

MobileResNet: A Generalized Lightweight Model for Accurate Leaf Disease Detection Across Diverse Plant Types

Rikathi Pal*, Somoballi Ghoshal†, Amlan Chakrabarti†

*Department of Computer Science and Automation, Indian Institute of Science, Bangalore, India
rikathi.pal@gmail.com

†A.K. Choudhury School of Information Technology, University of Calcutta, Kolkata, India
somoballi@gmail.com, acakcs@caluniv.ac.in

Abstract—Leaf disease detection is vital for effective and productive agriculture. Although many models exist for detecting leaf diseases, they are often optimized for specific leaf types and diseases. This paper proposes a generalized, simplified, and lightweight two-step model architecture called MobileResNet, which can detect diseases from images of any leaf type. Our model first identifies the plant type from the leaf image and then diagnoses the specific disease affecting it. MobileResNet enhances feature extraction by leveraging both standard convolution and depth-wise separable convolution and incorporating skip connections at strategic points to improve gradient flow and ensure efficient training. Our model performs exceptionally well on 18 varieties of major Indian crops, surpassing eight existing state-of-the-art deep learning architectures. This approach achieves 99.75% accuracy in leaf type identification and over 98% accuracy in identifying leaf disease types across various leading Indian crops, setting a new benchmark in the field.

Index Terms—Depthwise Separable Convolution, Global Average Pooling, Skip Connections, Neural Networks

I. INTRODUCTION

Agriculture sustains the livelihoods of millions and serves as the cornerstone of the nation's economy. However, farmers face numerous challenges, with crop diseases posing a significant threat to fruit and vegetable cultivation. These diseases severely impact marginalized farming communities, particularly those with limited land resources.

Traditional methods of disease identification and treatment are costly, time-consuming, and often require manual inspection and expert intervention. These approaches necessitate a level of expertise that may not be accessible to all farmers, consuming significant time and resources. Consequently, innovative solutions that democratize access to efficient disease detection and treatment techniques are urgently needed.

We present an innovative, lightweight technology tailored for low-powered devices, offering swift and precise results. Our solution merges the strengths of MobileNet and ResNet into a compact, 2-layer architecture, excelling in leaf type identification and disease diagnosis. This architecture, named MobileResNet, outperforms existing methods in efficiency and accuracy, delivering a powerful and timely tool for farmers.

MobileResNet is a novel, streamlined approach that sequentially identifies leaf types and diagnoses diseases with remark-

able precision. As illustrated in Figure 1, this architecture ensures top-tier performance while providing a comprehensive solution for monitoring plant health. These key features support our method's performance:

- **Enhanced Feature Extraction with Hybrid Convolution and Skip Connections:** Our model uses standard and depthwise separable convolutions to enhance feature extraction and reduce complexity, with strategic skip connections to improve gradient flow and training efficiency.
- **Two-Step Classification:** A two-phase classification methodology that first identifies the leaf type followed by the specific disease, leveraging intermediate feature maps for improved accuracy.
- **Lightweight Architecture:** Designed for efficient deployment in resource-constrained environments, ensuring broad accessibility and practical usability in Indian agricultural settings.

II. RELATED WORKS

Recent advancements in deep learning have significantly improved plant disease detection. Agarwal et al. achieved 91.2% accuracy for tomatoes with CNN [1], Shin et al. reached 98.11% for strawberry powdery mildew with ResNet-50 [2], and Tiwari et al. obtained 97.8% for potatoes using VGG19 [3]. Yadav et al. achieved 98.75% for peach bacteriosis with CNN [4], while Rao et al. reported 99% for grapes with AlexNet [5]. Harakannanavar et al. combined CNN and SVM for 99.6% accuracy in tomatoes [6]. Recent studies include a semi-automated system for grape leaf detection with 95.69% accuracy [7] and a CapsNet-based model for apple diseases with 91.37% accuracy [8]. PL-DINO [9], integrating CBAM with ResNet50 and EQL, achieved 70.3% mean average precision, outperforming Faster R-CNN and YOLOv7. Most recent works focus on specific types of leaves, with models that lack generalization. In contrast, our paper proposes a novel two-layer architecture, MobileResNet, designed for seamless detection of both leaf type and diseases, offering a more generalized approach.

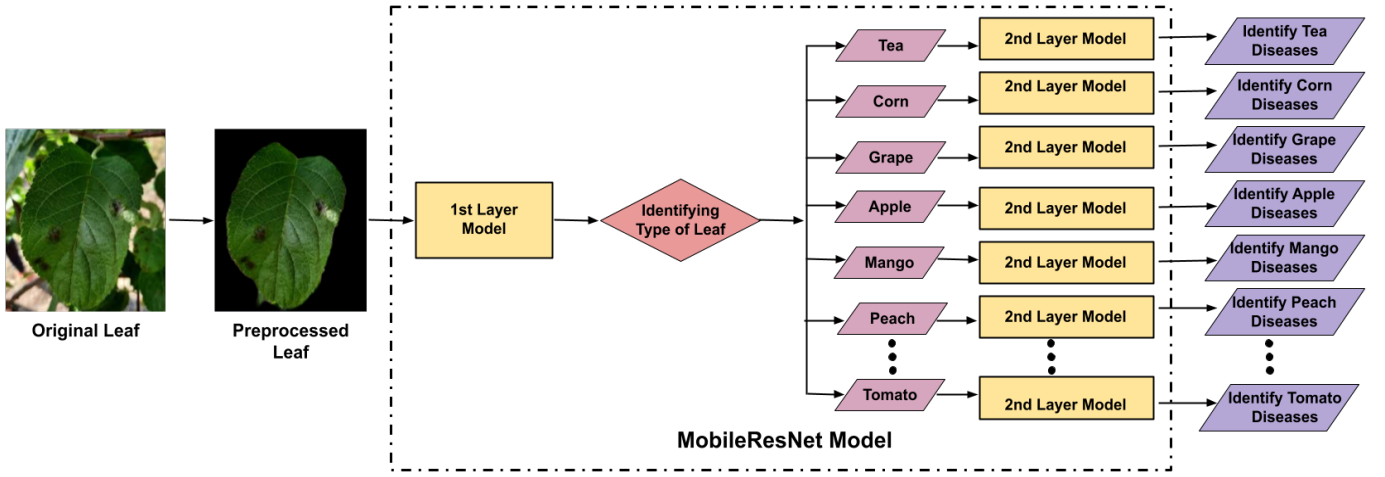


Fig. 1: Steps to identify the leaf type and disease in MobileResNet architecture

III. PROPOSED METHODOLOGY

Our methodology showcases a two-step classification architecture, as in Figure 2. This efficiently identifies leaf types and detects diseases, as illustrated in Figure 1. The process begins with image preprocessing, after which the refined image is fed into our proposed two-layer architecture (Figure 2). This system swiftly identifies the leaf type and diagnoses the disease in minimal time. Our model excels across 18 varieties of Indian crops, outperforming eight existing state-of-the-art methods.

A. Image Preprocessing

We capture the input image, followed by Otsu thresholding [10] for effective background removal and leaf isolation. Following, we apply a shearlet transform [11] for image denoising, which enhances the capture of geometrical features while reducing noise. To further enrich the dataset, we employ data augmentation techniques such as rotation, translation, scaling, and contrast stretching. Finally, the images are resized to a standard 448×448 resolution, ensuring consistent processing throughout the pipeline.

B. First Step Classification

In the initial classification stage, our primary goal is to accurately identify the leaf type and proceed accordingly to identify the disease. We achieve this by inputting a preprocessed image into the proposed model architecture, as shown in Figure 2. Our training dataset comprises a diverse range of leaf types commonly found in Indian agricultural settings, including tomato, strawberry, potato, bell pepper, peach, grape, corn, cherry, apple, mango, guava, rice, and tea. This diverse dataset mentioned in table I, ensures that our model can effectively recognize and differentiate between various leaf types with high precision.

The model architecture processes a preprocessed image of size $448 \times 448 \times 3$ through an initial classification layer. This layer utilizes standard and depthwise separable convolution

(DS Conv). The first 3×3 convolution layer reduces the spatial dimensions to $224 \times 224 \times 16$.

Subsequently, we apply a series of depthwise separable convolutions, each followed by batch normalization and ReLU activation. Specifically, the image is processed through:

- Two DS Conv layers, reducing dimensions to $112 \times 112 \times 32$
- Two DS Conv layers, further reducing to $56 \times 56 \times 64$
- Two DS Conv layers, reducing to $28 \times 28 \times 128$
- Four DS Conv layers, reducing to $14 \times 14 \times 256$
- Four DS Conv layers, reducing to $7 \times 7 \times 512$
- One DS Conv layer, further reducing to $7 \times 7 \times 1024$

To enhance feature preservation and gradient flow, we incorporate skip connections at strategic points. These connections link the output of the initial 3×3 convolution layer to the output of consecutive 3 DS Conv layers and from the combined output of the first skip connection to the output of 2 DS Conv layers. This setup addresses the issue of limited samples for specific leaf diseases by preserving essential features and improving gradient flow. The skip connections allow gradients to bypass intermediate layers, facilitating more efficient training and reducing overfitting. They also enable the reuse of low-level features, such as edges and textures, in later stages of the network.

Following the convolutional layers, we apply global average pooling (GAP) to summarize each feature map into a single value by averaging. GAP reduces the model's parameters, mitigating overfitting and ensuring spatial invariance by summarizing the entire feature map [12]. The pooled features are then fed into a fully connected (FC) layer and, subsequently, a softmax layer to obtain class probabilities for the first-layer classification. This layer expedites disease detection by autonomously identifying the leaf type thus eliminating the need for manual input from farmers. By focusing on the presence of features rather than their exact location, GAP simplifies the model, reduces computational complexity, and enables faster and more efficient processing, crucial for accurate leaf

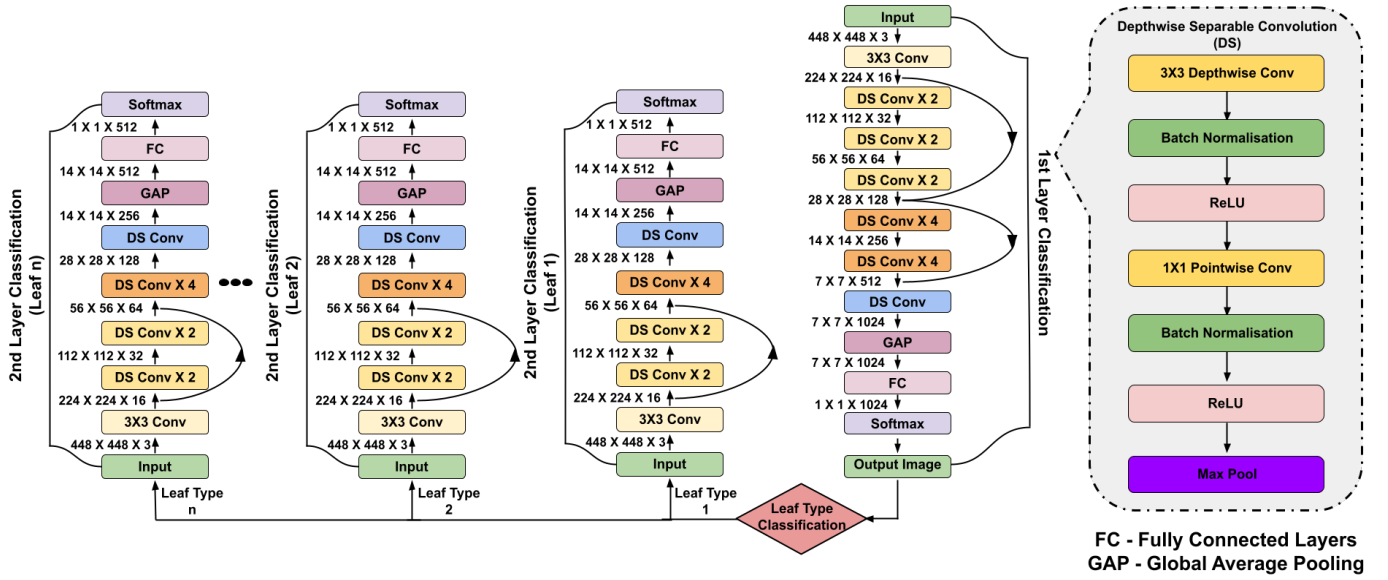


Fig. 2: MobileResNet Architecture

type detection. Once the leaf type is determined, the system transitions to the next classification layer to identify specific diseases affecting the leaf.

C. Second Step Classification

In the second classification phase, our focus shifts to identifying the various diseases affecting different crops. We have assembled a comprehensive catalog of diseases for each crop, ensuring thorough coverage of potential diseases. As detailed in Table I, the dataset includes multiple classes of leaves.

Following the initial identification of the plant type, our system advances to the disease classification stage for each plant. In this phase, intermediate feature maps from various stages of the first-layer classification are reused and processed further. Specifically:

- Features from the $56 \times 56 \times 64$ stage are processed through two additional DS Conv layers.
- Features from the $112 \times 112 \times 32$ stage are processed through four additional DS Conv layers.
- Features from the $224 \times 224 \times 16$ stage are processed through two additional DS Conv layers.

These intermediate outputs are then merged and processed through a 3×3 convolutional layer. During this second classification layer, skip connections are incorporated from the output of the 3×3 convolution to the output of consecutive 2 DS Conv layers. These connections enhance gradient flow and mitigate the vanishing gradient problem, which is crucial given the limited samples available for each leaf disease. By strategically implementing skip connections, we preserve important features and improve gradient flow within the model, ensuring more efficient and effective training.

Global Average Pooling (GAP) is applied after convolutional layers to reduce the number of parameters, making the model lightweight and less prone to overfitting. By averaging

each feature map into a single value, GAP simplifies the classification task while retaining key information about leaf diseases. This reduction in complexity improves generalization on unseen data and enhances translation invariance by focusing on feature presence rather than spatial location. Figure 2 showcases our hybrid architecture, a lightweight and efficient solution ideal for resource-constrained environments. By employing a two-tiered classification approach, it accurately identifies both leaf type and specific diseases. This technology provides farmers with precise crop health insights, enabling timely interventions and proactive management to enhance agricultural productivity.

IV. RESULTS

This section showcases the results obtained by implementing our proposed approach. Additionally, we have conducted a comparative analysis of the approach with the current state-of-the-art methodologies.

A. Dataset Description

In our research, we utilized a variety of online datasets and expanded our study to include real-life datasets from diverse geographic locations, detailed in Table I. This comprehensive approach allowed us to gain insights into various region-specific diseases. We also augmented these datasets using techniques like rotation, scaling, and sharpening to enhance their suitability for training and ensure robust analysis.

B. Experimental Configuration and Parameterization

Performance analysis and comparative studies were conducted on a high-performance platform featuring an Intel i5 processor (410 GHz) with 16 GB of RAM, and an NVIDIA Tesla K80 GPU with 12 GB of VRAM. The GPU runtime included an Intel Xeon CPU @ 2.20 GHz, 13 GB of RAM, and 12 GB of GDDR5 VRAM.

TABLE I: Dataset Description

Plant Type	Samples from Online Source	Samples from Geographic Areas	No. of Classes
Tomato [13]	18345	1341	10
Strawberry [13]	3598	98	2
Potato [13]	5702	402	3
Bell Pepper [13]	3901	490	2
Peach [13]	3896	326	2
Grape [13]	7222	248	4
Corn [13]	5457	368	4
Cherry [13]	3509	412	2
Apple [13]	7779	369	4
Mango [14]	2500	1246	8
Guava [15]	1527	354	5
Tea [16]	1885	150	8
Blueberry [13]	1896	196	2
Orange [13]	2010	210	2
Raspberry [13]	1781	181	3
Soybean [13]	2022	122	2
Squash [13]	1736	186	2
Rice [17]	1963	506	4

Our model was trained for 100 epochs with dynamic learning rate annealing, reducing the rate by a factor of 0.1 every 25 epochs. A larger initial learning rate improved generalization, as shown in Figure 3. We used a batch size of 32 for optimal efficiency and convergence. Sparse categorical cross-entropy was chosen as the loss function for better memory efficiency and faster training with integer labels.

C. Performance Evaluation Metrics

Class accuracy, also known as classification accuracy, measures the proportion of correctly classified instances out of the total instances, which is useful for assessing classifier performance across different categories. For a class C_i , it is calculated as:

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

where TN is true negative, TP is true positive, FP is false positive, and FN is false negative. The numerator represents correctly classified instances, while the denominator includes all cases.

Precision quantifies the proportion of correctly identified positive instances out of all predicted positives:

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

Recall measures the proportion of correctly identified positives out of all actual positives:

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

These metrics complement class accuracy by providing insights into model performance concerning false positives and false negatives.

D. Ablation Study

A key question about our approach is why we employ a two-layer classification system instead of classifying diseases directly. The answer is that many leaf diseases have overlapping features, making accurate classification challenging when diseases are identified in isolation for all plant types. Our method involves first identifying the crop, followed by disease classification specific to that crop. This context-specific approach significantly boosts accuracy—achieving 99.31% on average across all crops compared to 92.8% using only the second layer, as shown in Table II.

Moreover, this two-step system enhances efficiency. It reduces both training time and memory usage, resulting in a lighter, faster model that requires less computational power. This streamlined approach not only improves accuracy but also optimizes performance, making our model more practical for real-world use.

TABLE II: Ablation Study on the Use of Different Steps in MobileResNet Architecture

MobileResNet Steps	Accuracy (%)	Precision (%)	Recall (%)
Only 2nd Step	92.8	93.0	92.5
1st + 2nd Step	99.31	98.65	98.30

E. Comparative Study

The loss curve in Figure 4 demonstrates a steady decline in both training and validation losses, highlighting the model's consistent learning progress. The accuracy, precision, and recall for leaf-type identification remain high, as shown in Table III. Our MobileResNet architecture's first layer outperforms eight state-of-the-art techniques listed in the same table, showcasing its effectiveness in identifying leaf types.

TABLE III: Model Comparison: Accuracy, Precision, and Recall for Leaf Type Classification

Model	Accuracy (%)	Precision (%)	Recall (%)
AlexNet [18]	97.00	97.20	96.80
Inception ResNet [19]	95.50	95.70	95.30
MobileNet [20]	97.46	97.50	97.40
VGG 16 [21]	98.25	98.30	98.20
VGG 19 [21]	97.75	97.80	97.70
XceptionNet [22]	96.00	96.10	95.90
ZFNet [23]	96.50	96.60	96.40
ResNet [24]	97.64	97.70	97.60
MobileResNet	99.75	98.80	99

We evaluated our approach using various datasets and models for leaf disease identification, as detailed in Table IV. Our MobileResNet model consistently outperformed others across all crop types. The loss curves in Figure 5 show smooth trajectories, indicating effective convergence with stable training and efficient learning. This smoothness suggests strong generalization capabilities, confirming that our hyperparameters are well-chosen. Overall, our MobileResNet model demonstrates superior performance in both leaf type and disease detection, validating the robustness of our preprocessing pipeline and hyperparameter selection.

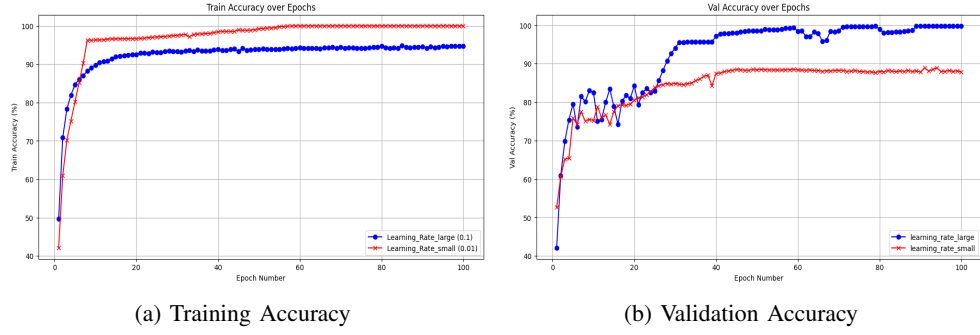


Fig. 3: Training and validation accuracy curves for MobileResNet with first step classification: Comparison of small vs. large learning rates

TABLE IV: Accuracy of different models on different datasets after 2nd layer classification

Crop Name	Alex Net (%) [18]	Inception ResNet (%) [19]	VGG 16 (%) [21]	VGG 19 (%) [21]	Xception (%) [22]	ZF Net (%) [23]	MobileNet (%) [20]	ResNet (%) [24]	MobileRes Net (%)
Tomato	94.63	90.08	92.86	88.38	90.03	98.67	98.72	97.63	98.85
Strawberry	97.83	99.83	97.23	95.64	96.58	99.24	99.5	98.34	99.84
Potato	98.56	98.64	97.48	97.02	98.11	98.56	98.64	96.25	99.82
Bell Pepper	94.35	90.56	95.48	92.14	93.78	96.84	97.63	95.42	99.7
Peach	98.4	97.6	96.4	94.3	95.4	96.8	97.84	97.32	99.7
Grape	99.5	98.46	99	97.15	97.19	99.25	98.24	95.55	99.89
Corn	98.51	98.61	98.63	97.14	95.89	99.68	98.25	97.63	99.25
Cherry	98.89	99.58	99.87	99.45	99.65	99.43	99.5	97.6	99.11
Apple	99.54	98.51	97.56	95.24	92.12	94.56	96	94.65	98.75
Mango	87.67	93.84	96.54	94.58	95.42	97.64	95.4	95.55	98.25
Guava	94.57	95.47	96.12	95.35	92.89	96.14	96	93	98
Tea	93.35	94.86	95.24	96.15	95.84	97.23	96.18	64.7	99.5
Blueberry	97.69	98.45	98.67	99.45	98.65	97.33	97.5	97.4	99.8
Orange	99.14	98	97.36	95.24	90.2	95.61	95.33	94.75	98.75
Raspberry	90.67	93.26	91.25	94.58	93	96.56	95.46	94.55	99.25
Soybean	95	95.18	96.15	97.25	91.79	95.14	94	93.55	99.56
Squash	94.5	95.72	96.25	97.19	96.88	98.65	97.88	89.87	99.6
Rice	92.8	94.56	90.16	95.00	93.55	94.55	96.58	90.64	99.86

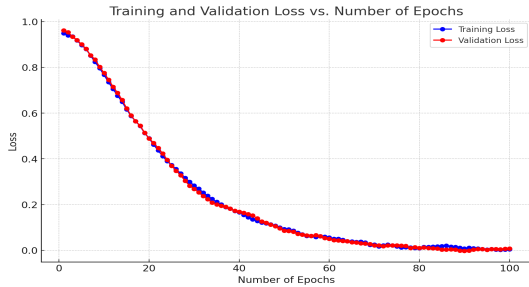


Fig. 4: Training and validation loss curve for leaf type identification

Figure 6 illustrates the performance of our MobileResNet model, showcasing its effectiveness in detecting leaf types and diseases across various leaves from diverse geographic locations. Based on the data presented in Table IV, it is evident that our model consistently achieves accuracy rates exceeding 98% across all crops. This remarkable consistency underscores the reliability of our model's predictive capabilities.

V. CONCLUSION AND FUTURE WORKS

We have successfully developed a lightweight, two-layer MobileResNet architecture that excels in leaf type classification and plant disease detection. Our model achieved 99.75% accuracy in classifying 18 different varieties of Indian crops and over 98% accuracy in detecting diseases for each crop. This performance not only highlights the effectiveness of our approach but also surpasses eight existing state-of-the-art methods.

In our future work, we aim to extend the impact of our research by developing a user-friendly application for farmers. This app will integrate our advanced detection model to provide real-time analysis and actionable insights. By offering practical recommendations and enhancing disease detection capabilities, the app will serve as a valuable tool for farmers, aiding in efficient plant management and maintaining optimal crop health. This initiative will significantly contribute to improving agricultural practices and crop yield management.

REFERENCES

- [1] M. Agarwal, A. Singh, and et al., "ToLeD: Tomato leaf disease detection using convolution neural network," *Procedia Computer Science*, vol. 167, pp. 293–301, 2020.

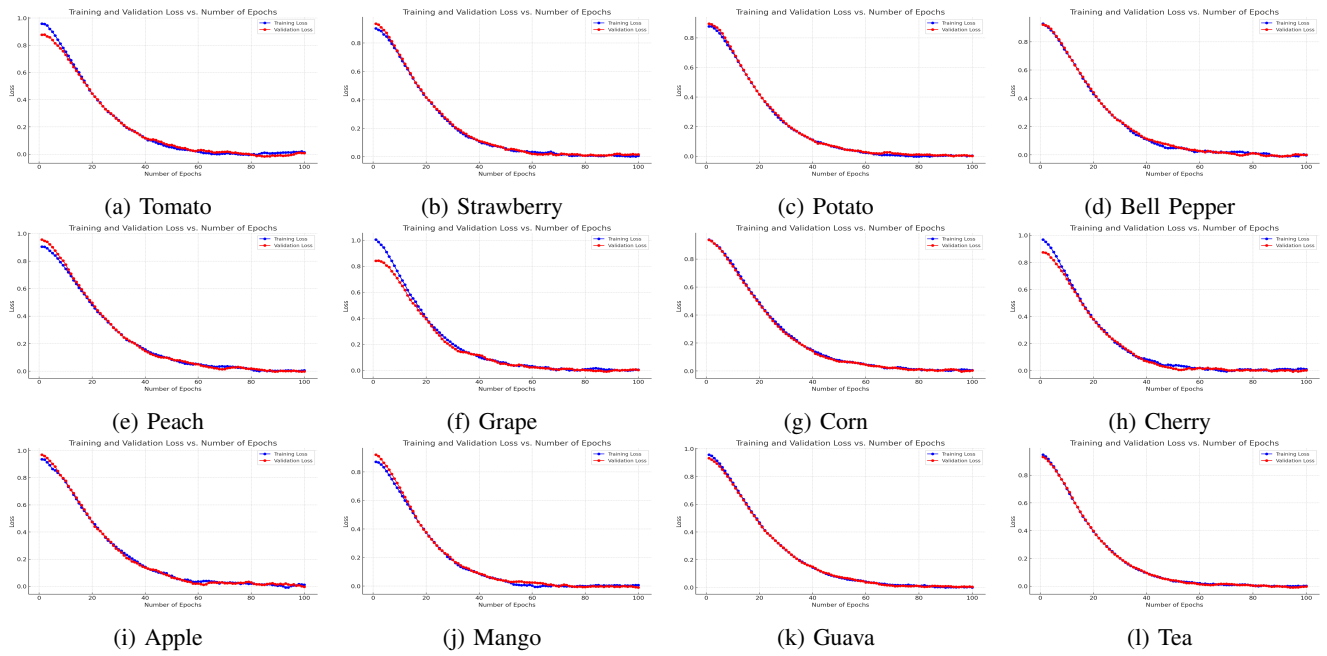


Fig. 5: Training and validation loss curve for leaf disease detection in some crops

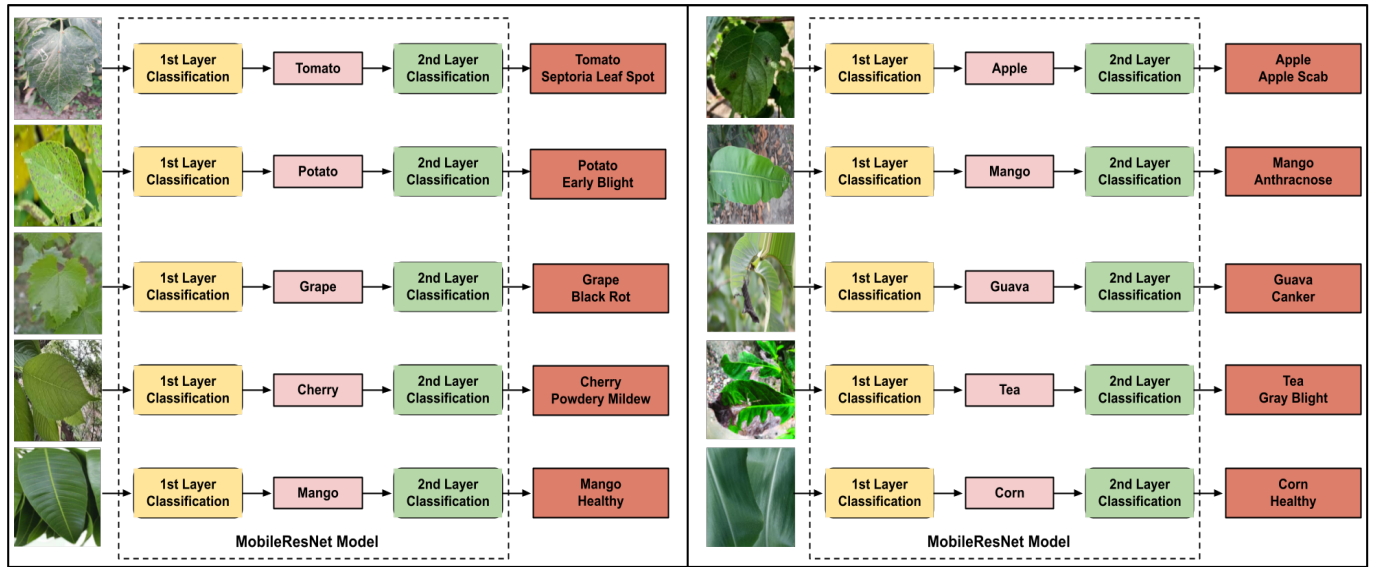


Fig. 6: Real-world leaf type and disease detection using MobileResNet architecture

- [2] C. Shin, Jaemyung and et al., "A deep learning approach for RGB image-based powdery mildew disease detection on strawberry leaves," *Computers and electronics in agriculture*, vol. 183, p. 106042, 2021.
- [3] A. Tiwari, Divyansh and et al., "Potato leaf diseases detection using deep learning," in *International conference on intelligent computing and control systems*. IEEE, 2020, pp. 461–466.
- [4] S. Yadav and et al., "Identification of disease using deep learning and evaluation of bacteriosis in peach leaf," *Ecological Informatics*, vol. 61, p. 101247, 2021.
- [5] U. S. Rao and e. a. Swathi, "Deep learning precision farming: grapes and mango leaf disease detection by transfer learning," *Global Transitions Proceedings*, vol. 2, no. 2, pp. 535–544, 2021.
- [6] S. S. Harakannanavar and et al., "Plant leaf disease detection using computer vision and machine learning algorithms," *Global Transitions Proceedings*, vol. 3, no. 1, pp. 305–310, 2022.
- [7] Kaur and et al., "A novel framework for semi-automated system for grape leaf disease detection," *Multimedia Tools and Applications*, vol. 83, no. 17, pp. 50 733–50 755, 2024.
- [8] Acharya and et al., "Apple foliar leaf disease detection through improved capsule neural network architecture," *Multimedia Tools and Applications*, vol. 83, no. 16, pp. 48 585–48 605, 2024.
- [9] W. Li and et al., "PL-DINO: An Improved Transformer-Based Method for Plant Leaf Disease Detection," *Agriculture*, vol. 14, 2024.
- [10] N. Otsu and et al., "A threshold selection method from gray-level histograms," *Automatica*, vol. 11, no. 285–296, pp. 23–27, 1975.
- [11] G. R. Easley and et al., "Shearlet-based total variation diffusion for denoising," *IEEE Transactions on Image processing*, vol. 18, no. 2, pp. 260–268, 2008.
- [12] R. Pal, A. B. Bhaumik, A. Murmu, S. Hossain, B. Maity, and S. Sen, "A Novel Feature Extraction Model for the Detection of Plant Disease

from Leaf Images in Low Computational Devices,” 2024. [Online]. Available: <https://arxiv.org/abs/2410.01854>

- [13] “New Plant Diseases Dataset,” <https://www.kaggle.com/datasets/vipooool/new-plant-diseases-dataset>, accessed: 2024-07-24.
- [14] “Mango Leaf Disease Dataset,” <https://www.kaggle.com/datasets/aryashah2k/mango-leaf-disease-dataset>, accessed: 2024-07-24.
- [15] “Guava Dataset,” <https://www.kaggle.com/datasets/noamaanabdulazeem/guava-dataset>, accessed: 2024-07-24.
- [16] “Identifying Disease in Tea Leafs,” <https://www.kaggle.com/datasets/shashwatwork/identifying-disease-in-tea-leafs>, accessed: 2024-07-24.
- [17] Prajapati and et al., “Detection and classification of rice plant diseases,” *Intelligent Decision Technologies*, vol. 11, no. 3, 2017.
- [18] A. Krizhevsky and et al., “Imagenet classification with deep convolutional neural networks,” *Advances in neural information processing systems*, vol. 25, 2012.
- [19] C. Szegedy and et al., “Inception-v4, inception-resnet and the impact of residual connections on learning,” in *Proceedings of the AAAI conference on artificial intelligence*, vol. 31, no. 1, 2017.
- [20] A. G. Howard and et al., “Mobilenets: Efficient convolutional neural networks for mobile vision applications,” *arXiv preprint arXiv:1704.04861*, 2017.
- [21] K. Simonyan and et al., “Very deep convolutional networks for large-scale image recognition,” *arXiv preprint arXiv:1409.1556*, 2014.
- [22] Chollet and et al., “Xception: Deep learning with depthwise separable convolutions,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1251–1258.
- [23] M. D. Zeiler and et al., “Visualizing and understanding convolutional networks,” in *Computer Vision–ECCV 2014*. Springer, 2014, pp. 818–833.
- [24] K. He and et al., “Deep residual learning for image recognition,” in *Computer Vision and Pattern recognition (CVPR)*, 2016, pp. 770–778.