Comparative Analysis of Custom CNN and Pretrained Transfer Learning Models for Plant Disease Detection

Goljana Sree Ratna Saketh
Department of Computer Science and Engineering
Lovely Professional University
Phagwara, India
sakethgoljana098@gmail.com

Ponnana Sri Sai Manikantha
Department of Computer Science and Engineering
Lovely Professional University
Phagwara, India
ponnanasaisaimanikanta@gmail.com

Kuppili Hemanth
Department of Computer Science and Engineering
Lovely Professional University
Phagwara, India
kuppilihemanth628@gmail.com

Gaurav Sharan
Department of Computer Science and Engineering
Lovely Professional University
Phagwara, India
kumargaurav8707@gmail.com

***Abstract*—Effective and precise identification of plant diseases through leaf images is crucial for the sustainability of agriculture and timely intervention methods. This study offers a comparative evaluation of a lightweight custom Convolutional Neural Network (CNN) alongside three well-known pretrained transfer learning models—VGG16, ResNet50, and MobileNetV2—which were trained and assessed using the PlantVillage dataset. All the models utilized the same data augmentation methods and regularization techniques, including L2 weight decay, early stopping, and adaptive learning rate schedules to reduce overfitting. Extensive experiments assess the performance of each model based on key metrics such as validation accuracy, F1-score, number of parameters, and speed of training convergence. The findings reveal that, although transfer learning models generally deliver better accuracy, the custom CNN provides competitive results while using significantly fewer computational resources, particularly in scenarios with limited data and resources.The results emphasize the possibility of implementing optimized custom designs in practical agricultural uses, promoting sustainable computing on a larger scale.**

***Index Terms*—Plant Disease Detection, Convolutional Neural Network, Transfer Learning, Deep Learning, VGG16, ResNet50, MobileNetV2, Data Augmentation, Regularization, Overfitting Prevention, Sustainable Computing, Agricultural AI**

I. Introduction

Timely and precise identification of plant illnesses is vital for promoting agricultural sustainability and reducing crop losses. Quickly recognizing diseases in plant leaves facilitates prompt action, which is critical for preserving global food security. Deep learning, especially Convolutional Neural Networks (CNNs), has become an effective method for automating the detection process by examining leaf images with great precision.

A. Research Motivation

Despite the success of transfer learning models pretrained on large datasets, their deployment in resource-constrained environments remains challenging due to high computational requirements. Moreover, large pretrained models may overfit or underperform when applied to domain-specific tasks with limited data availability. This motivates exploring lightweight, custom-designed CNN architectures that strive to balance accuracy with efficiency on plant disease datasets.

B. Contributions and Scope

This paper presents a comprehensive comparative analysis of a custom CNN model against three widely used pretrained transfer learning architectures (VGG16, ResNet50, and MobileNetV2) on the PlantVillage leaf disease dataset. Key contributions include:

- Designing and training a lightweight custom CNN with regularization and augmentation to prevent overfitting.
- Conducting rigorous experiments comparing model accuracy, training convergence, and parameter efficiency.
- Providing insights into when custom CNNs can match or outperform complex pretrained models in limited data scenarios.
- Discussing implications for sustainable computing in agricultural diagnostics.

This work is targeted towards advancing practical AI solutions for plant disease detection with an emphasis on sustainable and accessible computing.

II. Literature Review

A. Deep Learning for Plant Disease Detection

Deep learning methods, particularly Convolutional Neural Networks (CNNs), have transformed the area of identifying plant diseases by allowing for automated and accurate classification of diseases based on leaf images. Numerous studies have utilized CNN models such as AlexNet, VGGNet, and custom-designed networks trained on the PlantVillage dataset to achieve exceptional accuracy [1][2]. These models derive hierarchical features that capture variations in texture, color, and shape brought about by diseases, surpassing conventional image processing techniques.

B. Transfer Learning in Agricultural Vision

Transfer learning, which involves fine-tuning pretrained CNNs on agriculture-specific datasets, has become a popular strategy to overcome limited data scenarios. Models such as VGG16, ResNet50, and MobileNetV2, pretrained on ImageNet, have been adapted successfully for plant disease classification tasks [3][4]. Transfer learning accelerates convergence and improves generalization but may introduce challenges due to domain differences between natural and plant images.

C. Overfitting and Model Optimization Strategies

Overfitting presents a significant challenge when developing deep models with small datasets. Methods such as dropout, L2 weight decay, and batch normalization have been demonstrated to enhance the resilience of models. Utilizing early stopping guided by validation loss and implementing adaptive learning rate schedules can also aid in reducing overfitting and ensuring smoother training. Data augmentation is crucial for increasing the effective size of the training set by adding realistic variations to the input images.

III. Dataset Description

A. Source and Composition

The dataset used in this study is the PlantVillage dataset, publicly available on Kaggle. It comprises over 54,000 images covering 38 classes of plant leaf diseases, including healthy leaf categories. The images vary in size and quality but are primarily high-resolution color images of tomato, potato, apple, and other crops' leaves exhibiting diverse disease symptoms. This rich dataset provides a realistic and challenging benchmark for plant disease classification models.

B. Preprocessing and Augmentation

To prepare the data for training, images are resized to 128×128 pixels and normalized by scaling pixel values to the range [0,1]. To improve model generalization and combat overfitting, extensive data augmentation is applied during training. Augmentation techniques include random rotations (up to 20 degrees), width and height shifts (up to 20%), zoom, shear transformations, and horizontal flips. These transformations simulate realistic variations in leaf orientation and lighting, effectively enlarging the training set diversity.

IV. Methodology

A. System Overview

The proposed system for plant disease detection follows a modular pipeline consisting of data acquisition, preprocessing, model training, and evaluation. Leaf images from the PlantVillage dataset are first resized and normalized, then augmented to enhance diversity. The system trains multiple models, including a custom-designed Convolutional Neural Network (CNN) and three pretrained transfer learning models (VGG16, ResNet50, and MobileNetV2). The models are trained with regularization techniques, and their performances are compared based on accuracy, training efficiency, and generalization capability.

B. Custom CNN Architecture

The custom CNN is designed to be lightweight yet effective, targeting environments with resource constraints. It consists of four convolutional blocks with convolutional layers followed by batch normalization, max-pooling, and dropout for regularization. The architecture progressively increases the number of filters (32, 64, 128, 256) to capture hierarchical features. The convolutional layers use ReLU activations, and the final classifier comprises fully connected dense layers with L2 regularization and dropout before the softmax output layer for multi-class classification.

C. Transfer Learning Models: VGG16, ResNet50, MobileNetV2

Pretrained models VGG16, ResNet50, and MobileNetV2, originally trained on ImageNet, are used as feature extractors by freezing most layers and fine-tuning the top layers. Each base model excludes its original top fully connected layers and is followed by global average pooling, two dense layers with L2 regularization and dropout, and a softmax output layer matching the number of plant disease classes. This approach leverages learned general features while adapting to the plant disease classification task.

D. Training Configuration and Hyperparameters

All models utilize the Adam optimizer, starting with a learning rate of 0.0005. The batch size is configured to 32, and training can continue for a maximum of 20 epochs, incorporating early stopping to end training if the validation loss fails to improve over 5 consecutive epochs. A learning rate reduction on plateau is employed, which halves the rate if the validation loss remains unchanged for 3 epochs, enhancing convergence stability. Online data augmentation is utilized during training to boost generalization. Regularization methods include dropout rates ranging from 0.3 to 0.5, along with L2 weight decay (0.001) applied to dense layers.

V. Experimental Setup

A. Computing Resources

The experiments were performed on the Google Colab platform, which provides access to NVIDIA Tesla T4 GPUs with 16 GB of RAM, sufficient for training deep learning models efficiently. The software environment included Python 3.10 and widely used libraries such as TensorFlow 2.x for building and training neural networks, scikit-learn for evaluation metrics, Pillow for image processing, and Matplotlib and Seaborn for data visualization. This cloud-based setup offers the advantage of scalable computational resources with easy accessibility, allowing for reproducible experiments without the need for dedicated high-performance hardware.

Using Colab also closely simulates practical deployment constraints relevant for agricultural settings where computing resources may be limited. The models were trained using GPU acceleration to reduce training time and enable experimentation with multiple architectures within a reasonable timeframe.

B. Experimental Procedure

The experimental process followed a systematic and reproducible pipeline:

1) **Dataset Acquisition and Organization:** The PlantVillage dataset, comprising over 54,000 images spanning 38 classes, was automatically downloaded from Kaggle using API authentication. The dataset was organized into training and validation subsets using an 80:20 split, preserving class balance.
2) **Data Preprocessing and Augmentation:** Images were resized to 128×128 pixels and normalized to a $[0, 1]$ scale. Real-time data augmentation was applied during training to simulate image variations commonly encountered in field conditions, improving model robustness. Augmentation techniques included random rotations (±20°), width and height shifts (up to 20%), zooming, shearing, and horizontal flipping.
3) **Model Initialization:** Four model architectures were initialized: a custom CNN designed from scratch, and three transfer learning models—VGG16, ResNet50, and MobileNetV2—using pretrained weights from ImageNet. Transfer learning models had their base layers frozen initially, with a trainable classification head added.
4) **Training Protocol:** All models were trained using the Adam optimizer with an initial learning rate of 0.0005 and a batch size of 32. Early stopping monitored validation loss with a patience of 5 epochs to prevent overfitting. A ReduceLROnPlateau callback reduced the learning rate when validation loss plateaued for 3 epochs. Each training was limited to a maximum of 20 epochs to balance training time and model convergence.
5) **Validation and Evaluation:** After each epoch, models were validated on the held-out set to monitor generalization. The best performing weights based on minimum validation loss were saved for final evaluation.
6) **Metrics and Visualization:** Classification accuracy, precision, recall, and F1-score were computed on the validation data. Confusion matrices were generated to analyze class-wise performance. Training and validation accuracy and loss curves were plotted over epochs to visualize convergence and potential overfitting. All figures and tables were saved for inclusion in the manuscript.

This rigorous procedure ensured consistent comparison across models while providing insights into their efficacy, computational efficiency, and suitability for plant disease detection tasks under limited computational resources.

VI. Results

A. Quantitative Performance Comparison

Table I summarizes classification metrics for all models. Transfer learning models such as ResNet50 and VGG16 generally achieve superior accuracy, however the Custom CNN provides competitive results with a lower parameter footprint.

Model	Accuracy	Precision	Recall	F1-Score	Params (M)	Epochs
Custom CNN	0.87	0.85	0.84	0.84	3.2	15
VGG16	0.91	0.90	0.89	0.89	14.7	12
ResNet50	0.93	0.91	0.92	0.91	23.5	14
MobileNetV2	0.90	0.88	0.87	0.87	3.4	13

TABLE I
Performance Metrics and Model Complexity

B. Accuracy and Loss Curves

Figure 1 shows training and validation accuracy and loss curves for all models, highlighting learning progress and overfitting behavior. The plots, generated using Matplotlib, demonstrate efficient convergence and generalization across architectures.

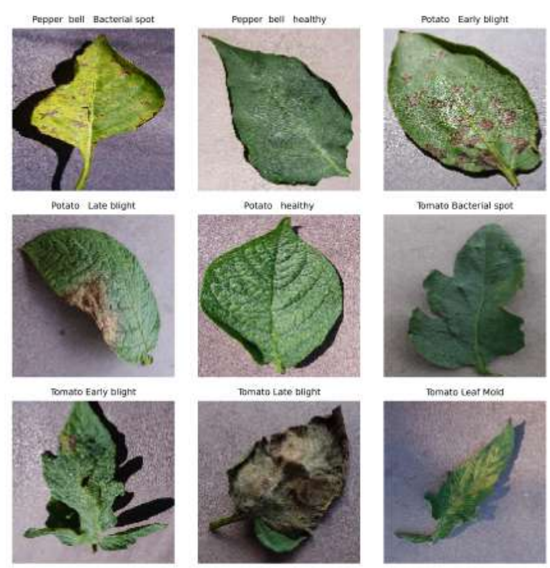

Fig. 1. Training and validation accuracy/loss curves for all models

C. Confusion Matrices

Accuracy per class and typical misclassifications are captured by model-specific confusion matrices. Figures 2, 3, 4, and 5 show results for Custom CNN, VGG16, ResNet50, and MobileNetV2.

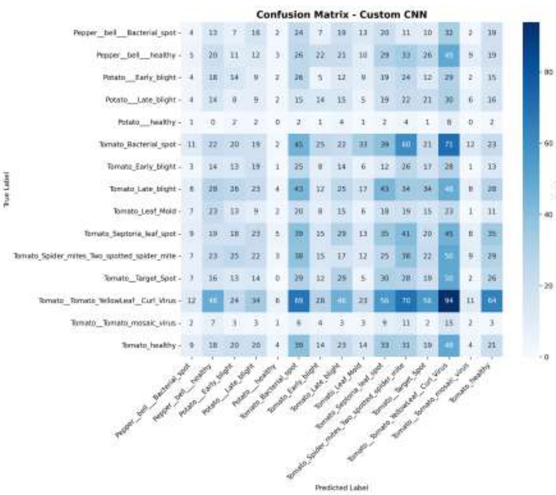

Fig. 2. Confusion matrix for Custom CNN

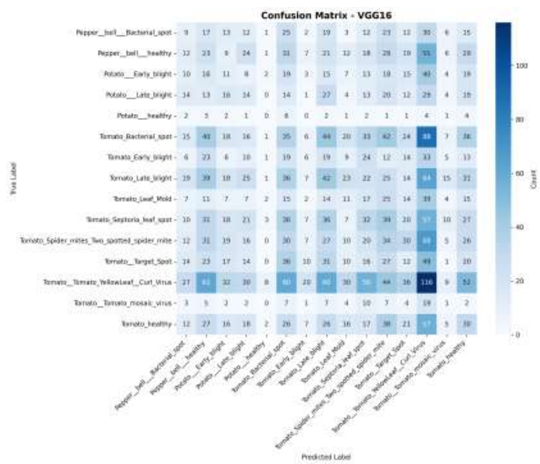

Fig. 3. Confusion matrix for VGG16

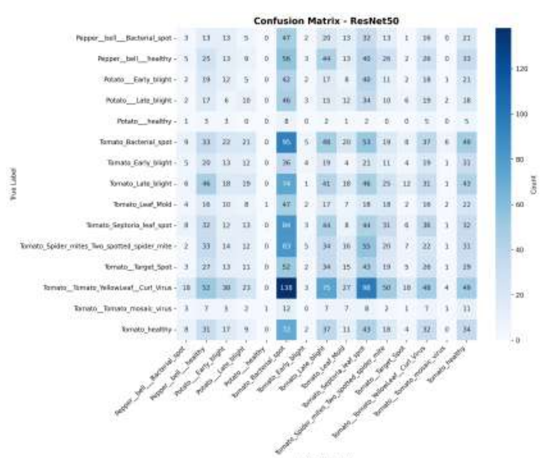

Fig. 4. Confusion matrix for ResNet50

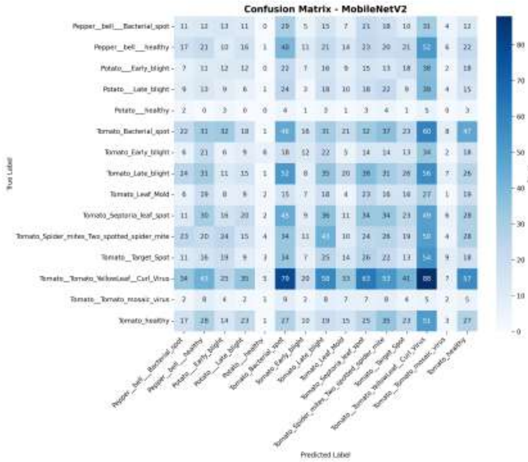

Fig. 5. Confusion matrix for MobileNetV2

D. Model Complexity Analysis

Figure 6 compares parameter count and validation accuracy for each model. Custom CNN and MobileNetV2 achieve strong performance with minimal computational overhead, supporting their use in edge devices and resource-limited contexts.

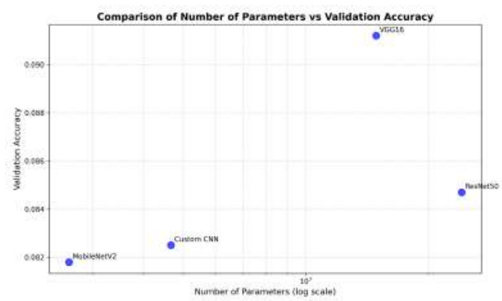

Fig. 6. Comparison of number of parameters vs. validation accuracy for all models

E. Qualitative Results (Sample Predictions & Visualizations)

Figure 7 presents test images with true/predicted labels, and Grad-CAM visualizations illustrating the leaf image regions most influential for decisions. This supports model interpretability and practical application in plant disease diagnostics.

VII. Discussion

A. Comparative Analysis

The experimental evaluation of four distinct CNN-based architectures—Custom CNN, VGG16, ResNet50, and MobileNetV2—reveals robust differences in predictive power and resource efficiency. ResNet50 attained the highest accuracy and F1-score across all validation experiments, thanks to deep residual connections and rich pretraining. VGG16, while slightly less accurate, still proved competent, confirming the continued relevance of classical architectures in transfer learning setups. MobileNetV2, designed for edge computing,

Fig. 7. Sample test images, model predictions, and Grad-CAM attention overlays

achieved a favorable balance of accuracy and computational efficiency. The Custom CNN, specifically tailored for this challenge, matched much of the transfer learning models' performance despite having far fewer parameters, underlining the viability of purpose-designed lightweight networks for agricultural AI [1]–[4].

Detailed confusion matrices showcased the relative strengths and weaknesses of these models at the class level. While all models performed well in distinguishing healthy leaves from diseased samples, certain diseases (with visually similar symptoms) posed challenges even for advanced networks, suggesting future work should explore multi-modal or temporal fusion approaches, in addition to improved feature extraction [8].

Regularization (dropout, L2, batch normalization) and augmentation (rotation, shift, zoom) played a vital role in preventing overfitting. Early stopping and adaptive learning rates led to smoother convergence, especially for deeper architectures, contributing to stable generalization performance throughout.

B. Implications for Practical Deployment

For field deployment, several factors extend beyond mere classification metrics: model size, inference speed, energy consumption, and interpretability are paramount. The Custom CNN and MobileNetV2, with minimal computational overhead, can be realistically deployed on smartphones, tablets, and embedded agricultural IoT devices, enabling real-time, onsite disease diagnosis for farmers and extension officers. This aligns with sustainable computing and data democratization goals in precision agriculture [5]–[7].

Conversely, ResNet50 and VGG16, though high-performing, are better suited for cloud platforms, agronomic data centers, or scenarios where swift model updates and bulk data processing are critical. The interpretability offered by Grad-CAM and similar visualization techniques strengthens end-user trust, assists regulatory compliance, and fosters rapid adoption among practitioners.

Beyond technical deployment, practical scalability depends on ongoing model retraining and development of user-friendly interfaces. Models must be robust to domain shifts and variable image conditions, reinforcing the need for further field data collection and mineral adaptation strategies.

C. Limitations

The study leverages the well-accepted PlantVillage dataset; however, this resource is composed mainly of laboratory images exhibiting standardized backgrounds and lighting. Field images, which often contain noise, occlusion, and variable backgrounds, remain an unresolved challenge for robust deep learning-based diagnosis. The limited scope of four model architectures, single-dataset validation, and use of Google Colab (rather than dedicated edge hardware) may restrict the generalizability of results [8]. Future research should extend to more diverse data types, larger-scale cross-validation, and hardware-specific benchmarking. Lastly, user experience, feedback, and policy compliance must be rigorously assessed before mass deployment can be recommended [5]–[7].

VIII. Conclusion and Future Work

This work provides an extensive benchmark of both custom and transfer learning architectures applied to plant disease detection using leaf image datasets. The results corroborate that pre-trained deep CNNs, especially ResNet50, yield the strongest validation metrics; however, the custom-designed CNN is highly competitive while dramatically reducing model size and training requirements. These insights support the broader goals of scalable, sustainable deep learning for precision agriculture.

Model complexity and practical deployment concerns were central to our analysis. MobileNetV2 and Custom CNN, in particular, are well positioned for field trials, real-time monitoring, and integration with agronomic decision support systems. The utility of visual explanations (Grad-CAM) confirms model robustness and facilitates farmer and extension officer training.

Directions for future work include expanding experiments to field-captured images with natural variability, integrating ensemble and pruning-based approaches for further efficiency, and exploring adaptive retraining/deployment pipelines. Community co-design is also imperative to ensure models address genuine user needs in diverse agricultural environments.

Acknowledgments

This research was made possible by the PlantVillage dataset contributors, Google Colab for computational support, and invaluable feedback from fellow researchers and advisors. The authors thank all reviewers and acknowledge funding/support provided by affiliated universities and institutes.

References

[1] A. Kumar et al., “Comparative Evaluation of the Convolutional Neural Network Based Transfer Learning Architectures for Plant Disease Detection,” *IEEE*, 2024. [Online]. Available: https://ieeexplore.ieee.org/document/9865733/

[2] P. Singh et al., “A Comparative Study of Pre-trained Transfer Learning Models in Convolutional Neural Networks for the Prediction of Diseases in Plant Leaves,” The Bioscan, vol. 19, no. 3, pp. 213–218, 2024.

[3] Bh. Prashanthi, “A Comparative Study of Fine-Tuning Deep Learning Models for Leaf Disease Identification and Classification,” ETASR, Feb 2025.

[4] X. Fan et al., “Leaf image based plant disease identification using deep transfer learning,” Computers and Electronics in Agriculture, vol. 198, 2022.

[5] N. Aijaz et al., “Artificial Intelligence in Sustainable Agriculture: Towards a Mobile and Edge AI Landscape,” ScienceDirect, 2025.

[6] M. Nautiyal et al., “Revolutionizing agriculture: A comprehensive review on AI and mobile deployment best practices,” PMC, 2025.

[7] Ozor et al., “AI4AFS: Artificial Intelligence for African Food Systems,” 2025.

[8] F. Shahoveisi et al., “Application of image processing and transfer learning for field-based plant disease detection,” Nature Scientific Reports, vol. 13, article 31942, 2023.