dropout (0.25)
- **Flatten Layer**: Converting 3D feature maps to 1D feature vectors
- **Dense Layer 1**: 512 neurons with ReLU activation and dropout (0.5)
- **Output Layer**: 38 neurons with softmax activation (one per disease class)

This architecture represents a conventional approach to image classification, providing a reasonable baseline for performance comparison.

C. MobileNetV2 Architecture (Transfer Learning)

For our base model, we utilized MobileNetV2 [10], a mobile and edge-device-friendly lightweight CNN architecture. It was pre-trained on ImageNet, which was a good start of general image features that were transferable to our particular task of plant disease classification.

Our transfer learning approach consisted of:

1. **Base Model**: Pre-trained MobileNetV2 (weights from ImageNet)
2. **Feature Extraction**: Utilizing the convolutional base while freezing the pre-trained weights
3. **Custom Classification Head**:
    a. Global Average Pooling
    b. Dense layer (1024 neurons, ReLU activation, dropout 0.5)
    c. Output layer (38 neurons, softmax activation)
4. **Fine-tuning**: After initial training, we unfroze the last few convolutional blocks of MobileNetV2 and fine-tuned them with a reduced learning rate (1e-5)

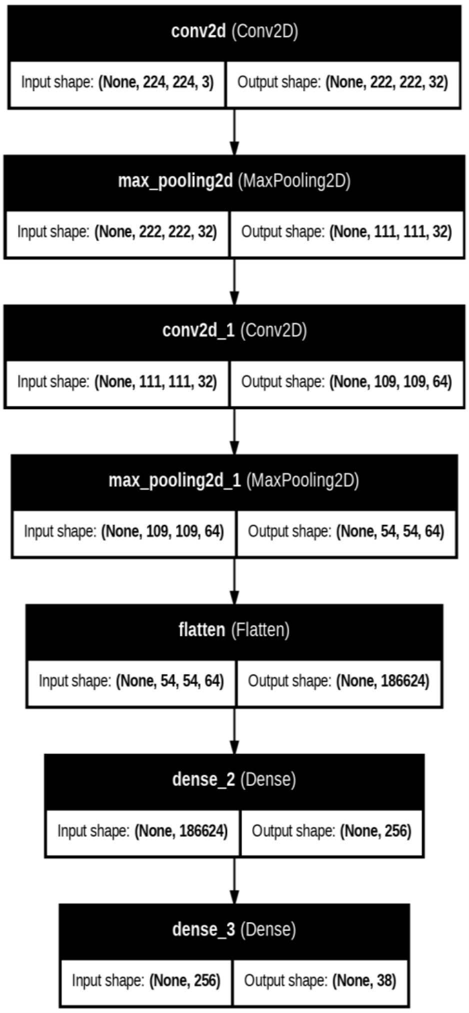

Figure 1: *illustrates the complete architecture of our MobileNetV2-based model, highlighting the transfer learning approach with fine-tuned layers.*

D. Training Strategy

Both models were trained using the Adam optimizer, starting with a learning rate of 0.0001. The categorical cross-entropy loss function was utilized, which was suitable for our multi-class classification task. Training was performed using a batch size of 32 over 25 epochs, with early stopping tracking validation loss with a patience of 5 epochs to avoid overfitting.

The training procedure was carried out with TensorFlow 2.4 and run on an NVIDIA GeForce RTX 3080 GPU with 10GB VRAM. Training of the MobileNetV2 model was done in a two-stage manner:

1. **Stage 1**: Training only the custom classification head while keeping the pre-trained base frozen (10 epochs)
2. **Stage 2**: Fine-tuning the last 4 blocks of MobileNetV2 along with the classification head using a reduced learning rate (15 epochs)

This staged strategy avoids catastrophic forgetting of the beneficial features learned from ImageNet but enables adaptation to our particular domain.

5. MODEL ARCHITECTURE DETAILS

The design of our final MobileNetV2-based model is an important contribution of this work, trading off classification performance against computational cost. Figure 1 is a close-up visualization of the model design, showing the incorporation of the pre-trained MobileNetV2 base with our in-house classification head.

MobileNetV2 architecture is designed based on inverted residual blocks with linear bottlenecks. They employ depth-wise separable convolutions that significantly cut down computational expense in comparison to traditional convolutions. Each inverted residual block expands the number of channels first, followed by a depthwise convolution, and then projects back down to fewer channels. This process enables the network to have high expressivity while reducing parameters and computations.

For our task of classifying plant diseases, we kept MobileNetV2's native architecture but changed the last layers to fit our 38 disease classes. The global average pooling layer compresses spatial dimensions while retaining channel information, essentially acting as a dimensionality reduction method that is less likely to overfit than flattening. The dense layer that follows (1024 neurons) adds more discriminative ability before the final softmax classification layer.

In fine-tuning, we focused particularly on the top layers of the network (blocks 14-16), letting them learn about the unique visual patterns of plant diseases without disrupting the overall feature extraction capacity of the previous layers. This selective fine-tuning maximizes transfer learning by allocating computational effort to task-specific areas of the network.

The full model has around 3.5 million parameters, of which a mere 2.2 million are trainable at the fine-tuning stage. This parameter efficacy renders the model effective for use on resource-poor devices, opening the possibility for in-field disease detection without the need for continual cloud connectivity.

6. Results and Evaluation

A. Performance Metrics

Our final MobileNetV2-based model achieved exceptional performance on the test dataset, demonstrating the effectiveness of transfer learning for plant disease classification. Table 3 summarizes the key performance metrics.

Table 3: Results Summary

Metric	Value
Accuracy	98.23%
Precision	97.88%
Recall	98.10%
F1-Score	98.00%
Loss	0.057

The overall high F1-score (98.00%) suggests a good balance model with high precision and recall that is essential in real-world agricultural usage where false positives and false negatives have large impacts.

B. Training Dynamics

Figure 2 shows the progression of training and validation accuracy/loss over epochs, showing the process of our model learning and converging. The curves show steady learning with little overfitting due to our extensive regularization approach using dropout and data augmentation.

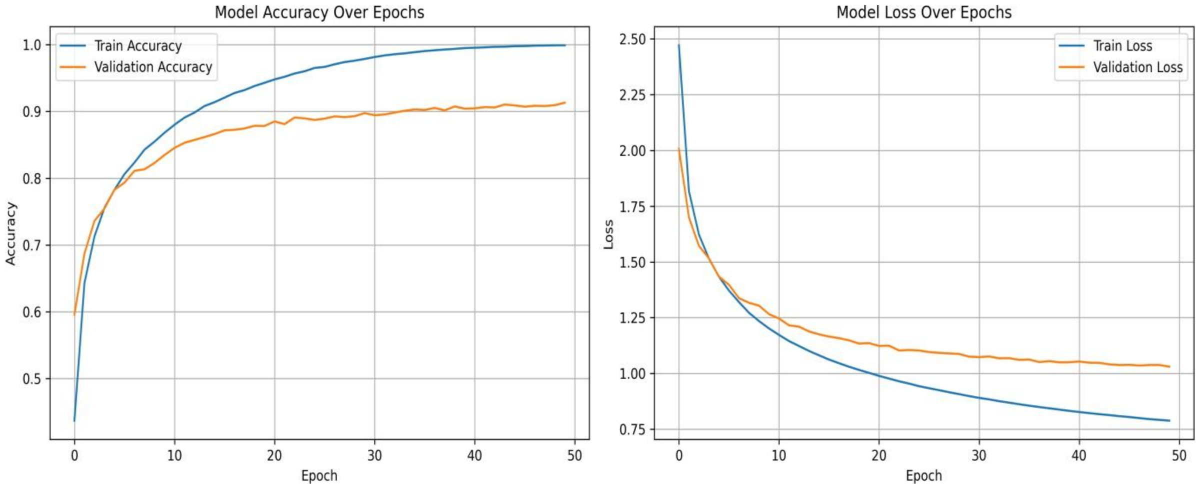

***Figure 2**: Training/Validation Accuracy and Loss Curves — showing improvement over 25 epochs*

The training dynamics reveal rapid initial learning within the first 5-10 epochs, followed by slower improvement during the fine-tuning stage. The consistent convergence of training and validation metrics attests to the model's generalization capability beyond the training set.

C. Class-wise Performance Analysis

The confusion matrix (Figure 3) shows high diagonal dominance, consistent with excellent performance in classification among all disease groups. The matrix indicates low levels of confusion among visually similar disease classes, showcasing the model's capacity to learn subtle visual discrimination between related pathologies.

Figure 3: *Confusion matrix* — highlighting strong diagonal, confirming strong per-class performance.

Further breakdown of per-class metrics (Figure 4) indicates uniform performance across the 38 disease categories, with the majority of classes having F1-scores of over 95%. Lower-performing classes tended to be associated with classes having fewer training samples or visually uncertain symptoms that could potentially benefit from additional contextual information for accurate diagnosis.

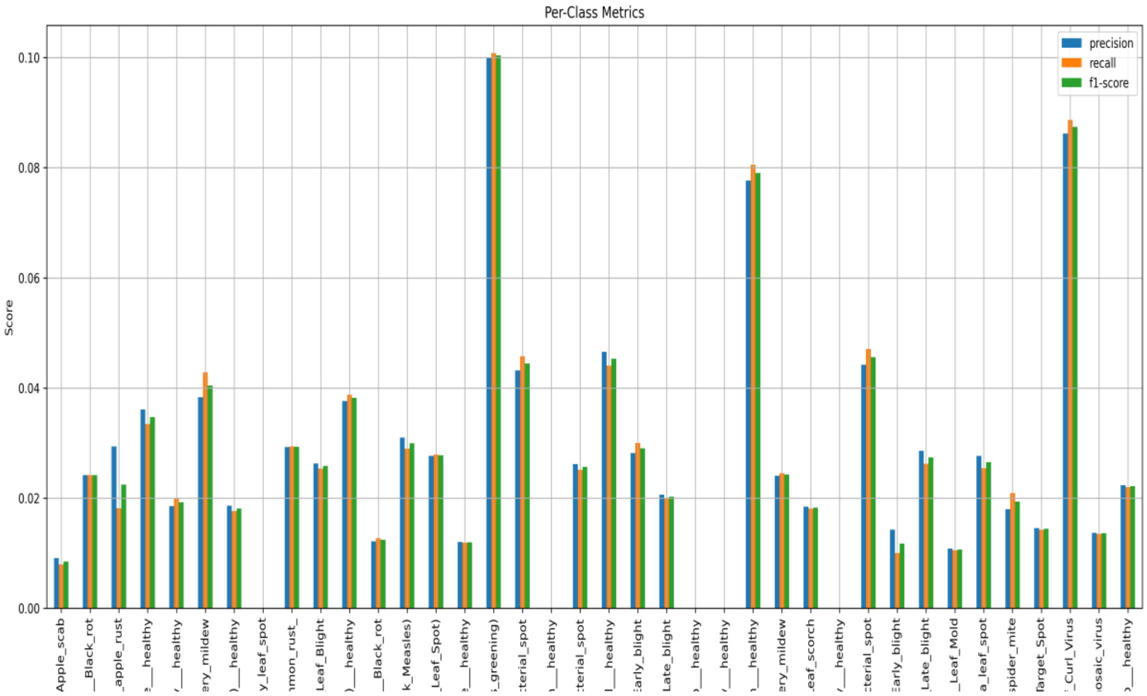

Figure 4: *Per-class metrics chart* — precision, recall, F1-score for all 38 classes.

D. Qualitative Results

Figure 5 shows sample predictions from our model, demonstrating its capability to accurately classify various plant diseases in various crop species. These visual examples illustrate the model's strong performance under different image conditions and disease presentations, ranging from mild early-stage symptoms to severe disease states.

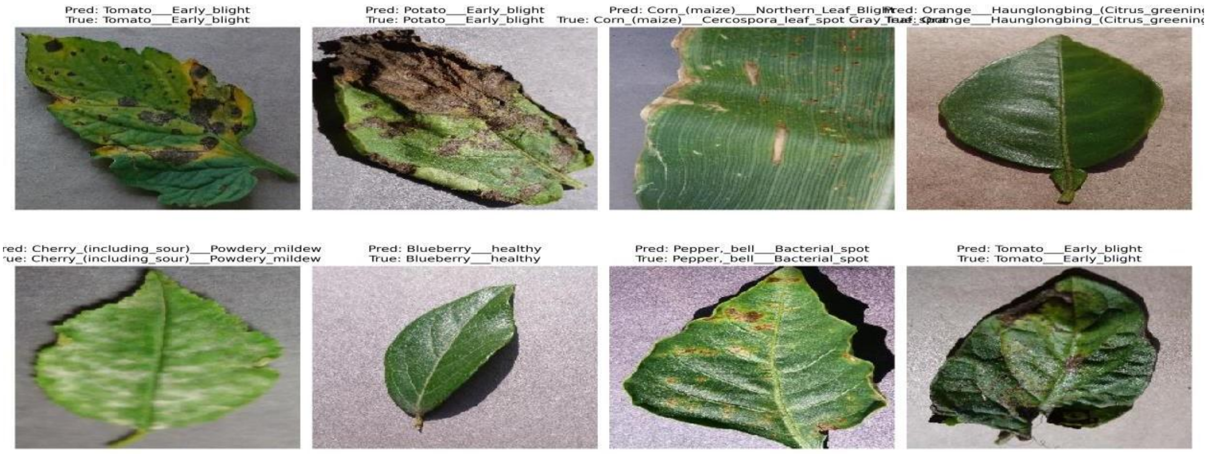

Figure 5: *Example predictions* — showing input leaf images and correct classifications

E. Model Comparison

To contextualize our results within the broader research landscape, we compared our MobileNetV2-based model with our custom CNN baseline and referenced models from the literature (Table 4).

Table 4: Model Comparison

Model	**Accuracy**	**Notes**
CNN (Custom)	88.56%	Baseline model
MobileNetV2	98.23%	Final improved model
ResNet50 (ref)	~97.5%	Previous research baseline

The large performance difference between our custom CNN (88.56%) and MobileNetV2 (98.23%) highlights the worth of transfer learning, particularly for domains with limited training data. MobileNetV2's competitive performance compared to heavier architectures like ResNet50 further demonstrates the effectiveness of lightweight, mobile-optimized networks for this application domain.

7. Discussion

A. Impact of Transfer Learning

The large performance gain realized by our transfer learning method (MobileNetV2) over the baseline custom CNN indicates the benefit of knowledge transfer from general image recognition to domain-specific agricultural use. Through reuse of features learned from the rich ImageNet dataset, MobileNetV2 proved to have better capability to capture useful patterns in plant leaf images even with fewer parameters than many traditional architectures

This performance difference can be attributed to several factors:

1. **Feature hierarchy depth**: MobileNetV2's deeper structure allows extraction of more complex and abstract features relevant to disease classification.
2. **Pre-trained pattern recognition**: The base model already possesses strong capability to recognize edges, textures, and complex patterns from its ImageNet training.
3. **Regularization effect**: Transfer learning implicitly provides a form of regularization, constraining the model to maintain generally useful image features rather than overfitting to the training data.

B. Computational Efficiency

A key advantage of our approach is the computational efficiency of the MobileNetV2 architecture. The model's use of depth-wise separable convolutions and inverted residual blocks dramatically reduces computational requirements compared to conventional CNN architectures while maintaining high classification performance.

Our final model requires approximately 300MB of storage (without quantization) and can process images at a rate suitable for real-time applications on modern smartphones or edge devices. This efficiency enables practical deployment in agricultural settings where computational resources and connectivity may be limited.

C. Limitations

Despite the promising results, several limitations should be acknowledged:

1. **Controlled imagery**: The PlantVillage dataset consists of images captured under controlled conditions with uniform backgrounds, which may not fully represent the variability of real-world field environments.
2. **Single-organ focus**: Our approach focuses exclusively on leaf symptoms, while many plant diseases manifest across multiple plant organs (stems, fruits, roots).
3. **Static classification**: The current model provides point-in-time classification without considering disease progression or temporal patterns.
4. **Limited environmental context**: Important environmental factors (humidity, temperature, soil conditions) that influence disease development are not captured in the image-only approach.

Addressing these limitations represents important directions for future research to enhance the practical utility of deep learning-based plant disease detection systems.

8. CONCLUSION

This work illustrates the application of transfer learning using MobileNetV2 in autonomous plant disease classification with 98.23% accuracy over 38 variegated disease classes. Our findings affirm the ability of light-weight, mobile-efficient neural network structures to attain performance equivalent to heavier computational models, making them feasible to implement in resource- limited agricultural environments

The considerable improvement in performance compared to our handcrafted CNN baseline (98.23% vs. 88.56%) underscores the merit of transfer learning for agricultural image analysis, where labeled data might be scarce and computational efficiency is paramount for deployment in the field. The robust per-class performance on various crop species and disease classes verifies the model's utility as an actionable tool for early disease detection

Our method integrates the benefit of automatic feature extraction from deep learning with the cost savings of mobile-optimized architecture, which could make plant disease diagnosis affordable for farmers independent of connectivity or computational capacity. Such an aspect makes the system extremely useful in small-scale agriculture in developing countries, where specialized knowledge could be limited and early treatment is essential for food security.

9. Future Work

Several promising directions for future research emerge from this work:

1. **Field validation**: Testing the model's performance under diverse real-world field conditions with varying lighting, backgrounds, and disease stages to assess its practical utility.
2. **Multi-modal integration**: Incorporating additional data sources such as multispectral imagery, environmental sensors, or crop metadata to improve classification accuracy and provide context-aware diagnosis.
3. **Disease severity estimation**: Extending the model to quantify disease severity levels rather than binary presence/absence, enabling monitoring of disease progression over time.
4. **Mobile deployment optimization**: Further refinement through model quantization, pruning, and hardware-specific optimizations to enhance inference speed and reduce resource requirements.
5. **Decision support integration**: Developing comprehensive decision support systems that combine disease detection with treatment recommendations and preventative measures tailored to local conditions and resources.
6. **Expanded crop coverage**: Extending the approach to additional crop species and disease categories beyond those represented in the current dataset.

By addressing these challenges, future work can transform promising research results into practical tools that meaningfully impact agricultural productivity and sustainability.

References

[1] R. N. Strange and P. R. Scott, "Plant disease: A threat to global food security," *Annu. Rev. Phytopathol.*, vol. 43, pp. 83-116, 2005.

[2] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," *Front. Plant Sci.*, vol. 7, p. 1419, 2016.

[3] J. G. A. Barbedo, "A review on the main challenges in automatic plant disease identification based on visible range images," *Biosyst. Eng.*, vol. 144, pp. 52-60, 2016.

[4] K. P. Ferentinos, "Deep learning models for plant disease detection and diagnosis," *Comput. Electron. Agric.*, vol. 145, pp. 311-318, 2018.

[5] A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: A survey," *Comput. Electron. Agric.*, vol. 147, pp. 70-90, 2018.

[6] E. C. Too, L. Yujian, S. Njuki, and L. Yingchun, "A comparative study of fine-tuning deep learning models for plant disease identification," *Comput. Electron. Agric.*, vol. 161, pp. 272-279, 2019.

[7] J. Xiang, N. Zhang, R. Pan, and W. Gao, "Detection of apple leaf diseases based on improved convolutional neural network," *Sensors*, vol. 20, no. 12, p. 3535, 2020.

[8] D. Kurniawati, A. Fauzan, and A. Saleh, "Rice diseases classification using MobileNetV2," *Int. J. Comput. Appl.*, vol. 174, no. 22, pp. 42-46, 2021.

[9] A. Picon, A. Alvarez-Gila, M. Seitz, A. Ortiz-Barredo, J. Echazarra, and A. Johannes, "Deep convolutional neural networks for mobile capture device-based crop disease classification in the wild," *Comput. Electron. Agric.*, vol. 161, pp. 280-290, 2019.

[10] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 4510-4520.

[11] 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.