International Journal of Innovative Research and Scientific Studies, 8(3) 2025, pages: 4962-4976

IJIRSS International Journal of Innovative Research & Scientific Studies

ISSN: 2617-6548

URL: www.ijirss.com

Comprehensive study of various machine learning models for plant disease detection: Analysis of deep models on tomato plant

Jesal Desai[1*], Amit Ganatra[2]

[1]U & P U Patel Department of Computer Engineering, Chandubhai S. Patel Institute of Technology, Faculty of Technology and Engineering, Charotar University of Science and Technology, Changa, India.
[2]Provost, Parul University (PU), Waghodia 391760, Vadodara, India.

Corresponding author: Jesal Desai (*Email: jaddesai654@gmail.com*)

Abstract

Agriculture is a crucial component for an overpopulated country like India, and thus, plant diseases present a substantial risk to the output of crops, thereby making timely identification and diagnosis crucial for guaranteeing robust economic development. Tomatoes, being a prominent agricultural commodity, are vulnerable to a diverse range of illnesses, which can be detected by observing foliage signs. Automated identification, categorization, and assessment of the severity of plant diseases utilizing Artificial Intelligence have become crucial for improving agricultural productivity. Advancements in machine learning, notably in the application of convolutional neural networks (CNNs), provide potential solutions for precisely detecting and categorizing tomato plant illnesses. These automated solutions minimize the requirement for manual inspection, which is both very labor-intensive and susceptible to errors. This work investigates machine learning methods for detecting plant diseases and provides an evaluation of advanced deep learning approaches used for detecting and classifying tomato plant diseases. MobileNet, ResNet, and DenseNet models exhibited greater performance among the six models examined. To improve the interpretability of deep learning models, Grad-CAM (Gradient-weighted Class Activation Mapping) is utilized. The performance of this method is evaluated using high-performing models such as MobileNet, ResNet, and DenseNet, which are commonly used for plant disease detection.

Keywords: Cotton plant disease detection, Faster R-CNN, hybrid models, localization, ResNet50, YOLOv8.

DOI: 10.53894/ijirss.v8i3.7697
Funding: This study received no specific financial support.
History: Received: 8 April 2025 / **Revised:** 10 May 2025 / **Accepted:** 13 May 2025 / **Published:** 10 June 2025

Competing Interests: The authors declare that they have no competing interests.
Authors' Contributions: Both authors contributed equally to the conception and design of the study. Both authors have read and agreed to the published version of the manuscript.
Transparency: The authors confirm that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.
Publisher: Innovative Research Publishing

1. Introduction

Plant growth depends on environmental conditions, and diseases can reduce crop yield. Early detection is key to minimizing losses and improving productivity. In 2018, agriculture contributed 17-18% to India's GDP, making it crucial for economic stability. AI, deep learning (DL), and machine learning (ML) are transforming agriculture, enabling precise disease detection. Climate change may increase disease risks, making early identification even more essential.

1.1. Advancements in CNN-Based Detection

CNNs have revolutionized plant disease classification by automatically extracting features, outperforming traditional ML models. Since LeNet by LeCun et al. [1], CNN architectures have rapidly evolved. Krizhevsky et al. [2] won the ImageNet Challenge in 2012, marking a major breakthrough. Later models, Zeiler and Fergus [3]; Szegedy et al. [4]; Simonyan and Zisserman [5]; He et al. [6] and Hu et al. [7] reduced classification errors from 25% to just 3%, improving real-world agricultural applications and food security.

1.2. Challenges in Plant Disease Detection

Despite advancements, there are a few challenges in using artificial intelligence. One of them is data limitations. Annotated datasets for diverse plant species are scarce. Environmental variability is one of the significant challenges, as light, temperature, and soil changes affect model accuracy. Overfitting is another challenge for all machine learning and deep learning models. Models may struggle to generalize across species.

Researchers struggle with Computational Constraints as running deep learning models on mobile/edge devices is difficult. Jackulin and Murugavalli [8] highlighted these issues, emphasizing the need for better dataset availability, robustness, and computational efficiency [9-11].

1.3. Motivation for Early Disease Detection

Timely disease detection is crucial for food security, crop health, and economic stability. Early identification helps farmers take preventive measures, reducing damage and pesticide overuse. Advanced detection methods, such as hyperspectral imaging and deep learning, save time and costs, supporting sustainable farming.

Paper structure is as follows: Section 2 reviews eight ML models for plant disease detection. Section 3 discusses publicly available datasets. Section 4 focuses on models and datasets for tomato diseases. Section 5 evaluates six ML models on nine tomato leaf diseases. Section 6 concludes the study.

2. Related Work

Research in plant disease detection includes several ML and DL models:

- CNN: Excels in image-based classification.
- SVM (Support Vector Machine): Strong classifier for structured data.
- Random Forest: Effective for complex datasets.
- KNN (K-Nearest Neighbors): Simple and useful for small datasets.
- YOLO (You Only Look Once): Ideal for real-time detection.
- Transfer Learning: Enhances accuracy using pre-trained models.

These methods help advance plant disease detection, improving agricultural efficiency and sustainability. Table 1 shows a summary of various machine learning-based models for plant disease detection.

Table 1.
Summary of Various Machine Learning based Models for Plant Disease Detection.

1. Plant Disease Detection Model's	2. Description	3. Key Features	4. Advantages	5. Disadvantages
CNN (Convolutional Neural Network)	Deep learning model used for image classification and segmentation.	Hierarchical feature learning, automatic feature extraction.	High accuracy with large datasets, robust to variations.	Requires extensive labeled data, computationally intensive.
SVM (Support Vector Machine)	Classifies data by finding the optimal hyperplane.	Effective in high-dimensional spaces, good for small datasets.	Effective with clear margins of separation.	Not ideal for large-scale image data, requires feature extraction.
Random Forest	Ensemble learning method using multiple decision trees.	Handles large datasets well, reduces overfitting.	Robust, interpretable, handles missing values well.	Can be slow with large datasets and less effective with highly unbalanced data.
KNN (K-Nearest Neighbors)	Classifies data based on proximity to known examples.	Simple and easy to understand, with no training phase.	Effective for smaller datasets, adapts to new data quickly.	Computationally expensive for large datasets, sensitive to noisy data.
YOLO (You Only Look Once)	Real-time object detection model.	Fast detection, real-time processing.	High speed, suitable for real-time applications.	Requires fine-tuning, less accurate for small objects.
Transfer Learning Models	Uses pre-trained models (e.g., VGG, AlexNet) for feature extraction.	Leverages existing models for faster training.	Reduces training time, requires less data.	Might not be as specialized for specific plant diseases.

Source: Alguliyev, et al. [12]; Swain, et al. [13]; Sajitha, et al. [14]; Demilie [15]; Simhadri, et al. [16] and Zeiler and Fergus [3]

2.1. CNN-Based Plant Disease Detection Models

CNN models like ResNet, DenseNet, Inception, VGGNet, AlexNet, MobileNet, Xception, and EfficientNet are widely used for plant disease detection due to their ability to extract features automatically. Table 2 summarizes their features, benefits, and drawbacks to help select the best model Szegedy et al. [4]; Simonyan and Zisserman [5]; Venkataramanan et al. [17]; Kolli et al. [18]; Hammou and Boubaker [19]; Kibriya et al. [20]; Bharali et al. [21]; Matin et al. [22]; Chen et al. [23]; Moid and Mousmi [24]; Saleem et al. [25]; Atila et al. [26]; Rajeena PP et al. [27].Table 2

Nagaraju and Chawla [28] identified key challenges in deep learning for plant disease detection, such as data acquisition, model optimization, feature extraction, overfitting reduction, and loss function improvements. Venkataramanan et al. [17] proposed a CNN model optimized for various plant species and disease types. Chowdhury et al. [29] introduced EfficientNet, which outperformed U-Net models, achieving 99.95% accuracy for binary classification and 99.12% for six-class classification. Anandhakrishnan and Jaisakthi [30] developed a Deep CNN model for tomato leaf disease detection, achieving 98.4% accuracy, outperforming traditional ML methods like SVM and MLP.

Khalid and Karan [31] used CNN and MobileNet for disease detection, achieving 89% and 96% accuracy, respectively, and employed XAI techniques like GradCAM for model interpretation. Chin et al. [32] improved YOLOv8 for plant disease detection, integrating GhostNet and Coordinate Attention, achieving a maximum accuracy of 72.2% with transfer learning.

2.1.1. ResNet-Based Models

ResNet models address vanishing gradient issues using shortcut connections. ResNet50, with 50 layers and Batch Normalization, improves learning. A ResNet34 model achieved 99.40% accuracy on a dataset of 15,200 crop leaf images Kumar et al. [33]. Mukti and Biswas [34] trained ResNet50 on 70,295 images across 38 disease classes, achieving 99.80% accuracy and surpassing VGG16, VGG19, and AlexNet.

2.1.2. DenseNet-Based Models

DenseNet, similar to ResNet, enhances feature propagation by connecting each layer to all subsequent layers. DenseNet121, with four DenseBlocks and transition layers, efficiently classifies plant diseases while using fewer parameters than ResNet50. Too et al. [35] found DenseNet superior for fine-tuning in agriculture, while Chen et al. [36] demonstrated its robustness in disease identification. Amara et al. [37] successfully applied DenseNet to banana leaf disease classification.

Ferentinos [38] evaluated DenseNet on the PlantVillage Dataset & CNN-Based Models. Other studies focused on InceptionV3's disease detection capabilities [39] and deep learning overfitting issues [40].

2.1.3. VGG-16 & VGG-19 Models

VGG-16 and VGG-19 use repeated convolutional layers with max-pooling. The difference lies in their layer depth. Rithik e al. [41] enhanced VGG-19 for tomato leaf disease detection (94% accuracy), while Nguyen et al. [42] combined it with image segmentation, achieving 99.72%. Alatawi et al. [43] applied VGG-16 for plant disease identification with 95.2% precision.

2.1.4. InceptionV3-Based Models

InceptionV3 processes features at multiple scales. Baheti et al. [44] achieved 88.98% training accuracy and 85.80% validation accuracy in tomato disease detection. Dutta et al. [45] used it with image augmentation for early and late blight detection, reaching 98.60% accuracy.

2.1.5. MobileNet-Based Models

MobileNet, optimized for mobile devices, balances accuracy and efficiency. Puranik [46] developed a MobileNetV3-based mango disease detection app with 98% accuracy. Bi et al. [47] applied MobileNet for apple leaf disease detection, offering a lightweight alternative to ResNet152 and InceptionV3.

2.1.6. SVM-Based Models

SVM is widely used for plant disease classification. Das et al. [48] found SVM superior to Random Forest and Logistic Regression. Rajagopal et al. [49] improved SVM with Fuzzy C-means and Particle Swarm Optimization, achieving high accuracy on a 55,400-image dataset. Dubey and Choubey [50] optimized SVM for rice disease detection (97.54% accuracy).

2.1.7. Random Forest-Based Models

Random Forest is effective for handling noisy data. Nancy and Kiran [51] integrated Random Forest with GLCM features for cucumber disease classification (98.62% accuracy). Praba and Krishnaveni. [52] used it for maize disease detection (97% precision), aiding early disease management.

2.1.8. KNN-Based Models

KNN is simple yet effective for disease detection. Kalyan and Rashmita [53] found that CNN (92.48%) outperformed KNN (74.14%) in leaf disease classification. Vaishnnave et al. [54] improved accuracy in groundnut leaf disease detection by replacing SVM with KNN.

2.1.9. YOLO-Based Models

YOLO enables real-time plant disease detection. Shill and Rahman [55] applied YOLOv3 and YOLOv4 for multi-species disease detection, achieving high precision. Another study, Mahesh and Mathew [56], used YOLOv3 for bacterial spot detection in bell peppers (90% accuracy). Wang and Liu [57] developed YOLOv8n-vegetable, improving mAP by 6.46% for greenhouse disease detection.

2.2. Transfer Learning for Plant Disease Detection

Transfer learning enhances model efficiency with limited labeled data by leveraging pre-trained models [36, 58-61]. A study by Chen et al. [36] combined VGGNet with GoogLeNet's Inception module and Swish activation, achieving over 92% accuracy. Mohanty et al. [62] tested AlexNet and GoogLeNet on the PlantVillage dataset, showing 98.21% accuracy even with just 20% training data.

Table 2.
Summary of Various CNN-Models for plant disease detection.

CNN based Plant Disease Detection Model's	**Description**	**Key Features**	**Advantages**	**Disadvantages**	**Use in Plant Disease Detection**
ResNet (Residual Networks)	Deep learning architecture with residual learning to handle vanishing gradients.	Deep networks with residual blocks, good for complex features.	Very accurate, effective for complex and deep networks.	Requires substantial computational resources, complex architecture.	Frequently used due to its strong performance in feature extraction, especially for complex diseases.
DenseNet (Densely Connected Networks)	Network where each layer receives input from all previous layers.	Several variants like DenseNet121, DenseNet169, depending on the number of layers. Improved gradient flow, feature reuse.	Reduces vanishing gradient problem, effective with fewer layers.	Higher computational cost, more complex architecture.	Very effective for plant disease classification due to its efficient learning of features, especially for complex datasets
Inception Network	Network with multiple types of convolutional filters and	22 layers (InceptionV3 is a popular version). Multi-scale feature	Effective for various object sizes, adaptable.	More complex and computationally demanding.	Efficient for large-scale image classification tasks, including plant

	pooling layers.	extraction, flexible architecture.			disease detection.
VGGNet	Deep CNN with a simple and uniform architecture.	16 or 19 layers; uniform 3x3 convolutions; max pooling.	High accuracy, effective feature extraction, straightforward design.	Requires substantial computational resources; can be prone to overfitting.	Popular for achieving high accuracy in plant disease detection.
AlexNet	Early deep CNN model with a simpler structure.	8 layers deep, with 5 convolutional layers and 3 fully connected layers; ReLU activation; dropout; max pooling.	Fast training, effective transfer learning, and reduced computational load.	Shallower than newer models; over-fitting risk; no batch normalization.	Often used for leaf disease classification with decent accuracy and relatively fast computation.
MobileNet	Lightweight CNN for mobile and embedded devices.	Depth-wise separable convolutions; low computational cost.	Efficient for mobile and real-time applications; lower resource usage.	Lower accuracy compared to deeper models; may miss subtle features.	Suitable for real-time and edge-device applications for plant disease detection.
Xception	Deep CNN with depth wise separable convolutions.	Similar depth to Inception networks but more efficient in terms of parameter usage. Depth-wise separable convolutions; high accuracy.	High performance with lower computational cost; effective feature extraction.	Computationally intensive; more complex architecture.	Provides high accuracy with relatively lower computational costs.
EfficientNet	Scalable CNN model balancing depth, width, and resolution.	Varies depending on the Efficient Net version (B0 to B7). Balanced scaling, efficient performance, and compound scaling.	High accuracy, efficient use of computational resources, state-of-the-art performance.	Complex scaling factors require careful tuning.	Combines high accuracy with efficient use of computational resources.
Custom CNN	Tailored CNN architecture for specific tasks.	Designed to fit specific needs; can be optimized for particular datasets.	Flexibility in architecture can be fine-tuned for specific tasks.	Design and tuning complexity; performance may vary widely.	Designed specifically for tasks like classifying diseases in crops such as tomatoes, potatoes, and wheat.

Source: Kolli, et al. [18]; Hammou and Boubaker [19]; Kibriya, et al. [20]; Bharali, et al. [21]; Matin, et al. [22]; Chen, et al. [23]; Moid and Mousmi [24]; Saleem, et al. [25]; Atila, et al. [26] and Rajeena PP, et al. [27]

3. Public Datasets for Plant Disease Detection

Several datasets for plant disease detection are investigated in this research. While some researchers choose to use publicly available datasets, others opt to take pictures in actual agricultural settings. To tackle any research problem, the very first thing you need is a complete, noise-free dataset that is well-labeled. It can be challenging to find a dataset that fits the needs of a study, yet doing so is crucial for reliable results. Here in Table 3, we showcase some of the most well-known open datasets that are widely utilized in plant disease detection research and are readily available. Listed here are some of the most popular public datasets for plant disease detection, along with details about them, including the datasets' sources, image resolution, number of images, plant species, disease categories, and number of images.

Table 3.
Description of public datasets available for plant diseases.

Ref	Year	Dataset	No. of Images	Number of Classes	Resolution	Species Included	Diseases Included	Application
Hughes and Salathé [63]	2015	Plant Village Dataset	~54,309	38 (plant species/disease categories)	Variable (~256x256 pixels)	Tomato, Potato, Apple, Maize, Grape, etc.	Powdery mildew, early blight, late blight, leaf mold, mosaic virus, etc.	This dataset has been extensively used with various CNN architectures, including InceptionV3, for fine-tuning models in plant disease detection.
Kaggle [64]	2020	Kaggle Plant Pathology 2020 Dataset	3,651	4 (healthy, rust, scab, multiple diseases)	High-resolution images	Apple leaves	Apple scab, rust, multiple diseases	These datasets are excellent for testing Inception models in distinguishing subtle leaf symptoms associated with different diseases.
Kaggle [65]	2021	Kaggle Plant Pathology 2021 Dataset	2,739 training images	12	High-resolution images	Apple leaves	Apple scab, rust, powdery mildew, cedar apple rust, and more	
AI Challenger [66]	2018	Agricultural Disease Images Dataset (AI Challenger)	~27,000	61 diseases across 10 crop species	400x400 pixels (variable)	Apple, maize, tomato, grape, and others	Rust, blight, leaf spot, mosaic, etc.	Suitable for Inception-based architectures to learn multi-scale features from a diverse range of crops and diseases.
UCI Machine Learning Repository [67]	2021	Leaf Disease Dataset (UCI Machine Learning Repository)	340 images	30 species	1024x768 pixels (variable)	Various species of plants (e.g., Acer, Quercus)	Not specifically disease-oriented but useful for species recognition	Though designed for species classification, this dataset can be adapted for disease detection in plant leaves using Inception-based models.
Computer Vision Problems in Plant Phenotyping (CVPPP) [68]	2014	CVPPP Leaf Dataset (Plant Phenotyping)	~2,000	N/A (for leaf segmentation)	High resolution	Various plant species	N/A (not disease-specific)	Inception models can benefit from the ability to first segment leaves from the background before applying disease detection algorithms.
Kaggle [69]	2020	Indian Leaf Dataset	~4,500 images	Multiple (healthy and diseased categories)	Variable	Various crops (paddy, maize, chili, etc.)	Leaf blight, leaf spot, powdery mildew	Can be used for training InceptionV3 models to identify disease-specific patterns in crops common in Indian agriculture.
UCI Machine Learning Repository [70]	2017	The Plant Seedlings Dataset	9,000 images.	12 classes, each representing a different plant species or seedling type.	High-resolution images (e.g., 256x256 pixels or higher)	Corn, Bean, Potato, Tomato, and others.	Classifying seedlings into different plant species, not for detecting or classifying diseases.	Useful for pre-training models like Inception for plant classification, and then fine-tuning for disease detection tasks.
Leminen Madsen, et al. [71]	2020	Open Plant Phenotyping Database (OPP)	7,590	Does not categorize data into specific classes. It provides comprehensive phenotyping data across various plant traits and conditions.	High-resolution images	47 different species.	Focuses on plant phenotyping and does not specifically categorize diseases. It provides detailed data on plant growth and development rather than on disease classification.	Ideal for detecting diseases or other plant stresses using transfer learning with Inception-based models.

Singh, et al. [72]	2019	PlantDoc	2,598	27 classes (17-10, disease-healthy)	Low resolution	13 species	17 different Diseases	Designed for plant disease detection in real-world conditions
Parraga-Alava, et al. [73]	2019	RoCoLe (Robusta coffee leaf images dataset)	1560	Six classes	High-resolution images	390 coffee plants	Red spider mite presence, rust level 1, rust level 2, rust level 3 and rust level 4	Used to train and validate the performance of machine learning algorithms used in binary and multiclass classification problems as well as in segmentation tasks specially related to plant diseases recognition
Rauf, et al. [74]	2019	Citrus Dataset	759	For Citrus fruits (Black Spot, Canker, Greening, Scab, and healthy with total number of 150 images) For Citrus Leaves (Black Spot, Canker, Greening, Melanose, and healthy with total number of 609 image)	256x256 pixels	-	Blackspot, Canker, Scab, Greening, and Melanose.	usable for the researchers to prevent plants from diseases using advanced computer vision techniques

4. Plant Disease Datasets and Tomato Disease Detection

The PlantVillage Dataset is a widely used resource for plant disease detection, containing over 54,000 images of healthy and diseased leaves from 38 plant species, including tomatoes, potatoes, and maize. The Kaggle [64] and Kaggle [65] focus on apple orchard diseases like rust and apple scab. The AI Challenger Agricultural Disease Photos Dataset includes 27,000 images of 10 crops and 61 diseases, aiding classification and detection tasks. Other datasets serve specific needs:

- UCI Machine Learning Repository – Used for plant species classification.
- CASA Wheat Disease Dataset – Focuses on wheat disease detection.
- CVPPP Leaf Dataset – Originally for leaf segmentation but also useful for disease diagnosis.
- Image Database of Indian Leaf Diseases – Covers rice, maize, and chili diseases like leaf spot and blight.
- Plant Seedlings Dataset – Tracks plant growth stages to detect abnormalities.
- Open Plant Phenotyping Database (OPP) – Contains plant health condition images for disease classification.
- Flavia Leaf Dataset – Includes 32 plant species for classification.
- Banana Leaf Disease Dataset – Dedicated to banana plant diseases.
- Indian Leaf Dataset – Features plant diseases common in Indian agriculture.

These datasets are publicly available on platforms like Kaggle, UCI, and Mendeley and are widely used to train CNNs, transfer learning models, and deep learning classifiers for plant disease detection.

Machine learning and deep learning have been widely used for tomato disease detection. Several models have achieved high accuracy using the PlantVillage dataset:

Mohanty et al. [62] trained a CNN model to classify 10 tomato diseases, achieving 99.35% accuracy.

Durmus et al. [75] developed a ResNet-50 model, solving the vanishing gradient issue with shortcut connections, achieving 98.42% accuracy.

Brahimi et al. [76] fine-tuned a VGG16 model using transfer learning, reaching 98.78% accuracy with minimal training effort.

Too et al. [35] implemented Inception-v3, using factorized convolutions to extract complex features, achieving 97.78% accuracy.

Ramcharan et al. [77] introduced MobileNet, optimized for mobile devices, ensuring real-time detection with efficiency.

Jiang et al. [78] developed a DenseNet-121 model, addressing the vanishing gradient problem and improving feature reuse, achieving 98.90% accuracy.

Nisha et al. [79] applied SVM classifiers with hand-crafted features, achieving 87.24% accuracy.

Arsenovic et al. [40] used Capsule Networks, preserving spatial relationships between features, achieving 99.12% accuracy.

Thangaraj et al. [80] applied transfer learning with VGG16-based CNN models to classify tomato diseases.

A summary of machine learning models used in above mentioned papers for tomato disease detection is presented in Table 4.

Table 4.
Machine Learning based Models for Tomato Plant Disease Detection

Year	Machine Learning Model	Approach	Dataset			Accuracy
			Name	No. of Images	No. of Classes	
2016	Convolutional Neural Networks (CNN)	CNN model with multiple layers for feature extraction and disease classification.	PlantVillage (Tomato subset)	18,160	10 (9 diseases, healthy)	99.35%
2018	Transfer Learning (VGG16)	Pre-trained VGG16 model fine-tuned for tomato disease detection.	PlantVillage (Tomato subset)	18,160	10 (9 diseases, healthy)	98.78%
2017	ResNet-50 (Deep Residual Networks)	Deep residual learning model (ResNet-50) for tomato leaf disease classification.	PlantVillage (Tomato subset)	18,160	10 (including diseases like early blight, late blight, and bacterial spot)	98.42%
2019	Inception-v3	Inception-v3 deep learning model, optimized for tomato disease classification.	PlantVillage (Tomato subset)	~18,160	10 (healthy, various disease types)	97.78%
2019	MobileNet	MobileNet model optimized for mobile applications and embedded systems for tomato disease classification.	PlantVillage (Tomato subset)	~18,160	10 (healthy, bacterial spot, late blight, early blight, etc.)	96.12%
2021	DenseNet-121	DenseNet-121 model used for plant disease classification.	PlantVillage	~18,160		98.90%

			(Tomato subset)			
2020	Support Vector Machine (SVM)	SVM with hand-engineered features such as color and texture for tomato disease detection.	Agricultural Disease Dataset (AI Challenger)	~1,000 (Tomato subset from AI Challenger)	3 (tomato healthy, tomato early blight, tomato late blight)	87.24%
2019	Capsule Networks	Capsule Networks for classifying tomato plant diseases.	PlantVillage (Tomato subset)	~18,160		99.12%
2021	Transfer learning techniques with a deep CNN architecture	Pre-trained models like VGG16 to identify tomato leaf diseases	PlantVillage dataset	3,000 images of tomato leaves	10	99.34%.

4.1. Datasets for Tomato Plant Disease Detection

PlantVillage Dataset (Tomato Subset), a widely used dataset for plant disease detection, with a large collection of tomato plant images. It contains images of healthy and diseased tomato leaves. Kaggle [64] (Tomato Subset). Released as part of the FGVC7 competition, this dataset contains images of tomato leaves with healthy and diseased leaves.

The Agricultural Disease Images Dataset (AI Challenger) includes labeled images of various crop diseases, including diseases affecting tomato plants. The Tomato Disease Dataset (Mendeley) is a smaller dataset hosted on Mendeley Data, focusing on leaf disease detection. The UCI Machine Learning Repository (Leaf Dataset - Tomato) contains plant leaf images and associated disease information. Although it covers multiple species, tomato leaf images can be extracted.

Table 5.
Summary of Public Datasets for Tomato Plant Disease Detection.

Dataset	Number of Images	Classes	Species	Disease	Availability
PlantVillage (Tomato Subset) Hughes and Salathé [63]	~18,160	10 (9 diseases, healthy)	Tomato	• Bacterial spot • Early blight • Late blight • Leaf mold • Septoria leaf spot • Spider mites (two-spotted spider mite) • Target spot • Mosaic virus • Yellow leaf curl virus • Healthy	PlantVillage on Kaggle
Kaggle Plant Pathology 2020 (Tomato) Kaggle [64]	3,651	4	Tomato (subset)	General disease classes (can include rust, scab, multiple diseases)	Kaggle Plant Pathology 2020
Agricultural Disease (AI Challenger) AI Challenger [66]	~27,000	61 diseases (multiple)	Tomato (subset)	Early blight, late blight, bacterial spot, mosaic virus	AI Challenger Dataset
Tomato Disease Dataset (Mendeley) UCI Machine Learning Repository [81]	~4,000	Multiple	Tomato	Early blight, late blight, bacterial spot, and more	Mendeley Tomato Dataset
UCI Machine Learning Leaf Dataset Lamrahi [82]	340	Multiple	Tomato	General disease information for leaf classification	UCI Leaf Dataset

These datasets are widely used in research and application development for tomato plant disease detection using machine learning and deep learning models. Most datasets are publicly available through Kaggle or Mendeley, with some field-collected data available through collaborations or institutional requests.

5. Evaluating ML Models for Tomato Plant Disease Detection

5.1. Disease Classification

For this research, we utilized the publicly available "New Plant Diseases Dataset (Augmented)," commonly referred to as the "Tomato Dataset," which was provided on Kaggle [83]. This dataset includes images of tomato leaves categorized into ten classes: nine for various diseases and one for healthy leaves. The diseases represented in the dataset are Bacterial Spot, Early Blight, Late Blight, Leaf Mould, Septoria Leaf Spot, Spider Mites (Two-Spotted Spider Mite), Target Spot, Tomato Yellow Leaf Curl Virus, and Tomato Mosaic Virus.

The dataset was augmented for this study and split into training and validation sets, comprising 18,345 and 4,585 images, respectively. The accompanying table illustrates the original distribution of images across each class and confirms the balance among classes. Table 6 provides a detailed breakdown of the number of training and validation images for each disease category

The images were resized to dimensions of 224 x 224 pixels. To enhance the dataset, we applied random zooming with a margin of up to 20% of the original size and performed horizontal flipping on some images. The validation set was employed to assess the performance of the six models trained on the training data.

Table 6.
No. of Training and Validation images for each disease type.

Class	No. of Training Images	No. of Validation Images
Bacterial Spot	1702	425
Early Blight	1920	480
Late blight	1851	463
Leaf Mould	1882	470
Septoria leaf spot	1745	436
Spider Mites - Two-spotted spider mite	1741	435
Target Spot	1827	457
Tomato Yellow Leaf Curl Virus	1961	490
Tomato Mosaic Virus	1790	448
Healthy	1926	481

The study categorizes tomato leaf diseases using specific notations for easier classification and analysis. D1 represents Tomato Bacterial Spot, while D2 and D3 correspond to Tomato Early Blight and Tomato Late Blight, respectively. Tomato Leaf Mold is labeled as D4, and Tomato Septoria Leaf Spot as D5. D6 denotes Tomato Spider Mites (Two-Spotted Spider Mite), whereas D7 refers to Tomato Target Spot. Viral infections are classified as D8 for Tomato Yellow Leaf Curl Virus and D9 for Tomato Mosaic Virus. Additionally, healthy tomato plants are marked as "Healthy" to differentiate them from diseased samples. These notations help streamline model evaluation and disease identification.

The InceptionNet model took the longest to train and had the highest misclassification rate. Its performance was similar to VGG-16 and VGG-19, with VGG-16 slightly outperforming VGG-19, likely due to its shallower depth reducing the vanishing gradient issue. The top three models were MobileNet, DenseNet121, and ResNet50.

As shown in Figure 1 all models excelled at detecting Tomato Yellow Leaf Curl Virus (D8), achieving an average F1-score of 0.962, the highest among all disease classes. They also performed well in identifying healthy leaves (F1-score: 0.945) and Tomato Bacterial Spot (D1) (F1-score: 0.94).

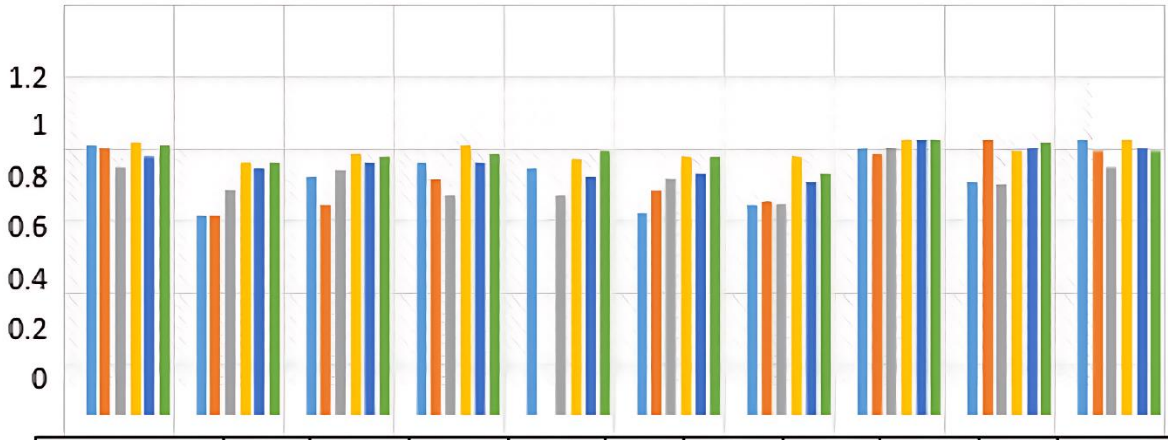
Figure 2 highlights overfitting and underfitting patterns. InceptionNet V3 showed underfitting, with low training and validation accuracy. VGG-19 displayed overfitting, with a significant gap between training and validation accuracy. VGG-16 and ResNet50 also showed some overfitting, but within an acceptable range. MobileNet trained the fastest, completing in 36 minutes.

The hardest diseases to classify were Tomato Early Blight (D2) and Tomato Target Spot (D7), with F1 scores of 0.816 and 0.811, respectively. These results indicate that some diseases are more challenging to detect, suggesting that more image data may improve model accuracy.

Table 7.
Summary of Machine Learning Models on Plant Disease Detection.

Machine Learning Model	No. of Layers	Time Taken (CPU time)	Training Accuracy	Validation Accuracy	(Correct, Misclassified)
VGG-16	16	39 min 41s	93.94%	89.79%	(3870, 715)
VGG-19	19	41 min 56s	92.03%	86.56%	(3969, 616)
MobileNet	30	36 min 8s	96.56%	94.29%	(4323, 262)
DenseNet-121	121	45 min 3s	95.15%	93.39%	(4281, 304)
ResNet-50	50	41 min 28s	94.75%	90.36%	(4143, 442)
InceptionNetV3	48	50 min 13s	89.01%	83.38%	(3823, 762)

As illustrated in Table 7, all models, with the exception of InceptionV3, exhibit a tendency to misclassify Disease D6 more frequently than other classes. Even the top-performing model, MobileNet, demonstrates its highest rate of misclassification with the Spider Mite disease. Furthermore, MobileNet, VGG-16, and VGG-19 tend to misclassify several diseases as Disease D2. Likewise, ResNet50 and DenseNet121 show the highest number of misclassifications for Disease D7.



	D1	D2	D3	D4	D5	D6	D7	D8	D9	Healthy
VGG-16	0.96	0.71	0.85	0.9	0.88	0.72	0.75	0.95	0.83	0.98
VGG-19	0.95	0.71	0.75	0.84	0	0.8	0.76	0.93	0.98	0.94
InceptionNet	0.88	0.8	0.87	0.78	0.78	0.84	0.75	0.95	0.82	0.88
MobileNet	0.97	0.9	0.93	0.96	0.91	0.92	0.92	0.98	0.94	0.98
ResNet-50	0.92	0.88	0.9	0.9	0.85	0.86	0.83	0.98	0.95	0.95
DenseNet-121	0.96	0.9	0.92	0.93	0.94	0.92	0.86	0.98	0.97	0.94

Figure 1.
Comparison of F1-Score for Different Model.

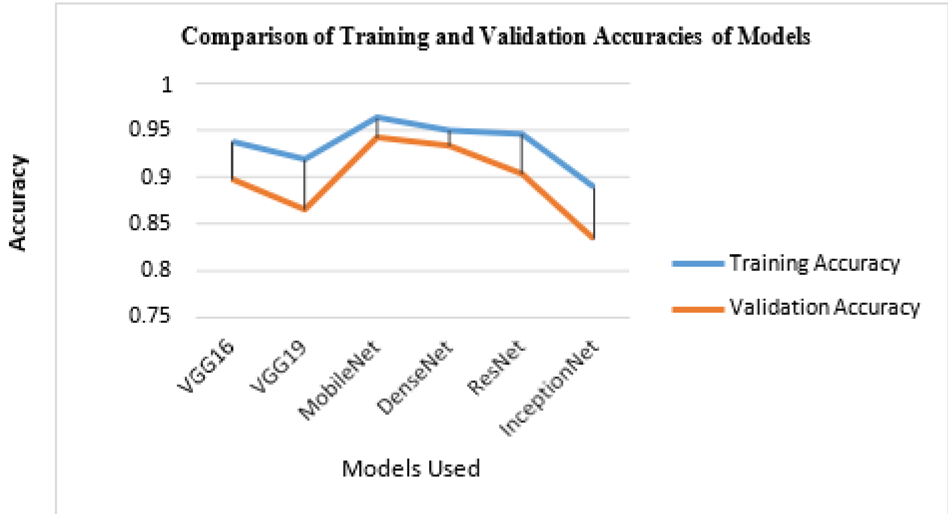


Figure 2.
Comparison of Training and Validation Accuracies of the Models used in the Study

5.2. Disease Visualization

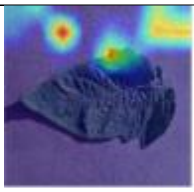
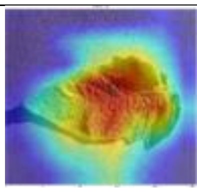
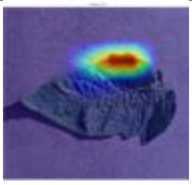
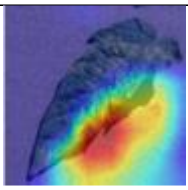
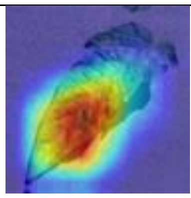
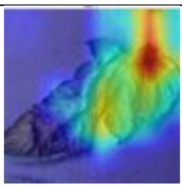
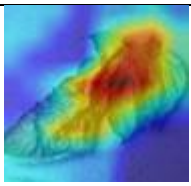
In this study, Grad-CAM (Gradient-weighted Class Activation Mapping) is used to make deep learning models for plant disease detection more interpretable. Grad-CAM, introduced by Selvaraju et al. [84], highlights the important areas in an image that influence a model's decision without needing to retrain or modify the model.

From the six models tested, MobileNet, ResNet, and DenseNet performed the best (as mentioned in Section 5.2). To evaluate how well they visualize plant diseases, images were analyzed in batches of 15, and Table 8 presents results for five sample images.

The Grad-CAM visualizations of these three models reveal key differences: DenseNet has a smaller but more consistent activation region, making its predictions more resistant to noise and outliers. ResNet shows inconsistent activation areas, often smaller and scattered across the leaf. Its deeper structure and skip connections may cause weaker activations, suggesting it needs a more diverse dataset. MobileNet sometimes activates areas outside the leaf, indicating certain layers are misfocusing. While it generally highlights leaf areas well, its wider activation zones may lead to overfitting when scaled up.

Overall, no single model is perfect for all plant disease classification challenges. A more accurate and adaptable model is needed, possibly using an ensemble approach where the best model is selected dynamically for each leaf instance.

Table 8.
Activation seeds in leaf images using Grad-CAM method on ResNet, DenseNet, and MobileNet.

Disease Name	DenseNet Result	MobileNet Result	ResNet Result
Spider_Mites	 99%	 100%	 99.9%
Leaf_mold	 100%	 100%	 99.9%
Late_Blight	 78%	 98%	 99.9%

6. Conclusion

This paper explores various machine learning models for plant disease detection, each with unique strengths. CNNs excel in image classification by learning spatial features automatically. SVM works well for small datasets with clear class boundaries, while Random Forests prevent overfitting. KNN is effective in lower-dimensional spaces, and YOLO enables real-time detection. Transfer learning improves accuracy by leveraging pre-trained models when data is limited.

The study evaluates deep learning models like VGG-16, VGG-19, InceptionNet, MobileNet, ResNet-50, and DenseNet-121 for tomato leaf disease classification, with MobileNet emerging as the best performer due to its high accuracy and fast training. However, certain diseases, such as Tomato Early Blight (D2) and Tomato Target Spot (D7), were harder to classify, indicating a need for more data. InceptionNet V3 showed poor accuracy and underfitting, while VGG-19 exhibited overfitting.

Grad-CAM proved useful in identifying model weaknesses and misclassification patterns, offering insights into areas for improvement. Future work will refine Grad-CAM for better disease localization and explore ensemble modeling to enhance classification reliability across diverse datasets. Additionally, an intelligent mobile app is planned to help farmers quickly identify plant diseases, enabling timely pesticide use for effective crop protection.

References

[1] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436-444, 2015. https://doi.org/10.1038/nature14539

[2] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84-90, 2017. https://doi.org/10.1145/3065386

[3] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," presented at the In European Conference on Computer Vision (ECCV), 818-833. https://doi.org/10.1007/978-3-319-10590-1_53, 2014.

[4] C. Szegedy *et al.*, "Going deeper with convolutions," in *Proceedings of the IEEE Conference on Computer Vision and pattern recognition. https://doi.org/10.1109/CVPR.2015.7298594*, 2015, pp. 1-9.

[5] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *In Proceedings of the International Conference on Learning Representations (ICLR). http://arxiv.org/abs/1409.1556*, 2014.

[6] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770-778. https://doi.org/10.1109/CVPR.2016.90*, 2016.

[7] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 7132-7141. https://doi.org/10.1109/CVPR.2018.00745*, 2018.

[8] C. Jackulin and S. Murugavalli, "A comprehensive review on detection of plant disease using machine learning and deep learning approaches," *Measurement: Sensors*, vol. 24, p. 100441, 2022.

[9] M. Shoaib *et al.*, "An advanced deep learning models-based plant disease detection: A review of recent research," *Frontiers in Plant Science,* vol. 14, p. 1158933, 2023.

[10] M. F. Guerri, C. Distante, P. Spagnolo, F. Bougourzi, and A. Taleb-Ahmed, "Deep learning techniques for hyperspectral image analysis in agriculture: A review," *ISPRS Open Journal of Photogrammetry and Remote Sensing,* p. 100062, 2024.

[11] M. Genaev and S. P. Mohanty, "Advances and challenges in image-based plant disease detection," *IEEE Xplore*, 2023.

[12] R. Alguliyev, Y. Imamverdiyev, L. Sukhostat, and R. Bayramov, "Plant disease detection based on a deep model," *Soft Computing,* vol. 25, no. 21, pp. 13229-13242, 2021. https://doi.org/10.1007/s00500-021-06176-4

[13] S. Swain, S. K. Nayak, and S. S. Barik, "A review on plant leaf diseases detection and classification based on machine learning models," *Mukt shabd,* vol. 9, no. 6, pp. 5195-5205, 2020.

[14] P. Sajitha, A. D. Andrushia, N. Anand, and M. Z. Naser, "A review on machine learning and deep learning image-based plant disease classification for industrial farming systems," *Journal of Industrial Information Integration,* vol. 38, p. 100572, 2024.

[15] W. B. Demilie, "Plant disease detection and classification techniques: a comparative study of the performances," *Journal of Big Data,* vol. 11, no. 1, p. 5, 2024.

[16] C. G. Simhadri, H. K. Kondaveeti, V. K. Vatsavayi, A. Mitra, and P. Ananthachari, "Deep learning for rice leaf disease detection: A systematic literature review on emerging trends, methodologies and techniques," *Information Processing in Agriculture,* 2024.

[17] A. Venkataramanan, D. K. P. Honakeri, and P. Agarwal, "Plant disease detection and classification using deep neural networks," *Int. J. Comput. Sci. Eng,* vol. 11, no. 9, pp. 40-46, 2019.

[18] J. Kolli, M. V. Dhara, and V. M. Manikandan, "Plant disease detection using convolutional neural network," presented at the In 2021 IEEE Bombay Section Signature Conference (IBSSC), pp. 1-6. IEEE, 2021.

[19] D. R. Hammou and M. Boubaker, *Tomato plant disease detection and classification using convolutional neural network architectures technologies. In: Ben Ahmed, M., Teodorescu, HN.L., Mazri, T., Subashini, P., Boudhir, A.A. (eds) Networking, Intelligent Systems and Security. Smart Innovation, Systems and Technologies, vol 237*. Singapore: Springer, 2022.

[20] H. Kibriya, A. Iram, and N. Amber, "Plant disease identification and classification using convolutional neural network and SVM," presented at the In 2021 International Conference on Frontiers of Information Technology (FIT), pp. 264-268. IEEE, 2021.

[21] P. Bharali, B. Chandrika, and B. Abhijit, "Plant disease detection by leaf image classification using convolutional neural network." In international conference on information, communication and computing technology." Singapore: Springer Singapore, 2019, pp. 194-205.

[22] M. M. H. Matin, A. Khatun, M. G. Moazzam, and M. S. Uddin, "An efficient disease detection technique of rice leaf using AlexNet," *Journal of Computer and Communications,* vol. 8, no. 12, pp. 49-57, 2020.

[23] H.-C. Chen *et al.*, "AlexNet convolutional neural network for disease detection and classification of tomato leaf," *Electronics,* vol. 11, no. 6, p. 951, 2022.

[24] M. A. Moid and A. C. Mousmi, "Transfer learning-based plant disease detection and diagnosis system using Xception," presented at the In 2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC), pp. 1-5. IEEE, 2021, 2021.

[25] M. H. Saleem, J. Potgieter, and K. M. Arif, "Plant disease classification: A comparative evaluation of convolutional neural networks and deep learning optimizers," *Plants,* vol. 9, no. 10, p. 1319, 2020.

[26] Ü. Atila, M. Uçar, K. Akyol, and E. Uçar, "Plant leaf disease classification using EfficientNet deep learning model," *Ecological Informatics,* vol. 61, p. 101182, 2021.

[27] F. Rajeena PP, A. SU, M. A. Moustafa, and M. A. Ali, "Detecting plant disease in corn leaf using efficientnet architecture—an analytical approach," *Electronics,* vol. 12, no. 8, p. 1938, 2023.

[28] M. Nagaraju and P. Chawla, "Systematic review of deep learning techniques in plant disease detection," *International journal of system assurance engineering and management,* vol. 11, no. 3, pp. 547-560, 2020.

[29] M. E. Chowdhury *et al.*, "Automatic and reliable leaf disease detection using deep learning techniques," *AgriEngineering,* vol. 3, no. 2, pp. 294-312, 2021.

[30] T. Anandhakrishnan and S. Jaisakthi, "Deep Convolutional Neural Networks for image based tomato leaf disease detection," *Sustainable Chemistry and Pharmacy,* vol. 30, p. 100793, 2022.

[31] M. M. Khalid and O. Karan, "Deep learning for plant disease detection," *International Journal of Mathematics, Statistics, and Computer Science,* vol. 2, pp. 75-84, 2024.

[32] P.-W. Chin, K.-W. Ng, and N. Palanichamy, "Plant disease detection and classification using deep learning methods: A comparison study," *Journal of Informatics and Web Engineering,* vol. 3, no. 1, pp. 155-168, 2024.

[33] V. Kumar, A. Hritik, and S. Jatin, "Resnet-based approach for detection and classification of plant leaf diseases," presented at the In 2020 International Conference on Electronics and Sustainable Communication Systems(ICESC), pp. 495-502. IEEE, 2020.

[34] I. Z. Mukti and D. Biswas, "Transfer learning based plant diseases detection using resnet50," presented at the 4th International Conference on Electrical Information and Communication Technology (EICT), Khulna, Bangladesh, 2019, pp. 1-6. https://doi.org/10.1109/EICT48899.2019.9068805, 2019.

[35] E. C. Too, L. Yujian, S. Njuki, and L. Yingchun, "A comparative study of fine-tuning deep learning models for plant disease identification," *Computers and Electronics in Agriculture,* vol. 161, pp. 272-279, 2019.

[36] J. Chen, J. Chen, D. Zhang, Y. Sun, and Y. A. Nanehkaran, "Using deep transfer learning for image-based plant disease identification," *Computers and electronics in agriculture,* vol. 173, p. 105393, 2020.

[37] J. Amara, B. Bassem, and A. Alsayed, "A deep learning-based approach for banana leaf diseases classification," 2017.

[38] K. P. Ferentinos, "Deep learning models for plant disease detection and diagnosis," *Computers and Electronics in Agriculture,* vol. 145, pp. 311-318, 2018. https://doi.org/10.1016/j.compag.2018.01.009

[39] M. Sibiya and M. Sumbwanyambe, "A computational procedure for the recognition and classification of maize leaf diseases out of healthy leaves using convolutional neural networks," *AgriEngineering,* vol. 1, no. 1, pp. 119-131, 2019.

[40] M. Arsenovic, M. Karanovic, S. Sladojevic, A. Anderla, and D. Stefanovic, "Solving current limitations of deep learning based approaches for plant disease detection," *Symmetry,* vol. 11, no. 7, p. 939, 2019.

[41] J. Rithik, P. Devi, S., and R. S. Brendan, M., "Plant disease classification using deep learning approach (VGG19)," presented at the International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2023, pp. 1715-1718, https://doi.org/10.1109/ICACCS57279.2023.10112823, 2023.

[42] T.-H. Nguyen, T.-N. Nguyen, and B.-V. Ngo, "A VGG-19 model with transfer learning and image segmentation for classification of tomato leaf disease," *AgriEngineering,* vol. 4, no. 4, pp. 871-887, 2022. https://doi.org/10.3390/agriengineering4040056

[43] A. A. Alatawi, S. M. Alomani, N. I. Alhawiti, and M. Ayaz, "Plant disease detection using AI based VGG-16 model," *International Journal of Advanced Computer Science and Applications,* vol. 13, no. 4, 2022.

[44] H. Baheti, A. Thakare, Y. Bhople, S. Darekar, and O. Dodmani, *Tomato plant leaf disease detection using inception V3. In: Kulkarni, A.J., Mirjalili, S., Udgata, S.K. (eds intelligent systems and applications lecture notes in electrical engineerin*. Singapore: Springer, 2023.

[45] M. Dutta, D. Gupta, Y. Gulzar, M. S. Mir, and A. B. Soomro, "Leveraging inception v3 for precise early and late blight disease classification in potato crops," *Traitement du Signal,* vol. 41, no. 2, 2024. https://doi.org/10.18280/ts.410213

[46] S. S. Puranik, "MobileNetV3 for mango leaf disease detection:an efficient deep learning approach for precision agriculture," presented at the International Conference for Emerging Technology (INCET), Belgaum, India, 2024, pp. 1-7. https://doi.org/10.1109/INCET61516.2024.10593318, 2024.

[47] C. Bi, J. Wang, Y. Duan, B. Fu, J.-R. Kang, and Y. Shi, "MobileNet based apple leaf diseases identification," *Mobile Networks and Applications,* pp. 1-9, 2022.

[48] D. Das, M. Singh, S. S. Mohanty, and S. Chakravarty, "Leaf disease detection using support vector machine," presented at the International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 2020, pp. 1036-1040. https://doi.org/10.1109/ICCSP48568.2020.9182128, 2020.

[49] M. Rajagopal, S. Kayikci, M. Abbas, and R. Sivasakthivel, "A novel technique for leaf disease classification using Legion Kernels with parallel support vector machine (LK-PSVM) and fuzzy C means image segmentation," *Heliyon,* vol. 10, no. 12, 2024.

[50] R. K. Dubey and D. K. Choubey, "Efficient prediction of blast disease in paddy plant using optimized support vector machine," *IETE Journal of Research,* vol. 70, no. 4, pp. 3679-3689, 2024.

[51] C. Nancy and S. Kiran, "Cucumber leaf disease detection using glcm features with random forest algorithm," *International Research Journal of multidisciplinary Technovation,* vol. 6, no. 1, pp. 40-50, 2024.

[52] V. Praba and K. Krishnaveni., "An enhanced random forest classifier to detect crop disease with texture and shape features of corn images (ERFCTS)," *Educational Administration: Theory and Practice,* vol. 30, no. 5, pp. 3720-3727, 2024.

[53] N. Kalyan and K. Rashmita, "A novel method for enhancing the accuracy in plant leaf disease detection using convolutional neural network over k-nearest neighbors classifier," in *In AIP Conference Proceedings, vol. 2853, no. 1. AIP Publishing*, 2024.

[54] M. P. Vaishnnave, K. S. Devi, P. Srinivasan, and G. A. P. Jothi, "Detection and classification of groundnut leaf diseases using KNN classifier," presented at the IEEE International Conference on System, Computation, Automation and Networking (ICSCAN), Pondicherry, India, 2019, pp. 1-5. https://doi.org/10.1109/ICSCAN.2019.8878733, 2019.

[55] A. Shill and M. A. Rahman, "Plant disease detection based on yolov3 and yolov4," presented at the International Conference on Automation, Control and Mechatronics for Industry 4.0 (ACMI), Rajshahi, Bangladesh, 2021, pp. 1-6. https://doi.org/10.1109/ACMI53878.2021.9528179, 2021.

[56] T. Y. Mahesh and M. P. Mathew, "Detection of bacterial spot disease in bell pepper plant using YOLOv3," *IETE Journal of research,* vol. 70, no. 3, pp. 2583-2590, 2024.

[57] X. Wang and J. Liu, "Vegetable disease detection using an improved YOLOv8 algorithm in the greenhouse plant environment," *Scientific Reports,* vol. 14, no. 1, p. 4261, 2024. https://doi.org/10.1038/s41598-024-54540-9

[58] M. Hussain, J. Bird, and D. Faria, "A study on cnn transfer learning for image classification. InUK workshop on computational intelligence 2018 Sep 5." Cham: Springer, 2018, pp. 191-202.

[59] B. Tlhobogang and M. Wannous, "Design of plant disease detection system: A transfer learning approach work in progress," presented at the In2018 IEEE International conference on applied system invention (ICASI) 2018 Apr 13 (pp. 158-161). IEEE, 2018.

[60] Y. Guo *et al.*, "Plant disease identification based on deep learning algorithm in smart farming," *Discrete Dynamics in Nature and Society,* vol. 2020, no. 1, p. 2479172, 2020.

[61] S. Coulibaly, B. Kamsu-Foguem, D. Kamissoko, and D. Traore, "Deep neural networks with transfer learning in millet crop images," *Computers in industry,* vol. 108, pp. 115-120, 2019.

[62] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," *Frontiers in plant science,* vol. 7, p. 215232, 2016.

[63] D. 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.

[64] Kaggle, "Plant pathology," Retrieved: https://www.kaggle.com/competitions/plant-pathology-2020-fgvc7, 2020.

[65] Kaggle, "Plant pathology," Retrieved: https://www.kaggle.com/competitions/plant-pathology-2021-fgvc8, 2021.

[66] AI Challenger, "AI challenger 2018 agricultural disease recognition dataset," Retrieved: https://challenger.ai/competition/agriculture, 2018.

[67] UCI Machine Learning Repository, "UCI leaf dataset," Retrieved: https://archive.ics.uci.edu/ml/datasets/leaf+data, 2011.

[68] Computer Vision Problems in Plant Phenotyping (CVPPP), "CVPPP plant phenotyping datasets," Retrieved: https://www.ibi.vu.nl/en/research-groups/visual-computing/plant-phenotyping.html, 2014.

[69] Kaggle, "Indian leaf dataset," Retrieved: https://www.kaggle.com/datasets/your-dataset-link, 2020.

[70] UCI Machine Learning Repository, "Plant seedlings dataset," Retrieved: https://archive.ics.uci.edu/ml/datasets/plant+seedlings+classification, 2017.

[71] S. Leminen Madsen, S. K. Mathiassen, M. Dyrmann, M. S. Laursen, L.-C. Paz, and R. N. Jørgensen, "Open plant phenotype database of common weeds in Denmark," *Remote Sensing,* vol. 12, no. 8, p. 1246, 2020. https://doi.org/10.3390/rs12081246

[72] D. Singh, N. Jain, P. Jain, P. Kayal, S. Kumawat, and N. Batra, "PlantDoc: A dataset for visual plant disease detection. ," in *InProceedings of the 7th ACM IKDD CoDS and 25th COMAD 2020 Jan 5 (pp. 249-253),* 2020.

[73] J. Parraga-Alava, K. Cusme, A. Loor, and E. Santander, "RoCoLe: A robusta coffee leaf images dataset for evaluation of machine learning based methods in plant diseases recognition," *Data in brief,* vol. 25, p. 104414, 2019.

[74] H. T. Rauf, B. A. Saleem, M. I. U. Lali, M. A. Khan, M. Sharif, and S. A. C. Bukhari, "A citrus fruits and leaves dataset for detection and classification of citrus diseases through machine learning," *Data in brief,* vol. 26, p. 104340, 2019.

[75] H. Durmus, E. O. Gunes, and M. Kirci, "Disease detection on the leaves of the tomato plants by using deep learning," in *Proceedings of the 6th International Conference on Agro-Geoinformatics*, 2017.

[76] M. Brahimi, M. Arsenovic, S. Laraba, S. Sladojevic, K. Boukhalfa, and A. Moussaoui, "Deep learning for plant diseases: detection and saliency map visualisation," *Human and machine learning: Visible, explainable, trustworthy and transparent,* pp. 93-117, 2018.

[77] A. Ramcharan *et al.*, "A mobile-based deep learning model for cassava disease diagnosis," *Frontiers in Plant Science,* vol. 10, p. 272, 2019.

[78] P. Jiang, Y. Chen, B. Liu, D. He, and C. Liang, "Real-time detection of apple leaf diseases using deep learning approach based on improved convolutional neural networks," *IEEE Access*, vol. 7, pp. 59069-59080, 2019.

[79] N. Nisha, G. Manogaran, and P. Vivekanandan, "Plant disease detection using support vector machine," *Journal of Computational and Theoretical Nanoscience,* 2020.

[80] R. Thangaraj, S. Anandamurugan, and V. K. Kaliappan, "Automated tomato leaf disease classification using transfer learning-based deep convolution neural network," *Journal of Plant Diseases and Protection*, vol. 128, no. 1, pp. 73-86, 2021. https://doi.org/10.1007/s41348-020-00403-0

[81] UCI Machine Learning Repository, "Leaf dataset," Retrieved: https://archive.ics.uci.edu/ml/datasets/Leaf, 2015.

[82] N. Lamrahi, "Tomato," Retrieved: https://WWW.kaggle.com/datasets/noulam/tomato, 2019.

[83] Kaggle, "Plant pathology 2020 - FGVC7," Retrieved: https://www.kaggle.com/competitions/plant-pathology-2020-fgvc7, 2020.

[84] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-cam: Visual explanations from deep networks via gradient-based localization," in *In Proceedings of the IEEE International Conference on Computer Vision (pp. 618-626)*, 2017.