Journal of Information Systems Engineering and Management
2025, 10(59s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Automated Detection of Tomato Plant Diseases: A Survey of Recent Advances

[1]G. Suguna, [2]S. Rathi,
[1]*Research Scholar, Department of Computer science and Engineering, Government College of Technology, Coimbatore*
[2]*Professor, Department of Computer science and Engineering, Government College of Technology, Coimbatore*

ARTICLE INFO

Received: 18 Dec 2024

Revised: 10 Feb 2025

Accepted: 28 Feb 2025

ABSTRACT

Tomato farming is of great importance to global agriculture but is ever troubled by different plant diseases. Such plant diseases are detected earlier and accurately to allow intervention in a timely manner to salvage losses in yield. This survey paper serves the purpose of giving clear insights into the recent emerging technologies of machine learning (ML) and deep learning (DL) models in detecting the early tomato plant diseases. It analyses their use in various methodologies-the conventional image processing techniques, supervised ML algorithms and deep learning network architectures as well as the convolutional neural networks (CNNs) and their variants. The review discusses the pros and cons of the different approaches, identifies the common challenges encountered when put into real-world applications and suggests avenues for future research. This survey would serve as a one-stop, consolidated understanding of what is currently considered to be the state-of-the-art for automated tomato disease detection for research and practitioners looking forward to developing effective and powerful agricultural monitoring systems.

Keywords: Tomato farming, plant diseases, early detection, machine learning (ML), deep learning (DL), image processing, supervised learning, convolutional neural networks (CNNs), automated disease detection

1. INTRODUCTION

Tomato (Solanum lycopersicum) is one of the most economically important horticultural crops. Its production is greatly affected by many disease pathogens, which include bacterial spots, early blight and leaf mold. The traditional methods of disease detection, which consist essentially of visual inspection of tomato plants by a selected group of experts, are time-consuming, generally subjective and labor-intensive. Destructive impact prevention through prompt detection is the best way of disease management; thus, the developing of automated, accurate disease detection systems forms an important area of study. This review paper looks into the recent developments in machine learning (ML) and deep learning (DL) approaches for the early diagnosis of tomato plant diseases, which hold immense promise in changing agricultural practices.

History of Techniques used for Detection of Plant disease:

The introduction of machine learning (ML) and deep learning (DL) techniques for tomato plant disease detection has grown substantially in history. Established in the 1990s, ML approaches, such as those using the famous decision trees, support vector machines (SVMs) and k-nearest neighbors (k-NN), were adapted into crop disease detection, including tomatoes. Further into the decade, simple image processing techniques were incorporated into such traditional ML methods to significantly improve the effectiveness of plant disease classification. A classification of tomato plant diseases using handcrafted features was carried out by SVM and k-NN by the year 2012.

Random Forest and SVM further increased the classification accuracy for tomato diseases in 2014. In fact, the real revolution was in 2015 when the use of Convolutional Neural Networks (CNNs) brought in new technology for plant disease detection based on deep learning. The subsequent years witnessed models based on deep learning growing like AlexNet, VGG16, and ResNet in 2016, followed by advanced models of CNN also in 2017 with better accuracies.

Journal of Information Systems Engineering and Management
2025, 10(59s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

By 2018, transfer learning had become a common avenue with the adoption of pre-trained models such as VGG16, ResNet and Inception, all for better performance. New techniques named DenseNet and EfficientNet were introduced in 2019 and were meant to detect diseases robustly. In 2020, another technological advancement was seen when Generative Adversarial Networks (GANs) were used in data augmentation to enhance the variety of datasets for overall improved model training.

Year 2021 saw the introduction of attention mechanisms and transformers, thereby improving feature extraction capabilities. These techniques slightly advanced towards explainable AI by 2022, as they applied Grad-CAM and SHAP to interpret deep learning models. With hyperspectral imaging, deep learning models increased their accuracy level in 2023, then progressed in 2024 by combining traditional image processing techniques with models like EfficientNet and VGG16 for context-aware detection of tomato plant diseases.

This survey intends to:

Some of the machine learning and deep learning methods involved in detecting diseases in tomato plants are given in Table 1.

1. Surveying among the different ML and DL methods employed in detecting and classifying the ailments of tomato plants.

2. Analyze the challenges in applying ML and DL towards the detection of diseases in tomatoes.

3. Provide an in-depth study on research that makes use of ML and DL approaches for tomato disease detection. Merit the key researched papers that have applied these two methods in their study.

4. Identify the best research area or most active and significant research in tomato plant disease detection. Locate areas that have high potentials in terms of developing or already shown possible efficient results in improving disease detection accuracies.

More machine learning and deep learning methods can be critical for detecting diseases of tomato plants. Survey concerning different ML and DL systems employed to detect and classify tomato plant diseases. Identify the challenges of successfully implementing ML and DL techniques in tomato plant disease detection. Discuss the issues, such as data shortages, environmental variations, and computational requirements. Conduct an in-depth investigation of the research applying ML and DL schemes in detecting diseases in tomato plants. Review and analyze key research papers and studies that have utilized ML and DL techniques on this regard. Find the most indicative and useful research focus on tomato plant disease detection.

Year	Technique/Development	Description
1990s	Traditional ML (Decision Trees, SVMs, k-NN)	Early use of machine learning algorithms for crop disease detection, including tomato diseases.
2000s	Image Processing + ML Algorithms	Basic image processing combined with traditional ML methods for disease classification.
2012	SVM, k-NN	Introduction of SVM and k-NN for classifying tomato plant diseases using handcrafted features.
2014	Random Forest, SVM	Improved classification accuracy for tomato diseases using Random Forest and SVM.
2015	Convolutional Neural Networks (CNNs)	Start of using CNNs for plant disease detection, including tomato diseases.
2016	AlexNet, VGG16, ResNet	Use of deep learning models like AlexNet, VGG16, and ResNet for tomato disease detection.
2017	Advanced CNN Architectures	Sophisticated CNN architectures employed for higher accuracy in tomato disease detection.
2018	Transfer Learning	Application of pre-trained models (VGG16, ResNet, Inception) for better performance.
2019	DenseNet, EfficientNet	Use of DenseNet and EfficientNet for robust tomato disease detection.
2020	GANs for Data Augmentation	Integration of GANs to diversify training datasets for improved model training.
2021	Attention Mechanisms, Transformers	Implementation of attention mechanisms and transformers for enhanced feature extraction.
2022	Explainable AI (Grad-CAM, SHAP)	Introduction of Grad-CAM and SHAP for interpretability of deep learning models.
2023	Hyperspectral Imaging + DL	Use of hyperspectral imaging combined with deep learning for higher accuracy in disease detection.
2024	Fusion of Image Processing + DL	Combining traditional image processing with EfficientNet and VGG16 for context-aware detection.

Table1: History of Techniques used for Detection of Plant disease

Find promising areas of research that have demonstrated effective results or potentiality to improve disease detection accuracy.

Journal of Information Systems Engineering and Management
2025, 10(59s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

2. TRADITIONAL MACHINE LEARNING APPROACHES

Automated plant disease detection has not come a long way. In fact, early implementations of plant disease detection employed conventional machine learning algorithms combined with image processing techniques for feature extraction. This concept mainly works in a two-step framework:

2.1 Image Preprocessing and Feature Extraction:

Image Preprocessing: This phase maintains the quality of input images and reduces noise. Basic techniques are image resizing, filtering and changing color spaces.

Feature Extraction: This phase decides which relevant feature(s) will be extracted from the preprocessed images for indicating differences between healthy and diseased plants. These are classical ways of feature extraction:

Color-based features: Involves analyzing the distribution of colors in the image such as mean, standard deviation and histograms [12] (Khan & Narvekar, 2022).

Texture-based features: Based on describing the spatial arrangement of pixel intensities using Gray Level Co-occurrence Matrices (GLCM) and Local Binary Patterns (LBP) [19] (Ahmad et al., 2024).

Shape-based features: Pull out geometric properties of diseased areas, mainly area, perimeter and shape descriptors. Nazari et al. (2022) [2] applied image processing for detecting the Alternaria disease and leafminer pest. Li et al. (2023) [5] and Zhang et al. (2022) [23] worked on desnoising and infrared image fusion, respectively, in order to improve the quality of classifying images. Further refinement of image quality was studied by Yeswanth and Deivalakshmi (2024) [16] using image super-resolution.

2.2 Supervised Learning Algorithms:

Once the features are extracted, a classifier-disease is assigned to the different classes of plant samples using supervised learning algorithms. The common algorithms used are:

Support Vector Machines (SVMs): Best for binary classification of high-dimensional data.

K-Nearest Neighbors (k-NN): Samples are classified based on the majority class of the nearest neighbors.

Random Forest: An ensemble of decision trees improves on classification accuracy and robustness. Traditional machine learning methods, although being appropriate in some cases, depend on hand-crafted features, which need time to design and may not encapsulate the complex patterns present in plant disease images.

3. DEEP LEARNING APPROACHES

Deep learning, notably CNNs, has dramatically changed the game for image-based plant disease detection by learning hierarchical features from raw images by itself.

3.1 Convolutional Neural Networks (CNNs):

CNNs are made up of various layers-convolutional layers, pooling layers and fully connected layers-and work together by classifying the input images. Convolutional layers learn to extract local features from the input image by convolving filters with input images. Pooling layers reduce the dimensions of feature maps, making the model invariant to changes in the size and orientation of images. The final layers connect to the output layer that infers the eventual classification of learned features. Studies, such as Anandhakrishnan and Jaisakthi (2022)[9] and Özbılge et al. (2022)[10], have shown that CNNs have been successful in recognizing diseases in tomatoes.

3.2 Transfer Learning:

Building a deep CNN involves large amounts of labeled data, a difficult task for plant disease detection. One use of transfer learning is to mitigate this issue, in which pre-trained models are built on large datasets (like ImageNet) and are then fine-tuned to the task of plant disease detection[4](Sanida et al., 2023; Djimeli-Tsajio et al., 2022)[7]. With this procedure, the training time is vastly reduced while model performance enhanced using limited data.

Journal of Information Systems Engineering and Management
2025, 10(59s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

3.3 Advanced CNN Architectures:

Most advanced CNN architectures have been studied to enhance the efficacy and accuracy of plant disease detection: Huang et al. (2023) presented [3] FC-SNDPN, a network to encapsulate space-dependent aspects. Li et al. (2023) have developed [5] LMBRNet, a network dedicated to tomato leaf disease identification. Zou et al. (2024)[5] and Sun et al. (2024)[21] further proved the fusion of Convolutional Neural Networks as well as Visual Transformers for improved performance metrics.

This processing is done at any level- from basic level processing to advanced one. Different categories of hierarchical, multi-stage and multi-view processing modules have been integrated for improving flexibility, stability and identification accuracy in plant leaves. Harris et al. (2016) and Das et al. (2018) used various deep learning networks for leaf disease detection. Deep spiking neural networks (DSNN) have been proposed in Chen et al. (2023).

3.5 Novel Neural Network Architectures:

Battiloro et al. (2024)[24] described generalized simplicial attention neural networks. Liu et al. (2021)[25] examined hybrid quantum-classical convolutional neural networks.

4. Advanced Techniques and Applications

4.1 Ensemble Learning:

Ensemble learning integrates predictions from multiple models in order to improve general accuracy and robustness. Weighted ensemble learning was used in the study of Javidan et al. (2024) [6] to effectively classify tomato leaf diseases.

4.2 Early Detection and Remote Sensing:

da Cunha et al. (2023) [11] demonstrated the potential of integrating remote sensing and artificial intelligence for the early detection of tomato bacterial spot disease in transplant seedlings. By employing hyperspectral imaging, the study captured subtle physiological changes in the seedlings before visible symptoms appeared. These spectral patterns were analyzed using advanced AI models, including machine learning and deep learning algorithms, to accurately classify diseased and healthy plants. This non-invasive, real-time approach enables early intervention and improves disease management strategies in nursery environments. The research highlights the growing relevance of precision agriculture tools, where the combination of spectral sensing and AI facilitates proactive crop monitoring and supports sustainable agricultural practices.

4.3 Explainable AI (XAI):

In 2024, Rajpal et al. [15] developed a deep learning model focused on extracting leaf disease segments with enhanced interpretability using Explainable AI (XAI) techniques. Their approach not only aimed at accurate segmentation of diseased areas on plant leaves but also emphasized transparency in model predictions, which is critical for building trust in AI-based agricultural solutions. By incorporating tools such as **Grad-CAM (Gradient-weighted Class Activation Mapping)** and **saliency maps**, the model provided visual explanations highlighting the specific regions and features that contributed to the classification decision. This enabled researchers and agronomists to validate the model's focus on biologically relevant symptoms, such as lesions, discoloration, or necrotic patches. The integration of XAI ensured that the deep learning system operated as a decision support tool, offering actionable insights while maintaining transparency, accountability and user confidence in real-world agricultural applications.

5. CHALLENGES AND FUTURE DIRECTIONS

Despite the substantial effort of having been made in recent years, there are still a number of challenges for ML and DL when it comes to early detection of tomato crop pathologies.

5.1 Real-world Variability:

Journal of Information Systems Engineering and Management
2025, 10(59s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

The symptoms of plant diseases may differ according to the different factors such as the lighting conditions, growth stages and environmental conditions. Developing strong models that generalize well for these variables is a serious challenge.

5.2 Data Scarcity and Imbalance:

Data source collecting huge labeled datasets of plant disease images is a long and costly affair. The datasets are often imbalanced, wherein healthy samples vastly outnumber diseased ones.

5.3 Computational Resources:

Rich resources in computing capacity are required for training and deploying various complex deep learning models, especially CNN architectures. It is very much significant to develop efficient lightweight models that can still be deployed on resource-constraint devices for real applications.

5.4 Early Detection Challenges

Detection of diseases at their nascent stages poses a problem with subtle symptoms that are sometimes indistinguishable from healthy plant variations. Whereas traditional machine learning methods can be deployed against tomato plant disease detection, they are outdone by deep learning (mainly CNNs) methods in respect of the following potential constraints:

Feature Extraction Bottleneck: Traditional ML requires the feature extraction process to be done manually, which is tedious and could give rise to feature sets not very optimal and may not generalize well. CNNs do this automatically.

Dealing Complexity and Variability: Traditional ML most likely has a problem dealing with the complexity and variability of plant disease images (lighting varies, angle variations, etc.), while CNNs do get designed to deal with this complexity.

Data Dependency: The performance of traditional ML plateaus with small to moderate data; in contrast, CNNs thrive on large datasets.

Generalization: Traditional ML models may not generalize well when put against new and unseen data compared to CNNs.

6. EXPERIMENTAL RESULTS:

The Plant Village dataset provided images of tomato leaves exhibiting nine disease types and one healthy class for the purpose of this Experiment. The images were subjected to feature extraction techniques focusing on shape, color and texture to acquire meaningful features. The features extracted were cataloged into a CSV file for further analysis. Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Decision Tree and Random Forest were machine-learning models trained and tested on this dataset. With good implementation over large-scale publicly available image datasets, these models offer a systematic approach to accurately classifying several tomato diseases through a dependable and efficient mechanism for early disease detection.

6.1 Methodology:

Figure 1 illustrates the complete dataflow for a plant disease detection system using the PlantVillage dataset. The process begins with **Data Acquisition**, where images of tomato leaves are sourced from the PlantVillage dataset. In the **Data Preprocessing** stage, the images are resized to 512×512 pixels, followed by Gaussian filtering for noise reduction and Otsu's Thresholding for image segmentation. Next, the **Feature Extraction** stage involves extracting shape, color and texture features from the preprocessed images. These features are then input into the **Classification** stage, where various machine learning algorithms such as SVM, KNN, Decision Tree and Random Forest, as well as deep learning models like CNN with VGG16, ResNet50 and InceptionV3 are used to classify the images.

Journal of Information Systems Engineering and Management
2025, 10(59s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

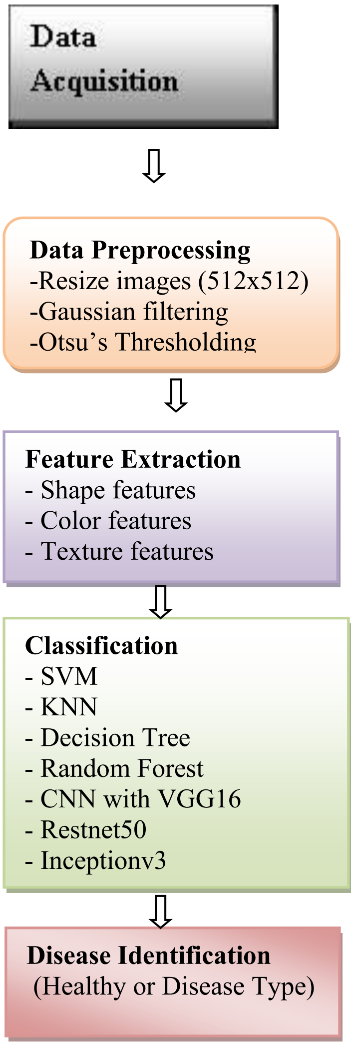

Fig 1: Dataflow Diagram

The final step is **Disease Identification**, which determines whether the leaf is healthy or affected by a specific disease type.

6.1.1 Data Acquisition

Approximately, 18200 images of tomato leaves obtained from the **PlantVillage** repository can be sorted according to 10 classes. The whole data set is 150MB in size and was downloaded using an IDM downloader.

6.1.2 Data Preprocessing

1. Images were resized to a 512×512 resolution to optimize training efficiency and computational performance.
2. Gaussian filtering was applied to smooth images by reducing noise and enhancing essential features.
3. Otsu's thresholding method was utilized for adaptive image segmentation, effectively distinguishing diseased regions from the background.

6.1.3 Feature Extraction

Key features such as shape, color and texture were extracted to enable precise classification of different tomato diseases.

6.1.3 Classification with Machine Learning Algorithms

Journal of Information Systems Engineering and Management
2025, 10(59s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

There has been an extensive effort to enhance tomato leaf disease classification by combining the art of traditional machine learning classifiers with CNN architectures like **VGG16, ResNet50** and **InceptionV3**. Features were extracted from images using pre-trained models and fed to classifiers-**SVM, KNN, Decision Tree** and **Random Forest**. VGG16 performed well in terms of accuracy with deep hierarchical feature extraction but it consumes more computational power. ResNet50, on the other hand, facilitated feature propagation through residual connections, reduced its vanishing gradient problem and hence improved classification performance. InceptionV3 on its side was effectively capturing multi-scale features via factorization convolution operations, thus enabling better characterization of diseases.

Techniques	Dataset	Accuracy	Precision	Recall	F1 Score
Decision Tree	Plant Village (with labeled features)	80%	79	80	80
Random Forest	Plant Village	92%	90	91	90
SVM	Plant Village	84%	84	84	84
KNN	Plant Village	90%	91	91	91

Table 1: Performance measures on Machine Learning Algorithms

The combination of deep learning feature extraction with machine learning classifiers promises good performance in providing a powerful solution for tomato disease classification.

Based on the table 1 tabular results, the performance measures indicate that **Random Forest** achieved the highest accuracy at **92%**, along with strong precision (90), recall (91) and F1 score (90), making it the most effective and reliable model for classifying tomato plant diseases using the PlantVillage dataset. **K-Nearest Neighbors (KNN)** closely follows with **90% accuracy** and the highest scores for precision, recall and F1 score (all 91), showing excellent consistency and balance, particularly in identifying true positives and minimizing false predictions. **Support Vector Machine (SVM)** displayed consistent performance with all metrics at **84%**, indicating a stable but slightly lower classification capability compared to the top performers. **Decision Tree**, while simple and interpretable, yielded the lowest performance with **80% accuracy** and slightly lower precision and recall, suggesting its limitations in handling more complex patterns. Overall, Random Forest stands out as the best classifier, with KNN also proving to be a highly competitive and dependable model.

Journal of Information Systems Engineering and Management
2025, 10(59s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

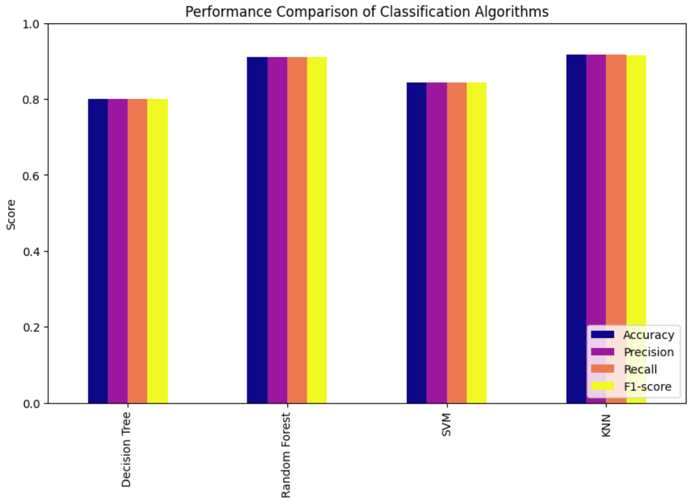

Fig 2: Performance Comparison based on metrics

The figure 2 presents bar graphs that show comparisons among the performance of different classification algorithms; namely Decision Tree, Random Forest, SVM and KNN, against the four evaluation metrics—Accuracy, Precision, Recall and F1-score. The x-axis represents each algorithm, while the y-axis indicates the score values, which range between 0 to 1. Random Forest and KNN were observed to have the highest scores among all metrics while Decision Trees were the ones with the lowest. There is also a legend that shows the colors incorporated for each of the metrics, aiding in the interpretation of the results.

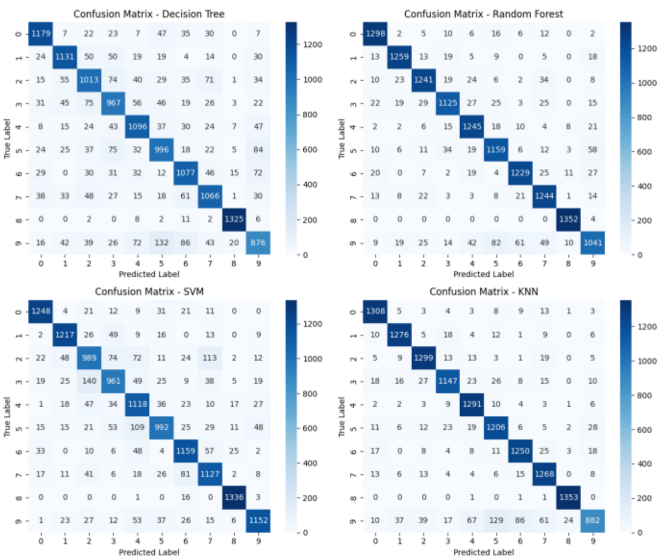

Fig 3: Comparative performance of Confusion Matrix

Confusion Matrixes (fig 3)1-4 show the comparative performance of various classification algorithms: Decision Tree, Random Forest, SVM and KNN. Each of the respective confusion matrices depicts the classification efficacy of the model in distinguishing among different classes through the means of correct and incorrect predictions.

Key observations:

The diagonal elements correctly classify instances, with greater values resulting in better performance.

•Random Forest and KNN yield very high performance in classification due to the high diagonal values and minimal misclassifications.

Journal of Information Systems Engineering and Management
2025, 10(59s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

•Decision Tree seems to misclassify more instances than any other model and thus has a relatively low performance.

•SVM is performing well but the off-diagonal elements show some misclassifications.

These matrices let us evaluate the effectiveness of each model in multi-class classification by analyzing false positives and false negatives.

6.1.4 Classification with Deep Learning Algorithms

Techniques	Dataset	Accuracy	Precision	Recall	F1 Score
CNN with VGG16	Tomato Leaves	92%	91	92	91
Restnet50	Tomato Leaves	94%	93	94	93
Inception v3	Tomato Leaves	90%	91	91	91

Table 2: Performance measures on Deep Learning Algorithms

Table 2 illustrates the performance comparison of deep learning models applied to tomato leaf disease classification reveals significant insights into their effectiveness. Among the three, **ResNet50** emerges as the best-performing model, achieving the highest **accuracy of 94%**, along with **precision of 93**, **recall of 94** and **F1 score of 93**. These metrics indicate that ResNet50 not only classifies the diseased and healthy leaves accurately but also maintains a strong balance between identifying true positives and minimizing false positives or negatives, making it highly reliable for real-world applications in agriculture.

The **CNN with VGG16** model also demonstrates excellent performance with an **accuracy of 92%**. It achieves **precision of 91**, **recall of 92** and **F1 score of 91**, which shows its ability to consistently and effectively detect tomato leaf diseases. VGG16 is particularly favored for its simpler architecture and computational efficiency while still maintaining high classification performance, making it suitable for resource-constrained environments.

On the other hand, **InceptionV3**, while slightly behind in terms of accuracy (**90%**), shows a well-balanced performance across **precision, recall and F1 score—all at 91**. This consistency highlights its robustness and reliability in handling diverse and complex image data, despite not leading in raw accuracy.

In summary, all three deep learning models—ResNet50, VGG16 and InceptionV3—are capable of delivering high-performance results in tomato plant disease detection. However, **ResNet50 stands out as the most powerful and accurate model**, making it highly suitable for deployment in precision agriculture systems. **VGG16 offers a good trade-off between performance and efficiency**, while **InceptionV3 excels in maintaining consistency across evaluation metrics**, making it a dependable model in varying conditions.

Journal of Information Systems Engineering and Management
2025, 10(59s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

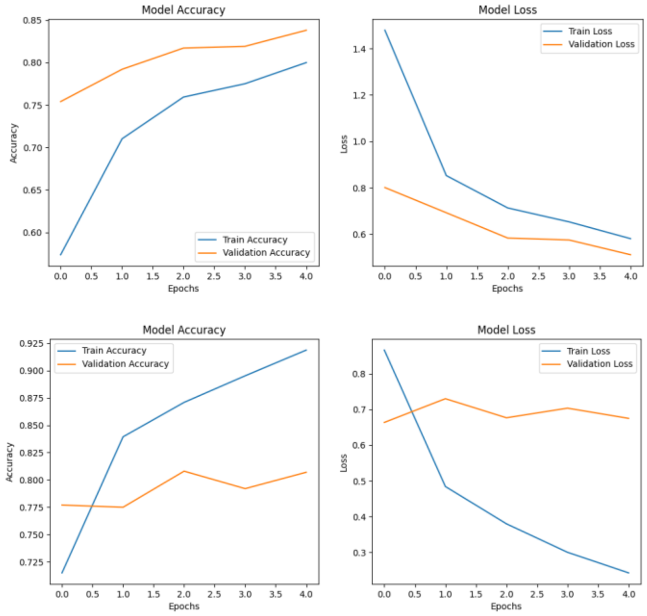

Fig 4: Accuracy of the Model

The graphs fig 4 illustrate the training progress of the complete model for five epochs. The left side displays the model accuracy, with a significantly rising training accuracy gradually achieving over 92 percent, while the validation accuracy fluctuates about 80 percent, suggesting possible over fitting. The graph on the right contains model loss with a dramatic drop in training loss while validation loss remains more stable but occurred with slight fluctuations. This infers that the model is not generalizable as well.

The second mentioned graphs would be showing a smooth learning process generally as an InceptionV3-based model is one of the most efficient architectures with factorized convolutions and auxiliary classifiers that help in eliminating overfitting. Training and validation accuracy would converge without sharp fluctuations while validation loss must be lower since it would better extract features. In case of over fitting one might opt for dropout or data augmentation techniques to improve further generalization.

6. COMPARATIVE ANALYSIS

S.no	Author and Title of the paper	Year of Publication	Aim	Dataset	Techniques	Drawback	Accuracy
1.	Rahman, Sami Ur, et al. “Image processing based system for the detection, identification and treatment of tomato leaf diseases”	Multimedia Tools and Applications (2023)	The goal is to facilitate early diagnosis and effective management, reducing crop losses and improving agricultural productivity.	The dataset comprises 400 images of tomato leaves.	1. Gray Level Co-occurrence Matrix (GLCM) algorithm. 2. Support Vector Machine (SVM) classifier.	1.Data Dependency, 2.Computational Resources, 3.Environmental Variability, 4.Generalizability	92%
2.	Nazari, Kobra, Mohammad	Journal of the	The study underscores the	The study utilized 272	Neural networks,	1.Limited Dataset Size	84.71%

Journal of Information Systems Engineering and Management
2025, 10(59s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

	Javad Ebadi, and Kamal Berahmand. Et al. " Diagnosis of alternaria disease and leafminer pest on tomato leaves using image processing techniques"	Science of Food and Agriculture (2022)	potential of integrating image processing and machine learning techniques for plant disease and pest diagnosis, contributing to more efficient and sustainable agricultural practices.	tomato leaf images from Vali-e-Asr University of Rafsanjan's farm, comprising 100 healthy leaves and 172 infected leaves at initial stages	and an Adaptive Network-Based Fuzzy Inference System (ANFIS) classifier	2.Specific to Certain Diseases 3.Accuracy Constraints 4.Dependence on Image Quality 5.Computational Requirements	
3.	Huang, Xibei, et al. "Tomato leaf disease detection system based on FC-SNDPN"	Multimedia tools and applications (2023)	To develop an automated and accurate system for detecting and identifying tomato leaf diseases, particularly in the context of crops grown in southern China	Dataset consisting of **1,000 images** of tomato leaves	Fully Convolutional–Switchable Normalization Dual Path Network (FC-SNDPN) and VGG-16	1.Limited Dataset Diversity 2.Dependency on Image Quality 3.Model Over fitting 4.Computational Resources Requirement	95.40%
4.	Sanida, Theodora, et al."Tomato leaf disease identification via two–stage transfer learning approach"	Smart Agricultural Technology (2023)	To develop an efficient and accurate system for identifying and classifying diseases in tomato leaves using a two-stage transfer learning framework.	**PlantVillage**, which contains a total of **54,000 images** of tomato leaves	VGGNet and Inception blocks	1.Dependency on Large Dataset 2.Image Quality Sensitivity 3.Overfitting Risk 4.Limited Disease Coverage 5.High Computational Requirements	90.23%

Journal of Information Systems Engineering and Management
2025, 10(59s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

5.	Li, Mingxuan, et al. "Identification of tomato leaf diseases based on LMBRNet"	Engineering Applications of Artificial Intelligence (2023)	To develop an advanced deep learning model, LMBRNet, for accurate and efficient identification of tomato leaf diseases	Dataset comprising over **8,000 images** of tomato leaves.	Comprehensive Grouped Differentiated Residual (CGDR) structure and Multiple Residual Connections	1.Dependency on Large Datasets 2.High Computational Requirements 3.Limited Generalization in Some Scenarios 4.Complexity of Model Architecture 5.Sensitivity to Image Quality	92.7%
6.	Javidan, Seyed Mohamad, et al."Tomato leaf diseases classification using image processing and weighted ensemble learning"	Agronomy Journal 116.3 (2024)	To develop an efficient and accurate system for classifying tomato leaf diseases by combining image processing techniques with ensemble learning	Plant Village dataset with **8,000 images** of tomato leaves	1.Support Vector Machine (SVM) 2.Decision Tree 3.Random Forest 4.K-Nearest Neighbors (KNN) 5.Naïve Bayes 6.Discriminant Analysis	1.Dependency on Feature Extraction 2.Limited Generalization 3.Image Quality Sensitivity 4.LimitedCoverage of Diseases 5.Ensemble Learning Overhead	Simple Majority Voting: 93.49% Weighted Majority Voting: 95.58%
7.	Djimeli-Tsajio, Alain B., et al."Improved detection and identification approach in tomato leaf disease using transformation and	Journal of Plant Diseases and Protection 129.3 (2022)	The primary goal is to enhance the feature extraction process by leveraging pre-trained deep learning models and applying image	**PlantVillage dataset**, which consists of **over 50,000 images** of plant leaves	VGGNet and SVM	1.Dependence on Large Datasets 2.Image Quality 3.Computational Complexity	91.92%

Journal of Information Systems Engineering and Management
2025, 10(59s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

	combination of transfer learning features"		transformations, which improve the classification performance			4.Limited Disease Coverage 5.Overfitting Risk 6.High Resource Requirements	
8.	Kaur, Prabhjot, et al."An approach for characterization of infected area in tomato leaf disease based on deep learning and object detection technique"	Engineering Applications of Artificial Intelligence 115 (2022)	To combine deep learning methods with object detection techniques to automatically detect and segment the infected regions in leaf images	Dataset consisting of **1,000 images** of tomato leaves	1.Convolutional neural networks (CNNs) 2.YOLO (You Only Look Once)	1.Dependency on 2.Large Datasets Sensitivity to Image Quality 3.High Computational Cost 4.Limited Disease Coverage 5.Overfitting 6.Complexity in Real-World Deployment	97.6%
9.	Anandhakrishnan, T., and S. M. Jaisakthi. "Deep Convolutional Neural Networks for image based tomato leaf disease detection"	Sustainable Chemistry and Pharmacy 30 (2022)	The paper focuses on leveraging the power of CNNs to automatically analyze and classify tomato leaf images, identifying various diseases with high accuracy	Dataset consists of **over 50,000 images** that cover various diseases in tomatoes	CNN(Convolutional Neural Networks)	1.Computational Cost 2.Overfitting 3.Limited Disease Coverage 4.Generalization Issues	98.3%
10.	Özbılge, Emre, et al. "Tomato disease recognition using a	IEEE Access 10 (2022)	The goal is to create a model that requires less computational power and	Dataset contains **14,676 images** for training and	CNN and Optimization Techniques	1.Limited Dataset Variety 2.Dependency on Image	95.75%

Journal of Information Systems Engineering and Management
2025, 10(59s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

	compact convolutional neural network"		memory, making it suitable for deployment in resource-constrained environments such as farms and rural areas	testing the model.		Quality 3.Model Complexity	
11.	da Cunha, Vitor A. Gontijo, et al."Early detection of tomato bacterial spot disease in transplant tomato seedlings utilising remote sensing and artificial intelligence"	Biosystems Engineering 234 (2023)	The goal is to enable precise and timely identification of the disease, allowing for early intervention and better management of the disease to prevent its spread	Dataset utilizes **1,200 images** from a dataset of tomato seedlings	1.Remote Sensing 2.Artificial Intelligence (AI) 3.Support Vector Machines (SVM)	1.Dependence on High-Quality Remote Sensing Data 2.Complexity of Implementation 3. Data Availability and Variety 4.High Computational Requirements	96.3%
12.	Khan, Saiqa, and Meera Narvekar et al., "Novel fusion of color balancing and super pixel based approach for detection of tomato plant diseases in natural complex environment"	Journal of King Saud University -Computer and Information Sciences 3 4.6 (2022)	Aims to enhance the accuracy and reliability of disease detection under varying lighting conditions and complex backgrounds, contributing to early disease management and improved agricultural outcomes.	1.Internet downloaded image dataset(57 images) 2. Real world dataset 3.Combined dataset	1.K-means clustering 2.Pyramid of HOG(PHOG) 3.Gray Level Co-occurrence Matrix (GLCM) 4.Random Forest (RF)	1.Dependence on Color Features 2.Complexity of Super pixel Segmentation 3. Performance in Low-Light or High-Contrast Conditions 4.Sensitivity to Background Variability 5.Generalization Across Different	93.12%

Journal of Information Systems Engineering and Management
2025, 10(59s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

						Tomato Varieties	
13.	Khan, Khalil, et al. "End-to-End Semantic Leaf Segmentation Framework for Plants Disease Classification"	Complexity 2022	To develop an integrated system that automates the process of plant disease diagnosis by combining semantic leaf segmentation with disease classification	Dataset of 20,000 images	CNN based Model	1.Data Imbalance 2. Limited Focus on Disease Severity Estimation 3.Interpretability Issues	97.6%
14.	Pal, Arunangshu, and Vinay Kumar. "Plant Leaf Disease severity classification using agriculture detection framework"	Engineering Applications of Artificial Intelligence (2023)	It employs image processing and machine learning techniques to detect and evaluate disease severity, aiding in effective agricultural decision-making.	**8,000 images** of plant leaves	1.Convolutional Neural Networks (CNNs) 2.Transfer Learning 3.Support Vector Machines (SVMs) and Random Forest	1.Sensitivity to Environmental Factors 2.Limited Interpretability 3.Dependence on Image Quality 4. Over fitting Risk	90%
15.	Rajpal, Ankit, et al. "Explaining deep learning-based leaf disease identification"	Soft Computing (2024)	To explore and provide an interpretative framework for deep learning models used in plant leaf disease identification	**50,000 images** of plant leaves	1.Convolutional neural networks (CNNs) 2. Explainable AI (XAI)	1.Computational Resources and Time 2.Interpretability Challenges in Complex Models 3. Dependency on High-Quality Data	95%
16.	Yeswanth, P. V., and S. Deivalakshmi" ASFESRN: bridging the gap in real-time corn leaf	Multimedia Systems 30.4 (2024)	To develop a real-time framework that improves the accuracy of corn leaf disease detection by	**1000 images** of corn leaves	1.ASFESRN (Adaptive Super-Resolution and Feature Enhancement for Real-	1.Dependence on Image Quality 2.Computational Complexity	95%

Journal of Information Systems Engineering and Management
2025, 10(59s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

	disease detection with image super-resolution"		enhancing the quality of low-resolution images through super-resolution techniques.		time Networks) 2. Convolutional Neural Networks (CNNs)	3. Generalization to Other Crops 4.Interpretability Issues	
17.	Sun, Changxia, et al."Research on tomato disease image recognition method based on DeiT"	European Journal of Agronomy 162 (2024)	The paper focuses on leveraging the power of transformer-based architectures to extract relevant features from images of tomato leaves.	54,000 images of tomato leaves, with 10,000 images per disease type	1.Vision Transformer (DeiT) 2.EMA-DeiT Model 3.Transfer Learning	1.Limited Dataset Diversity 2.Generalization Issues 3.Lack of Explainability	89.6%
18.	Harakannanavar, Sunil S., et al. "Plant leaf disease detection using computer vision and machine learning algorithms"	Global Transitions Proceedings 3.1 (2022)	The aim of the paper is to develop an automated system for detecting and classifying plant leaf diseases using computer vision and machine learning algorithms.	54,000 images across multiple plant species	1.Convolutional Neural Networks (CNNs) 2.Support Vector Machines (SVM) 3.K-Nearest Neighbors (KNN)	1.Dataset Dependency 2.Limited Generalization 3.Quality of Input Images 4. Interpretability	95%
19.	Ahmad, Wakeel, Syed M. Adnan, and Aun Irtaza. "Local triangular-ternary pattern: a novel feature descriptor for plant leaf disease detection"	Multimedia Tools and Applications 83.7 (2024)	The aim of the paper is to introduce the Local Triangular-Ternary Pattern (LTP) It seeks to enhance classification accuracy and robustness by capturing subtle texture variations in leaf images.	54,000 images across various plant species and diseases	1.Local Triangular-Ternary Pattern (LTP) 2.Support Vector Machines (SVM), 3.Random Forest	1.Limited Generalization 2. Noise Sensitivity 3. Interpretability	Varies from 94.50%-97.80%
20.	Zou, Fendong, et al."ECVNet:	Agronomy 14.12	The aim of the paper is to	**custom dataset**	1.Efficient Convolutiona	1.Complexity and	96.35%

Journal of Information Systems Engineering and Management
2025, 10(59s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

	A Fusion Network of Efficient Convolutional Neural Networks and Visual Transformers for Tomato Leaf Disease Identification”	(2024)	develop a fusion network combining Efficient CNNs and Visual Transformers for improved tomato leaf disease identification	containing **4,000 tomato leaf images**	l Neural Networks (EfficientCNNs) 2.Visual Transformers (ViTs)	Computational Demand 2.Sensitivity to Image Quality	
21.	Sun, Yubing, et al. “Tomato Leaf Disease Classification by Combining EfficientNetv2 and a Swin Transformer”	Applied Sciences 14.17 (2024)	The paper aims to enhance tomato leaf disease classification by combining **EfficientNetv2** and the **Swin Transformer**, leveraging their strengths in local feature extraction and global context modeling.	Tomato Leaf Disease Dataset(18,160 images) and Tomato Pests Dataset(4,263 images)	1.EfficientNet 2.Swin Transformer	1.Increased Complexity 2.Training Time and Resource Consumption 3.Interpretability Issue	92.70%
22.	Li, Zhe, et al. “Image denoising algorithm based on gradient domain guided filtering and NSST”	IEEE Access 11 (2023)	The paper aims to enhance image denoising by combining gradient domain guided filtering with non-sub sampled shearlet transform (NSST)	5 to 30 images for experimental validation	1.Gradient Domain Guided Filtering 2.Weighted kernel norm minimization (WNNM) 3.Non-subsampled shearlet transform (NSST)	1.Parameter Tuning 2. Performance in Different Noise Levels 3.Computational Complexity	-NA-
23.	Zhang, Liming, et al.”An infrared and visible image fusion algorithm based on ResNet- 152”	Multimedia Tools and Applications 81.7 (2022)	To develop a robust image fusion technique that combines the complementary information from infrared and visible	100 infrared and visible image pairs	1.ResNet-152	1.Complexity of Hyper parameter Tuning 2.Limited Generalization 3. Fusion	97.07%

Journal of Information Systems Engineering and Management
2025, 10(59s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

			images			Quality in Challenging Conditions	
24.	Battiloro, Claudio, et al. "Generalized simplicial attention neural networks"	IEEE Transactions on Signal and Information Processing over Networks (2024)	to develop a novel attention mechanism, Generalized Simplicial Attention (GSA), that extends traditional attention models to simplicial complexes	-NA-	1.Generalized Simplicial Attention (GSA) 2.Neural Network Architecture 3.Self-Attention Mechanism	1.Limited Practical Datasets 2.Adaptability to Non-Simplicial Data 3. Scalability	89%
25.	Liu, Junhua, et al. "Hybrid quantum-classical convolutional neural networks"	Science China Physics, Mechanics & Astronomy 64.9 (2021)	This hybrid approach enhances tasks like image processing and feature extraction by leveraging quantum speedup alongside traditional CNN capabilities.	1.MNIST contains 70,000 images 2.CIFAR-10 consists of 60,000 images 3.CIFAR-100 has 60,000 images 4.Fashion MNIST includes 70,000 images	1.Classical Convolutional Neural Networks (CNNs) 2.Quantum Circuits	1.Quantum Hardware Limitations 2.Integration Complexity 3.Training Challenges 4.Resource Intensive 5.Quantum Data Encoding	90%

Summary of Comparative Analysis:

The reviewed studies from 2022 to 2024 highlight a variety of AI and image processing approaches for tomato leaf disease detection, each with distinct goals, datasets, techniques and limitations. Techniques ranged from classical methods like GLCM and SVM to advanced deep learning architectures such as CNNs, YOLO, FC-SNDPN, and LMBRNet, often enhanced by transfer learning and ensemble methods. Datasets used varied widely in size—from a few hundred to over 50,000 images—primarily sourced from PlantVillage or collected in-field. While models like the one proposed by Anandhakrishnan and Jaisakthi (2022) achieved up to 98.3% accuracy using CNNs, and

Kaur et al. (2022) reached 97.6% with object detection, many studies faced common drawbacks such as high computational demands, dependency on large and high-quality datasets, limited disease coverage, generalization issues, and overfitting risks. Despite these challenges, the research collectively demonstrates significant progress in automating tomato disease detection, with a trend toward developing accurate, scalable and deployable solutions using AI.

The integration of advanced artificial intelligence (AI) and computer vision techniques for early and accurate detection of plant diseases, particularly in tomato and corn crops. Techniques like CNNs, SVMs, Random Forests, Vision Transformers (e.g., DeiT, Swin), and hybrid models such as EfficientNet-ViT combinations are frequently employed, often achieving high accuracy (89–97.8%). Challenges persist in terms of generalization, interpretability, dependence on high-quality datasets, computational demands, and robustness under varying environmental conditions. Some studies focus on semantic segmentation, disease severity classification and explainable AI to enhance usability and decision-making in real agricultural settings, while others explore novel descriptors (e.g., LTP) and quantum-classical networks for performance boosts.

7. COMPLEXITY ANALYSIS

Table 3: Comparative complexity analysis

Technique	Dataset	Accuracy	Precision	Recall	F1 Score	Algorithmic Complexity	Computational Demand	Interpretability	Real-World Applicability
Decision Tree	Plant Village (Labeled)	80%	79	80	80	O(n log n)	Low (CPU-based)	High	Moderate
Random Forest	Plant Village	92%	90	91	90	O(n log n) × t (t = trees)	Medium	Medium	High
SVM	Plant Village	84%	84	84	84	$O(n^2)$ (non-linear kernels)	Medium to High	Medium	High
KNN	Plant Village	90%	91	91	91	$O(n^2)$	High (Memory Intensive)	High	Moderate
CNN with VGG16	Tomato Leaves	92%	91	92	91	$O(n^3)$	Very High (GPU required)	Low	High
ResNet50	Tomato Leaves	94%	93	94	93	$O(n^2 \cdot \log n)$	High (Optimized architecture)	Medium	High
InceptionV3	Tomato Leaves	90%	91	91	91	$O(n^2 \cdot \log n)$	High	Medium	High

The comparative complexity analysis in table 3 reveals that **deep learning models** like **ResNet50**, **VGG16** and **InceptionV3** achieve **higher accuracy (90–94%)** and F1 scores (91–93),** making them highly suitable for real-world tomato leaf disease detection despite their **high computational demands** and **lower interpretability**. In contrast, **traditional machine learning models** such as **Decision Trees**, **Random Forests**, **SVM** and **KNN** offer better interpretability and lower computational costs but generally yield **lower accuracy (80–92%)**. Among them, **Random Forests** and **KNN** stand out for balancing performance and feasibility, though they require feature engineering or face efficiency issues at scale. Overall, **deep learning models are preferred for accuracy and scalability**, while traditional models offer **transparency and efficiency** on smaller or simpler datasets.

Threatening hurdles:

- One of the most demanding tasks would be building solid models that can generalize well with these different variations.
- Collecting huge labelled datasets of images showing plant diseases is quite tedious and costly.

Journal of Information Systems Engineering and Management
2025, 10(59s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

- Building efficient and lightweight models deployable on resource-starved devices is quite necessary for the real world.
- Detecting diseases in the initial stages is extremely difficult; the symptoms are not easily perceptible or may come under the healthy plant variations.

Future prospects:

- For future research, it intends to conceive improved and generalized models that may function under real-life variations.
- Investigating new data augmentation techniques, synthetic data generation and few-shot learning methods to reduce the dependency on data.
- Designing efficient and lightweight models that could be deployed on mobile devices and all the other edge devices.
- Enhancing the interpretability and explainability of DL models to provide insights into their decision-making process.

8. CONCLUSION

Machine learning and deep learning exhibit great potential for automation and improvement in early detection of tomato plant diseases. Deep learning methods, especially CNNs are capable of effectively classifying and localizing diseases in leaf images. In this survey, we have reviewed the key methodologies, recent developments and challenges in the field. Addressing these challenges and the promising research routes would allow for the development of robust, efficient, and practical solutions for early disease detection for sustainable agriculture and the security of food worldwide.

REFERENCES:

1. Rahman, Sami Ur, et al. "Image processing based system for the detection, identification and treatment of tomato leaf diseases." Multimedia Tools and Applications 82.6 (2023):9431-9445.

2. Nazari, Kobra, Mohammad Javad Ebadi, and Kamal Berahmand. " Diagnosis of alternaria disease and leafminer pest on tomato leaves using image processing techniques." Journal of the Science of Food and Agriculture 102.15 (2022): 6907-6920.

3. Huang, Xibei, et al. "Tomato leaf disease detection system based on FC- SNDPN." Multimedia tools and applications 82.2 (2023): 2121-2144.

4. Sanida, Theodora, et al. "Tomato leaf disease identification via two–stage transfer learning approach." Smart Agricultural Technology 5 (2023): 100275.

5. Li, Mingxuan, et al. "Identification of tomato leaf diseases based on LMBRNet." Engineering Applications of Artificial Intelligence 123 (2023): 106195.

6. Javidan, Seyed Mohamad, et al. "Tomato leaf diseases classification using image processing and weighted ensemble learning." Agronomy Journal 116.3 (2024): 1029-1049.

7. Djimeli-Tsajio, Alain B., et al. "Improved detection and identification approach in tomato leaf disease using transformation and combination of transfer learning features." Journal of Plant Diseases and Protection 129.3 (2022): 665-674.

8. Kaur, Prabhjot, et al. "An approach for characterization of infected area in tomato leaf disease based on deep learning and object detection technique." Engineering Applications of Artificial Intelligence 115 (2022): 105210.

Journal of Information Systems Engineering and Management
2025, 10(59s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

9. Anandhakrishnan, T., and S. M. Jaisakthi. "Deep Convolutional Neural Networks for image based tomato leaf disease detection." Sustainable Chemistry and Pharmacy 30 (2022): 100793.

10. Özbılge, Emre, et al. "Tomato disease recognition using a compact convolutional neural network." IEEE Access 10 (2022): 77213-77224.

11. da Cunha, Vitor A. Gontijo, et al. "Early detection of tomato bacterial spot disease in transplant tomato seedlings utilising remote sensing and artificial intelligence." Biosystems Engineering 234 (2023): 172-186.

12. Khan, Saiqa, and Meera Narvekar. "Novel fusion of color balancing and superpixel based approach for detection of tomato plant diseases in natural complex environment." Journal of King Saud University-Computer and Information Sciences 34.6 (2022): 3506-3516.

13. Khan, Khalil, et al. "End-to-End Semantic Leaf Segmentation Framework for Plants Disease Classification." Complexity 2022.1 (2022): 1168700.

14. Pal, Arunangshu, and Vinay Kumar. "AgriDet: Plant Leaf Disease severity classification using agriculture detection framework." Engineering Applications of Artificial Intelligence 119 (2023): 105754.

15. Rajpal, Ankit, et al. "Explaining deep learning-based leaf disease identification." Soft Computing (2024): 1-24.

16. Yeswanth, P. V., and S. Deivalakshmi. "ASFESRN: bridging the gap in real-time corn leaf disease detection with image super-resolution." Multimedia Systems 30.4 (2024):175.

17. Sun, Changxia, et al. "Research on tomato disease image recognition method based on DeiT." European Journal of Agronomy 162 (2024): 127400.

18. Harakannanavar, Sunil S., et al. "Plant leaf disease detection using computer vision and machine learning algorithms." Global Transitions Proceedings 3.1 (2022): 305-310.

19. Ahmad, Wakeel, Syed M. Adnan, and Aun Irtaza. "Local triangular-ternary pattern: a novel feature descriptor for plant leaf disease detection." Multimedia Tools and Applications 83.7 (2024): 20215-20241.

20. Zou, Fendong, et al. "ECVNet: A Fusion Network of Efficient Convolutional Neural Networks and Visual Transformers for Tomato Leaf Disease Identification." Agronomy 14.12 (2024): 2985.

21. Sun, Yubing, et al. "Tomato Leaf Disease Classification by Combining EfficientNetv2 and a Swin Transformer." Applied Sciences 14.17 (2024): 7472.

22. Li, Zhe, et al. "Image denoising algorithm based on gradient domain guided filtering and NSST." IEEE Access 11 (2023): 11923-11933.

23. Zhang, Liming, et al. "An infrared and visible image fusion algorithm based on ResNet-152." Multimedia Tools and Applications 81.7 (2022): 9277-9287.

24. Battiloro, Claudio, et al. "Generalized simplicial attention neural networks." IEEE Transactions on Signal and Information Processing over Networks (2024).

25. Liu, Junhua, et al. "Hybrid quantum-classical convolutional neural networks." Science China Physics, Mechanics & Astronomy 64.9 (2021): 290311