Artificial Intelligence and Machine Learning Approaches for Plant Disease Detection: A Comprehensive Review of Image, Sensor, IoT and VOC-Based Techniques

Jayprakash H. Pardeshi[1], Dr. Hari V. Dube[2]

[1]Assistant Professor, Department of Computer Science, Padmashri Vikhe Patil College of Arts, Science and Commerce, Pravaranagar, Maharashtra, Email ID: jayprakashpardeshi@gmail.com

[2]Assistant Professor, Department of Computer Science, Padmashri Vikhe Patil College of Arts, Science and Commerce, Pravaranagar, Maharashtra, Email ID: dubehvd72@gmail.com

Abstract

Plant diseases are a major constraint to agricultural productivity, crop quality and food security. Conventional diagnosis is commonly based on visual inspection by farmers or experts and, when required, laboratory techniques such as microscopy, culturing and molecular assays. These approaches can be laborious, time-consuming, subjective and difficult to scale for continuous field monitoring. Recent advances in Artificial Intelligence (AI), Machine Learning (ML), Deep Learning (DL), computer vision, Internet of Things (IoT), chemical sensing and electronic-nose technology have created new opportunities for automated, rapid and non-destructive plant disease detection. Classical ML algorithms such as Support Vector Machine (SVM), Random Forest (RF), k-Nearest Neighbour (kNN), Decision Tree and Artificial Neural Network (ANN) have been used with engineered image and sensor features. DL models, including Convolutional Neural Networks (CNNs), ResNet, MobileNet, EfficientNet, YOLO and Vision Transformers (ViTs), can learn discriminative representations directly from images. Recent systematic reviews show rapid growth in this field but also identify important limitations related to dataset diversity, environmental variability, class imbalance, model generalisation and the gap between laboratory and field performance. In parallel, electronic noses and volatile organic compound (VOC) sensors provide a complementary chemical-sensing modality that can detect changes associated with plant stress and disease. This review examines image-based ML/DL, VOC and E-nose approaches, IoT integration, evaluation metrics, current challenges and future research directions. Particular attention is given to multimodal AI that combines plant images, VOC signatures and environmental parameters for early tomato disease detection. The review identifies multimodal sensing, field validation, sensor-drift compensation, explainable AI and lightweight edge deployment as important research opportunities.

Keywords: Artificial Intelligence; Machine Learning; Deep Learning; Plant Disease Detection; Computer Vision; Electronic Nose; Volatile Organic Compounds; IoT; Tomato Disease; Precision Agriculture.

1. Introduction

Agricultural production is increasingly challenged by fungal, bacterial, viral and other plant diseases. Disease can reduce yield, quality and market value and may require rapid intervention to limit spread. The conventional approach of visual examination remains useful, but its accuracy depends on the experience of the observer and on the visibility of symptoms. Early infection may not produce obvious symptoms, while different diseases or abiotic stresses can produce similar visual patterns. Laboratory diagnosis can improve confidence, but it generally requires sampling, transport, equipment and trained personnel.

AI has emerged as an important technology for automating disease recognition. Image-based systems can acquire leaf or fruit images using smartphones, cameras or UAVs and use ML/DL models to classify healthy and diseased samples. A 2024 systematic review by Pacal et al. analysed 160 studies published from 2020 to 2024 and examined disease classification, object detection, segmentation, datasets and challenges. [1] A 2025 review by Upadhyay et al. further surveyed computer vision and DL techniques, including CNNs, transformers and other modern architectures, while emphasizing practical issues in precision agriculture. [2] Recent reviews continue to identify generalisation, small or imbalanced datasets, environmental noise and real-world field performance as major concerns. [3–5]

A limitation of purely image-based diagnosis is that it generally relies on visible or detectable symptoms. Plant biochemical responses to pathogens can alter emissions of volatile organic compounds (VOCs). VOC monitoring therefore offers a complementary route to disease diagnosis. A 2023 review in Biosensors and Bioelectronics discussed VOC-monitoring technologies and noted both their diagnostic potential and challenges in practical deployment. [6] Electronic noses use arrays of partially selective chemical sensors and pattern-recognition algorithms to distinguish complex odour patterns. [7,8] Combining image, VOC and environmental data could therefore move plant disease detection from single-modality classification toward multimodal and potentially earlier diagnosis.

2. Review Methodology

This review is structured using the principles of transparent systematic-review reporting. PRISMA 2020 provides a recognised framework for reporting why a review was performed, how evidence was identified and selected, and how results were synthesised. [9] For the present review, literature was organised around five themes: (i) conventional diagnosis, (ii) image-based ML/DL, (iii) VOC and electronic-nose sensing, (iv) IoT and edge-AI integration, and (v) multimodal disease detection.

The search scope emphasises recent literature from 2020–2026 while retaining selected foundational studies where necessary. Representative search terms include “plant disease detection”, “machine learning plant disease”, “deep learning plant disease”, “computer vision agriculture”, “electronic nose plant disease”, “volatile organic compounds plant disease”, “VOC sensors agriculture”, “IoT plant disease detection” and “multimodal plant disease detection”. Priority was given to peer-reviewed journal and conference literature and recent systematic reviews. The review does not claim a formal quantitative meta-analysis because datasets, crops, sensors, evaluation protocols and target diseases vary substantially across studies.

3. Conventional Plant Disease Detection

Conventional diagnosis can be divided into visual diagnosis and laboratory diagnosis. Visual diagnosis considers leaf colour, lesions, spots, wilting, necrosis, powdery growth, fruit abnormalities and other symptoms. Its major advantages are low cost and immediate availability. However, performance varies with observer expertise, lighting, plant variety and symptom stage.

Laboratory diagnosis may include microscopy, pathogen isolation, immunological assays and molecular methods such as PCR. Such methods can provide strong evidence for pathogen identification but are less convenient for continuous field monitoring. These limitations motivate automated sensing technologies that can operate closer to the point of production.

4. Machine Learning for Plant Disease Detection

A conventional ML pipeline typically consists of image acquisition, preprocessing, segmentation, feature extraction, feature selection and classification. Preprocessing may include resizing, denoising, contrast enhancement, colour normalisation and augmentation. Segmentation separates the plant or diseased region from the background. Engineered features may describe colour, texture and shape.

4.1 Image Features

Colour: RGB, HSV, Lab and normalised colour indices.
Texture: contrast, entropy, homogeneity, energy and GLCM-derived descriptors.
Shape: area, perimeter, circularity, major axis and minor axis.
Lesion characteristics: number, size, distribution and colour variation.

4.2 Classical ML Algorithms

Algorithm	Strength	Typical use	Limitation
SVM	Strong classification with small/high-dimensional datasets	Image or sensor features	Kernel and parameter selection
Random Forest	Robust, nonlinear, feature importance	Image and sensor/tabular data	Large ensembles can be computationally heavier
kNN	Simple and intuitive	Small feature datasets	Prediction cost increases with dataset size
ANN	Learns nonlinear relationships	Image-derived or sensor data	Needs tuning and representative training data
XGBoost	Strong structured-data performance	Sensor/environmental features	Hyperparameter tuning required

5. Deep Learning and Computer Vision

Deep learning reduces dependence on manually engineered features by learning hierarchical representations directly from data. CNNs have become the dominant family for image-based plant disease classification. A typical CNN contains convolutional layers for local feature extraction, nonlinear activation functions, pooling or downsampling, and classification layers. Transfer learning allows a model pretrained on a large image dataset to be adapted to a crop-specific dataset, which can be valuable when agricultural datasets are relatively small.

5.1 CNN and Transfer Learning

CNNs can learn edges, textures, lesion patterns and higher-level disease representations. Transfer learning using architectures such as VGG, ResNet, DenseNet, EfficientNet and MobileNet is widely used. A 2025 review specifically examined CNN-based plant leaf disease detection and reported strong use of VGG, EfficientNet, GoogleNet and ResNet families. [10]

5.2 ResNet

ResNet uses residual connections to improve optimisation of deeper networks. Its ability to learn complex representations makes it useful for fine-grained disease classification, although computational requirements increase with model size.

5.3 MobileNet and EfficientNet

MobileNet uses lightweight operations and is attractive for mobile and edge applications. EfficientNet scales depth, width and input resolution in a coordinated way. These architectures are particularly relevant when a disease detector must operate on resource-constrained devices.

5.4 YOLO and Object Detection

Classification assigns a label to an image, whereas object detection can locate diseased regions. YOLO-family detectors are therefore useful when several leaves or disease lesions occur in a single image. Detection is valuable for field photographs where the background is uncontrolled and symptoms are spatially distributed.

5.5 Vision Transformers

Vision Transformers model relationships among image patches using attention mechanisms. Recent reviews identify ViTs, hybrid CNN–ViT models and self-supervised learning as emerging approaches for improving robustness and reducing dependence on extensive manual feature engineering. [3–5]

6. Image-Based Disease Detection: Advantages and Limitations

Image-based detection is non-destructive, fast and relatively inexpensive when RGB cameras or smartphones are used. It can support mobile applications, greenhouse monitoring, UAV surveys and automated crop scouting. However, image models are affected by illumination, shadows, background clutter, camera distance, leaf orientation, cultivar differences and disease severity. A model trained on clean laboratory images may therefore fail when applied to field images. Recent literature repeatedly identifies the laboratory-to-field domain gap as a central challenge. [2–5]

Approach	Main benefit	Major concern
RGB + classical ML	Low-cost, interpretable feature pipeline	Manual feature engineering
YOLO/object detection	Localises disease regions	Needs bounding-box training data
Multispectral/hyperspectral	Additional spectral information	Higher cost and data complexity
Vision Transformer	Long-range feature relationships	Data/computation requirements

7. Datasets and Generalisation

Dataset quality is as important as model architecture. Public datasets have enabled rapid progress, but many commonly used datasets contain images collected under controlled conditions. Problems include limited disease classes, class imbalance, repeated visual patterns, restricted cultivars, laboratory backgrounds and insufficient geographic diversity. Recent reviews specifically emphasise that high reported accuracy does not guarantee robust field performance. [1–5]
A stronger evaluation strategy should separate training, validation and test samples at the plant or field level rather than allowing near-duplicate images to appear across splits. External field testing is especially important. Performance should be reported across different lighting conditions, growth stages, cultivars and locations wherever possible.

8. Electronic Nose and VOC-Based Plant Disease Detection

Plants emit complex mixtures of volatile organic compounds. Biotic and abiotic stresses can change the quantity and composition of emitted VOCs. This makes VOC analysis a promising non-destructive sensing route. An electronic nose attempts to recognise patterns in complex vapours using an array of chemical sensors and pattern-recognition algorithms rather than identifying every compound individually. Reviews describe applications in plant disease and pest detection and highlight rapid response and non-invasive operation as important advantages. [6–8,11]
The basic E-nose workflow is: plant or headspace sampling → sensor array exposure → signal acquisition → preprocessing → feature extraction → pattern recognition → disease classification. Sensor arrays can produce multidimensional responses such as baseline resistance, maximum response, response time, recovery time and integrated response area. These features can be combined with environmental measurements to compensate for operating-condition variability.

8.1 VOC Sensors and Sensor Arrays

Common low-cost metal-oxide gas sensors can respond to broad classes of gases, while more selective sensor technologies may be used where chemical specificity is required. For research prototypes, an array with different sensing characteristics can produce a disease-specific response pattern. The goal is not necessarily to map each sensor to a single VOC; rather, the collective response can be treated as a fingerprint for ML classification.

8.2 ML for VOC Patterns

PCA is useful for visualising whether healthy and diseased samples occupy different regions of feature space. LDA can maximise separation among labelled classes. SVM and Random Forest can be used for supervised classification. ANN and XGBoost can model nonlinear relationships. CNNs can also be applied to structured sensor-response matrices or temporal signals when sufficient data are available.

8.3 Important E-Nose Challenges

Sensor drift caused by ageing, contamination and long-term operation.

Temperature and relative humidity dependence.
Sampling-chamber design and headspace equilibration.
Variation caused by plant age, cultivar, nutrition and environmental stress.
Limited crop-specific and disease-confirmed VOC datasets.
Need for field validation and standardised sampling protocols.
A 2025 review on E-Nose technology for early plant disease detection highlights the promise of rapid, non-invasive diagnosis while also identifying sensor drift, standardisation and field-level validation as important unresolved issues. [11]

9. IoT-Based Plant Disease Detection

IoT can provide continuous acquisition and communication of plant and environmental data. A typical architecture contains sensors, a microcontroller such as an ESP32, wireless connectivity, an edge or cloud processing layer and a user interface. Environmental parameters may include temperature, relative humidity, soil moisture, leaf wetness and light intensity. IoT and deep learning have been reviewed as complementary technologies for plant disease detection, with IoT supporting data acquisition and communication and DL supporting automated pattern recognition. [12]
Edge AI is especially relevant where internet connectivity is unreliable or where immediate alerts are required. Lightweight models can reduce latency and data-transfer requirements. However, model compression, memory constraints, energy consumption and hardware-specific optimisation must be considered during system design.

10. Multimodal AI: Image + VOC + Environment

A key research direction is to combine complementary sensing modalities. Images describe visible symptoms; VOC sensors describe chemical responses; environmental sensors provide context. The combined feature vector can be expressed conceptually as X = [X_image, X_VOC, X_environment]. Fusion may occur at the feature level, intermediate representation level or decision level.
Multimodal fusion may reduce weaknesses associated with a single sensing modality. For example, an image may be ambiguous when two diseases have similar lesions, while VOC responses may provide an additional discriminative signal. Conversely, VOC signals may be influenced by humidity, and image information can provide a useful independent measurement. The value of fusion must be demonstrated experimentally through ablation studies comparing each modality with the combined model.

11. Application to Tomato Disease Detection

Tomato is a suitable crop for evaluating multimodal disease detection because it is affected by several economically important diseases and because leaves, fruits and stems can provide image and chemical signals. A research system can target healthy plants and selected diseases such as early blight, late blight, bacterial spot, leaf mold and Septoria leaf spot, provided that disease labels are confirmed using an appropriate plant-pathology protocol.
For a low-cost prototype, an E-nose can use an eight-sensor array with different gas-response characteristics, together with temperature and humidity sensors. An ESP32 can acquire sensor data and transmit them to a local computer or cloud service. A camera can capture the corresponding plant image. Each sample should be assigned a unique ID so that the VOC record, environmental record and image are synchronised.

11.1 Proposed Experimental Pipeline

Select healthy and confirmed diseased tomato plants.
Standardise sampling time, plant age range and sampling chamber conditions.
Capture RGB images before or after VOC sampling using a consistent protocol.
Measure VOC sensor-array responses together with temperature and relative humidity.
Repeat measurements across days and plants to capture biological variability.
Preprocess sensor signals, remove unstable initial readings and derive response features.
Train and compare SVM, Random Forest, XGBoost and selected neural models.
Train image models such as MobileNet, ResNet or EfficientNet using field-oriented images.
Perform feature/decision fusion of image, VOC and environmental information.
Validate the final model on plants and field conditions not represented in training.

12. Performance Evaluation

Accuracy alone can be misleading, particularly for imbalanced disease datasets. Precision, recall, F1-score, sensitivity, specificity and confusion matrices should be reported. For binary classification, accuracy is defined as (TP + TN)/(TP + TN + FP + FN), precision as TP/(TP + FP), recall as TP/(TP + FN), and F1 as the harmonic mean of precision and recall. Multi-class macro-averaged and weighted metrics should be reported where appropriate.
For VOC systems, evaluation should additionally include PCA/LDA visualisations, sensor repeatability, response and recovery characteristics, drift over time, environmental sensitivity, sampling duration and robustness to day-to-day variation. For edge-AI systems, model size, inference time, memory use and power consumption are also relevant.

Metric	Purpose	Recommended use
Accuracy	Overall correct predictions	Balanced datasets
Precision	Reliability of positive predictions	Disease alert systems

Recall/Sensitivity	Ability to find diseased samples	Early warning
F1-score	Balance between precision and recall	Imbalanced datasets
Specificity	Ability to reject healthy samples	False-alarm control
ROC-AUC	Ranking/separation ability	Binary or one-vs-rest analysis
Inference time	Real-time suitability	Edge/mobile deployment

13. Research Gaps

The literature indicates several opportunities. First, much research remains image-centric, whereas fewer systems integrate chemical and visual signals in a single experimentally validated dataset. Second, the majority of benchmark datasets do not fully represent field variability. Third, VOC-based diagnosis needs better standardisation of sampling chambers, flow rates, environmental compensation and sensor calibration. Fourth, pre-symptomatic detection should be tested using time-series experiments rather than inferred solely from a classifier's accuracy. Fifth, disease detection should be distinguished from abiotic stress and nutrient deficiency, which can alter both appearance and VOC emissions.

Another important gap is the lack of longitudinal datasets. A plant monitored from healthy status through early infection and visible disease would allow researchers to determine how early a model can detect a change. Such time-series data could support temporal models and provide a scientifically stronger definition of “early detection”.

Research gap	Why it matters	Potential solution
Controlled-to-field domain gap	High lab accuracy may not generalise	Field-diverse datasets and external validation
Sensor drift	VOC model performance changes over time	Calibration, reference samples and drift compensation
Environmental interference	Temperature/RH alter sensor response	Record environmental variables and compensate in ML
Limited VOC datasets	Difficult to learn reliable fingerprints	Longitudinal, disease-confirmed datasets
Visible-symptom dependence	May miss early infection	Combine VOC with image and time-series data
Explainability	Farmers need interpretable decisions	Grad-CAM/XAI and sensor-feature importance

14. Future Research Directions

Multimodal AI combining RGB images, VOC sensor arrays and environmental measurements.
Longitudinal experiments from healthy plants to early and advanced disease stages.
Lightweight CNN/ViT models for mobile and edge devices.
Sensor-drift compensation and adaptive calibration.
Explainable AI using visual attention maps and sensor-feature importance.
Self-supervised and few-shot learning for limited agricultural datasets.
Field-scale datasets covering cultivars, locations, seasons and environmental conditions.
Integration of disease severity estimation with disease classification.
Real-time IoT dashboards and farmer alerts.
Standardised VOC sampling and disease-confirmation protocols.

The strongest opportunity for electronics- and IoT-oriented research is to develop a low-cost multimodal platform in which an ESP32 or similar edge controller synchronises an E-nose array, environmental sensors and an RGB camera. AI models can then be compared at three levels: VOC-only, image-only and fused multimodal prediction. This experimental comparison would provide a clear contribution and would avoid claiming that multimodal fusion is superior without evidence.

15. Proposed Multimodal Framework for Tomato Disease Detection

The proposed framework consists of five layers. Layer 1 is sensing, containing an RGB camera, eight VOC/gas sensors and environmental sensors. Layer 2 is acquisition, using a microcontroller and a standardised sampling chamber. Layer 3 is preprocessing and feature extraction, including image normalisation, VOC baseline correction and environmental compensation. Layer 4 is AI inference, where image and sensor models are trained independently and then fused. Layer 5 is decision support, producing disease class, confidence, severity estimate and an early-warning message.

Conceptual flow: Tomato plant → RGB image + VOC/E-nose + temperature/RH/leaf-wetness → ESP32/data acquisition → preprocessing → feature extraction → multimodal fusion → ML/DL classifier → disease and confidence → IoT dashboard/alert.

The research should include ablation experiments: (A) VOC only, (B) image only, (C) environmental data only, (D) VOC + environment, (E) image + VOC, and (F) image + VOC + environment. If the fused model consistently improves external-test performance, the study can provide strong evidence for multimodal AI.

16. Conclusion

AI and ML have transformed plant disease detection from manual inspection toward automated, data-driven diagnosis. Classical ML remains useful for engineered image and sensor features, while DL provides powerful representation learning for image classification, detection and segmentation. Nevertheless, real-world deployment is constrained by dataset bias, environmental variability, disease similarity, computational limitations and the laboratory-to-field domain gap. Recent systematic reviews consistently identify these challenges. [1–5]

Electronic-nose and VOC sensing provide a complementary chemical perspective. VOC monitoring can capture physiological responses associated with plant stress and disease, but practical deployment requires careful attention to sensor drift, environmental effects, sampling protocol and field validation. [6,11] The most promising direction is therefore not simply replacing image diagnosis with VOC sensing, but integrating multiple modalities. For tomato disease detection, the combination of RGB images, VOC signatures and environmental parameters can form a strong research framework for investigating early and robust disease diagnosis. Future studies should use longitudinal, disease-confirmed field datasets and directly compare unimodal and multimodal models. Such work can connect electronics, sensors, IoT and AI with practical precision agriculture.

References

1. I. Pacal, I. Kunduracioglu, M. H. Alma, M. Deveci, S. Kadry, J. Nedoma, V. Slany, R. Martinek et al., “A systematic review of deep learning techniques for plant diseases,” Artificial Intelligence Review, vol. 57, art. 304, 2024. DOI: 10.1007/s10462-024-10944-7.
2. A. Upadhyay, N. S. Chandel, K. P. Singh, S. K. Chakraborty, B. M. Nandede, M. Kumar, A. Subeesh, K. Upendar, A. Salem et al., “Deep learning and computer vision in plant disease detection: a comprehensive review of techniques, models, and trends in precision agriculture,” Artificial Intelligence Review, vol. 58, art. 92, 2025. DOI: 10.1007/s10462-024-11100-x.
3. T. Nyawose, R. C. Maswanganyi, and P. Khumalo, “A Review on the Detection of Plant Disease Using Machine Learning and Deep Learning Approaches,” Journal of Imaging, vol. 11, no. 10, 326, 2025. DOI: 10.3390/jimaging11100326.
4. Y. Kaya and E. Gürsoy, “A review of deep learning architectures for plant disease detection,” Turkish Journal of Biology, vol. 49, no. 5, pp. 459–497, 2025. DOI: 10.55730/1300-0152.2761.
5. “Revolutionizing plant disease diagnosis through vision-based intelligence and next-generation computing,” Computers & Electrical Engineering, 2025. DOI: 10.1016/j.compeleceng.2025.110695.
6. “Challenges and applications of volatile organic compounds monitoring technology in plant disease diagnosis,” Biosensors and Bioelectronics, vol. 237, 115540, 2023. DOI: 10.1016/j.bios.2023.115540.
7. “Plant Pest Detection Using an Artificial Nose System: A Review,” Sensors, vol. 18, no. 2, 378, 2018.
8. “Advances in Electronic Nose Sensors for Plant Disease and Pest Detection,” Engineering Proceedings, vol. 48, 14, 2023.
9. M. J. Page et al., “The PRISMA 2020 statement: an updated guideline for reporting systematic reviews,” BMJ, vol. 372, n71, 2021. DOI: 10.1136/bmj.n71.
10. T. D. Salka, M. B. Hanafi, S. M. S. A. A. Rahman, D. B. M. Zulperi et al., “Plant leaf disease detection and classification using convolution neural networks model: a review,” Artificial Intelligence Review, vol. 58, art. 322, 2025.
11. S. P., Heyram, S. P. K., P. S. Kurian, B. A. S., and L. Sabu, “E-NOSE Technology: An Eminent Tool for Early Detection of Plant Diseases,” Journal of Scientific Research and Reports, vol. 31, no. 10, pp. 379–396, 2025. DOI: 10.9734/jsrr/2025/v31i103580.
12. “Role of Internet of Things and Deep Learning Techniques in Plant Disease Detection and Classification: A Focused Review,” Sensors, vol. 23, 7877, 2023.