

Article – e010

A Review of Machine Learning and Deep Learning Techniques for Tomato Leaf Disease Classification

Sahil Sher¹✉, **Kunal Batra²**✉

^{1,2} *Institute of Engineering & Technology, Sage University Indore Indore (M.P.), India*

Received: 2026-08-29

Revision Received: 2026-09-24

Accepted: 2026-10-12

ABSTRACT

Tomato is an economically important crop, and disease infection can reduce plant health, productivity, and fruit quality. Image-based machine learning (ML) and deep learning (DL) methods have therefore been investigated for automated recognition of tomato leaf diseases. This narrative review examines the progression from handcrafted feature-based ML methods to convolutional neural networks (CNNs), transfer learning, lightweight architectures, attention mechanisms, and recent model-compression and multi-task approaches. The review considers the data characteristics, preprocessing strategies, learning paradigms, evaluation conditions, computational requirements, and evidence of generalization reported by representative studies. The literature shows that conventional ML can perform well on small, carefully prepared datasets but depends on the quality of handcrafted representations. CNN-based methods learn task-specific visual representations and generally provide stronger performance on multi-class tomato leaf image datasets. However, very high reported accuracies are frequently obtained on controlled or curated datasets, particularly PlantVillage-derived data, and therefore should not be interpreted as direct evidence of field performance. Studies using field images or explicitly measuring model size, FLOPs, inference speed, or mobile/edge execution provide stronger evidence concerning practical deployment. The principal research gaps are limited field diversity, illumination and background variation, class imbalance, weak cross-dataset validation, inconsistent evaluation protocols, and insufficient reporting of deployment metrics. Future research should therefore evaluate not only classification accuracy but also robustness, independent validation, computational cost, and reproducible deployment conditions.

KEYWORDS: Tomato Leaf Disease, Image Classification, Machine Learning, Deep Learning, Convolutional Neural Network, Transfer Learning, Lightweight Networks, Field Generalization.

I. INTRODUCTION

Tomato disease recognition from images has developed from conventional image-processing pipelines into a broad research area involving ML, DL, and computer vision. Early systems commonly isolated the leaf or diseased region, extracted color, texture, shape, wavelet, or spectral descriptors, and then applied a classifier. Mokhtar et al., for example, combined Gabor-wavelet features with SVM kernels for two tomato diseases and reported 99.5% accuracy on a small dataset containing 100 images for each disease [1]. Such results demonstrate that handcrafted representations can be effective when the visual problem is constrained, but they also illustrate why results obtained from small and controlled datasets should not be generalized to field conditions without additional validation.

Deep learning changed the workflow by learning discriminative representations directly from images. Fuentes et al. developed an object-detection system for tomato diseases and pests using Faster R-CNN, R-FCN, and SSD with VGG and ResNet feature extractors. Their approximately 5,000-image dataset

was collected from farms in different areas of the Korean Peninsula and contained nine annotated disease/pest categories plus a background category [2]. This study is important because it differs from leaf-only classification: it addressed recognition and localization in complex scenes rather than assigning a single class to a cropped leaf.

Subsequent studies increasingly focused on image classification. Karthik et al. evaluated residual and attention-enhanced residual CNNs using PlantVillage tomato images representing early blight, late blight, and leaf mold; the attention-based residual network achieved 98% accuracy under five-fold cross-validation [3]. Trivedi et al. reported 98.49% accuracy using a CNN-based system for ten tomato leaf categories, including healthy leaves and nine disease categories [4]. These studies illustrate the improvement in learned representation, but they also highlight a recurring issue in the literature: the numerical value of accuracy depends strongly on the dataset, class definition, preprocessing, and validation design.

The recent literature has increasingly addressed efficiency. Ahmed et al. proposed a MobileNetV2-based lightweight pipeline using illumination correction and runtime augmentation on 18,160 PlantVillage tomato images spanning one healthy and nine disease classes. The reported accuracy was 99.30%, with a 9.60 MB model and 4.87 million floating-point operations [5]. Ullah et al. introduced DTomatoDNet, a 19-learnable-layer network trained on 10,000 images from a ten-class tomato leaf dataset and reported 99.34% accuracy [6]. Zhang et al. further reported 98.71% average accuracy for LDAMNet on a ten-category tomato dataset while reporting 0.91 million parameters and 0.142 GFLOPs [7]. Such studies shift the discussion from accuracy alone toward the performance-efficiency trade-off.

The purpose of this review is therefore not to reproduce or summarize a particular thesis or implementation. Instead, it synthesizes representative published studies and critically examines what their experimental evidence supports, where comparisons are valid, and which limitations remain. In particular, the review distinguishes classification from object detection and segmentation, separates controlled-dataset evidence from field-image evidence, and treats computational efficiency as a measurable property rather than an assumed deployment benefit.

II. REVIEW METHODOLOGY

This work is a narrative literature review rather than a systematic review or meta-analysis. The objective was to identify representative published studies that illustrate the evolution of tomato disease image analysis and to compare their methodological evidence rather than to calculate a pooled effect size. Literature was identified from scholarly sources including IEEE Xplore, ScienceDirect, SpringerLink, PubMed/PMC, Frontiers in Plant Science, and Google Scholar. The review window was January 2015 to April 2026, allowing inclusion of early feature-based approaches as well as recent lightweight and field-oriented studies.

Search terms were combined around four concept groups: (i) crop/task: “tomato leaf disease,” “tomato plant disease,” “tomato disease classification,” “tomato disease detection”; (ii) method: “machine learning,” “SVM,” “CNN,” “ResNet,” “MobileNet,” “attention,” “transfer learning,” “knowledge distillation,” “object detection”; (iii) data: “PlantVillage,” “PlantDoc,” “field dataset,” “real-field”; and (iv) deployment: “lightweight,” “mobile,” “edge,” “FLOPs,” “inference time,” and “model size.”

Studies were included when they (a) addressed tomato disease or pest recognition from images, (b) described a reproducible ML/DL method, and (c) provided identifiable information about the dataset, disease categories, evaluation protocol, or computational characteristics. Review articles were used for contextual synthesis, while original research papers were prioritized for performance claims. Studies were excluded from the comparative performance table when the available publication record did not provide enough information to identify the dataset/class setting or the reported metric unambiguously. Results from different datasets and protocols were not pooled or ranked.

The final comparative synthesis therefore emphasizes evidence quality and experimental context. Accuracy values are reported only as study-specific observations; they are not treated as directly comparable scores across unrelated datasets.

III. TOMATO LEAF IMAGE ANALYSIS: TASKS AND PIPELINES

Tomato disease image analysis contains several distinct computer-vision tasks. Image classification assigns an image, usually a cropped leaf, to one disease or healthy category. Object detection identifies one or more disease or pest regions and supplies their locations, typically with bounding boxes. Segmentation assigns pixels to a leaf, lesion, or disease region. These tasks should not be used interchangeably. Fuentes et al. performed object detection and localization [2], whereas most of the studies reviewed here perform image classification [3]–[7]. A classification model may use segmentation as a preprocessing step without itself being a segmentation model.

A typical classification workflow contains image acquisition, preprocessing, optional region isolation, representation learning or feature extraction, classification, and evaluation. In conventional ML, feature engineering is explicit: color statistics, texture descriptors, shape descriptors, wavelets, or spectral information are calculated before classification. In CNN-based methods, feature extraction and classification are learned jointly. Recent architectures add attention, transfer learning, knowledge distillation, or lightweight convolutional blocks to alter the balance between representational power and computational cost.

Preprocessing remains important because leaf appearance is affected by illumination, camera characteristics, background, orientation, and disease severity. Ahmed et al. explicitly used contrast-limited adaptive histogram equalization to address lighting effects [5]. Zhang et al. used piecewise linear enhancement and oversampling to address feature visibility and class imbalance [7]. These examples show that preprocessing should be analyzed as part of the experimental design rather than treated as a generic preliminary operation.

IV. MACHINE LEARNING APPROACHES

A. Handcrafted Features and SVM

Mokhtar et al. provide a representative early ML study. Their method used Gabor-wavelet features followed by SVM classification with alternative kernels and evaluated powdery mildew and early blight images [1]. The reported 99.5% accuracy is notable, but the experimental setting was small—100 images per disease—and the task involved only two disease categories. The study therefore supports the feasibility of feature-based classification rather than a conclusion that SVM is superior to later deep networks.

The main strength of this class of methods is explicit feature construction. Handcrafted descriptors can be computationally manageable and can work with limited data. Their limitation is sensitivity to feature selection and to changes in the visual distribution. When illumination, background, symptom scale, or disease stage changes, a descriptor optimized for one image collection may not transfer reliably. Later deep-learning studies were motivated partly by this representation bottleneck [3], [5].

B. Conventional ML in the Broader Literature

Recent reviews and research papers continue to report color, texture, shape, SVM, random-forest, KNN, and hybrid feature-classifier systems. Their common methodological advantage is that training can be less data-intensive than end-to-end DL. However, the literature also indicates that handcrafted-feature methods often have difficulty scaling to heterogeneous images and complex backgrounds. Thus, the important comparison is not simply algorithm name versus algorithm name; it is whether the representation remains stable when the acquisition conditions change.

V. DEEP LEARNING APPROACHES

A. CNN-Based Classification

CNNs have become the dominant approach for tomato leaf image classification because convolutional feature learning reduces the need to specify color and texture descriptors manually. Karthik et al.

demonstrated residual learning and attention for three PlantVillage tomato diseases and obtained 98% accuracy with five-fold cross-validation [3]. Trivedi et al. reported 98.49% accuracy for a ten-class tomato disease problem [4]. The important evidence is not only the accuracy values but also the fact that both studies operate on curated leaf-image datasets, which limits direct inference about field robustness.

Guha Paul et al. developed a custom CNN and compared augmentation settings on a dataset containing ten tomato disease classes plus healthy leaves. The study also implemented the selected model in an Android application [8]. This provides stronger evidence of application integration than a statement that a model is “suitable for mobile use,” although an application demonstration should still be distinguished from an independently measured deployment benchmark.

B. Transfer Learning and Residual Networks

Transfer learning reduces the amount of task-specific representation learning required by starting from a model trained on a large source dataset. ResNet-based architectures are particularly common because residual connections facilitate optimization of deeper networks. The literature shows strong classification performance from transfer-learning systems, but the benefits depend on the target dataset and the extent of fine-tuning. A high score on a curated dataset does not by itself establish cross-domain generalization.

The 2025 ensemble study by Sharma et al. combined fine-tuned ResNet50 and MobileNetV2 on a balanced 11,000-image, ten-class Kaggle dataset. The reported test accuracy was 99.91%, while the individual ResNet50 and MobileNetV2 models achieved approximately 90% and 91.82%, respectively [9]. The study used an 80/10/10 train-validation-test split. This illustrates how a model can achieve very high performance under a balanced, fixed dataset protocol, while the same paper acknowledges the need for field images from varied geographic conditions for stronger robustness evidence.

C. Lightweight CNNs and Attention

Efficiency-oriented research is increasingly important because agricultural deployment may involve smartphones, embedded processors, or edge devices. Ahmed et al. reported 99.30% accuracy with 9.60 MB model size and 4.87 million FLOPs using a MobileNetV2-based pipeline [5]. Ullah et al. reported 99.34% accuracy with DTomatoDNet on 10,000 images from a ten-class dataset [6]. The latter study used an 80/20 training-test division and was executed on a laptop, providing a computationally lighter architecture but not a full independent edge-device latency benchmark.

Zhang et al. proposed LDAMNet, combining dual attention with a lightweight architecture. On their ten-category tomato dataset, the model achieved 98.71% average accuracy with 0.91 million parameters and 0.142 GFLOPs [7]. The study also performed a cross-dataset experiment on a four-class rice dataset, showing that the model could retain high recognition performance in a different crop domain. Nevertheless, cross-crop transfer should not be interpreted as equivalent to field generalization within tomato because the image distribution and disease semantics are different.

Liu et al. used multi-task distillation to jointly exploit tomato disease classification and severity estimation. Their data were assembled from AI Challenger, PlantDoc, and Taiwan tomato disease sources, and the study excluded a tomato canker category because only three samples were available. The authors reported that an MTDL-optimized EfficientNet improved classification accuracy over a single-task ResNet101 while using a much smaller parameter fraction [10]. This study is important because it connects efficiency to a richer diagnostic task rather than accuracy alone.

D. Explainability and Recent Architectures

Recent models increasingly include attention and explainability. Assaduzzaman et al. introduced XSE-TomatoNet, an EfficientNetB0-based model with squeeze-and-excitation blocks and multi-scale feature fusion, and reported 99.11% accuracy in the article abstract [11]. The same paper reports an average validation accuracy of 98.35% under five-fold cross-validation and 98.88% under ten-fold cross-validation. Its internal comparison table also reports a 99.41% accuracy figure for the proposed model

under a separate comparison setting; these values should therefore not be conflated. The work also used LIME, SHAP, Grad-CAM, and Grad-CAM++ for interpretability. Such methods can help examine whether a classifier focuses on symptomatic regions, but explanation quality should be assessed separately from classification accuracy.

Ghosh et al. evaluated several modern architectures, including VGG19, ViT, EfficientNetV2, ConvNeXt, and a hybrid model, on a ten-class tomato dataset. Their results show that architecture complexity does not automatically translate into a proportional gain in accuracy, while the study also demonstrates that lightweight alternatives can provide competitive results under the same experimental setting [12]. This supports the broader conclusion that model selection should consider both performance and computational burden.

VI. DATASETS, VALIDATION CONDITIONS, AND COMPARATIVE EVIDENCE

PlantVillage remains a major source of tomato leaf images in the reviewed literature. Its value is reproducibility and standardized class definitions. However, images captured under controlled or relatively clean conditions can reduce the difficulty of the recognition problem. The distinction is visible in the literature: Fuentes et al. collected images in agricultural environments with complex backgrounds and localized disease/pest regions [2], whereas several classification studies rely on curated leaf datasets [3]–[7], [9], [11], [12]. Jelali’s review of field-based tomato studies similarly emphasizes that relatively few studies use real agricultural field datasets [13].

This difference is critical when interpreting accuracy. A 99% result obtained on a balanced dataset with similar acquisition conditions cannot be assumed to represent 99% performance on independent field images. Cross-dataset testing, geographic variation, unseen cultivars, varying disease severity, and uncontrolled illumination provide more demanding evidence of generalization. The reviewed literature contains examples of cross-dataset testing, but these remain less common than within-dataset evaluation [7], [13].

TABLE I. VERIFIED COMPARATIVE SUMMARY OF REPRESENTATIVE STUDIES

Study	Task	Dataset / images	Classes	Method	Reported result	Efficiency / validation
Mokhtar et al. [1]	Classification	Tomato leaf images; 100 per disease	2	Gabor wavelet + SVM	99.5% accuracy	Small dataset; kernel comparison
Fuentes et al. [2]	Object detection	~5,000 field images	9 disease/pest categories + background	Faster R-CNN, R-FCN, SSD + VGG/ResNet	Recognition/localization reported; no pooled accuracy used here	80/10/10 split; field images
Karthik et al. [3]	Classification	PlantVillage	3 diseases	Residual CNN + attention	98% accuracy	5-fold cross-validation
Trivedi et al. [4]	Classification	Tomato leaf dataset	10	CNN with preprocessing/segmentation	98.49% accuracy	Within-dataset evaluation
Ahmed et al. [5]	Classification	PlantVillage tomato	10	MobileNetV2 + classifier	99.30% accuracy	9.60 MB; 4.87M FLOPs

		subset; 18,160 images				
Ullah et al. [6]	Classifica tion	Kaggle tomato leaf dataset; 10,000 images	10	DTomatoDNet	99.34% accuracy	80/20 train-test; 19 learnable layers
Zhang et al. [7]	Classifica tion	Kaggle merged tomato dataset	10	LDAMNet with dual attention	98.71% average accuracy	0.91M paramete rs; 0.142 GFLOPs
Paul et al. [8]	Classifica tion	Tomato dataset	11 (10 diseases + healthy)	Custom CNN / VGG transfer learning	95% accuracy and recall reported	Android/ web applicati on
Sharma et al. [9]	Classifica tion	Kaggle; 11,000 images	10	ResNet50 + MobileNetV2 ensemble	99.91% test accuracy	80/10/10 split; balanced classes
Liu et al. [10]	Classifica tion + severity	AI Challeng er + PlantDo c + Taiwan sources	Multi- source; class count varies	Multi-task distillation	Improved classification with reduced parameters	Cross- source dataset; 9.46% paramete r use reported for one comparis on
Assaduzza man et al. [11]	Classifica tion	Tomato disease dataset	10	XSE-TomatoNet / EfficientNetB0	99.11% in abstract; 10-fold validation avg. 98.88%	10-fold validatio n avg. 98.88%; XAI used
Ghosh et al. [12]	Classifica tion	Tomato dataset	10	VGG19, ViT, EfficientNetV2, ConvNeXt, hybrid	Up to 98% for hybrid in reported comparison	Architect ure comparis on

The table is intentionally descriptive rather than a ranking. The reported results come from different datasets, class definitions, image conditions, train/test partitions, and evaluation protocols. Consequently, the percentages should not be interpreted as a common leaderboard. For example, the 99.30% reported by Ahmed et al. was obtained on a ten-class PlantVillage tomato subset, whereas Fuentes et al. addressed object detection in complex field images [2], [5]. Similarly, the 99.91% result of Sharma et al. was obtained on a balanced 11,000-image dataset with a fixed 80/10/10 split [9]. These are different experimental questions.

VII. CRITICAL ANALYSIS OF THE REVIEWED EVIDENCE

A. Dataset and Class Design

The reviewed studies demonstrate substantial variation in dataset size and class design. Small classical-ML studies may contain only hundreds of images, whereas recent CNN studies use thousands of images. Increasing sample count is useful, but dataset size alone does not guarantee diversity. A balanced dataset can simplify optimization, while a highly imbalanced dataset may require augmentation or reweighting. Ahmed et al. explicitly addressed class imbalance using runtime

augmentation [5], whereas Sharma et al. used a balanced dataset and therefore did not require oversampling [9].

B. Preprocessing and Illumination

Preprocessing is not a neutral operation: it changes the input distribution seen by the classifier. CLAHE-based illumination correction in Ahmed et al. [5] and image enhancement in Zhang et al. [7] show that image quality can be part of the model design. However, the literature provides limited standardized evaluation of the same classifier under multiple illumination and camera conditions. More controlled robustness experiments are therefore needed.

C. Generalization

The strongest recurring limitation is the difference between controlled datasets and real agricultural images. PlantVillage-based experiments offer reproducibility, but field environments introduce clutter, overlapping leaves, shadows, variable camera distances, partial leaves, and different symptom stages. Fuentes et al. directly addressed complex field scenes through object detection [2]. Jelali's 2024 review further found that relatively few tomato studies used real field datasets [13]. Thus, claims of generalization should be tied to evidence such as external test sets, cross-location evaluation, or field-image experiments.

D. Computational Requirements

Computational efficiency is increasingly measurable. Ahmed et al. reported model size and FLOPs [5], Zhang et al. reported parameters and FLOPs [7], and Liu et al. compared parameter usage across architectures [10]. By contrast, simply describing a model as "lightweight" or "mobile-friendly" does not establish actual deployment performance. A rigorous deployment claim should ideally report hardware, input resolution, preprocessing time, inference latency, memory use, model size, and power or energy consumption.

E. Evaluation Consistency

Accuracy is the most frequently reported metric, but it is insufficient for imbalanced or deployment-oriented problems. Precision, recall, F1-score, confusion matrices, parameter count, FLOPs, and inference latency provide complementary information. The reviewed literature uses different combinations of these measures. Consequently, future reviews and experiments should report a common minimum set of metrics and clearly distinguish validation accuracy, test accuracy, cross-validation performance, and external-dataset performance.

VIII. RESEARCH GAPS

The reviewed evidence supports five connected research gaps.

1) Dataset diversity and field representation. Most high-accuracy classification studies use curated or public datasets, while field-image evidence is less common. Fuentes et al. provide a notable field-oriented example, and Jelali's review documents the scarcity of real-field tomato studies [2], [13]. The gap is therefore not simply "more images are needed"; the more specific need is geographically and environmentally diverse images with reliable expert labels.

2) Illumination and acquisition robustness. Ahmed et al. explicitly incorporated illumination correction, and Zhang et al. used enhancement and oversampling [5], [7]. These studies support the importance of acquisition variability, but a common benchmark for changes in lighting, camera, distance, and background is still lacking.

3) Generalization beyond the source dataset. High within-dataset accuracies are common, including 99.30% [5], 99.34% [6], and 99.91% [9]. These values demonstrate performance under their respective protocols, not universal field accuracy. Independent test sets and cross-location validation are required to determine whether the learned representations transfer.

4) Efficiency versus accuracy. Lightweight methods such as Ahmed et al. [5], Ullah et al. [6], Zhang et al. [7], and Liu et al. [10] show that efficiency can be incorporated into model design. However, only a

subset of studies report deployment-relevant measurements. Future work should evaluate parameter count, FLOPs, RAM, latency, model size, and energy on the target hardware.

5) Task-specific evaluation. Classification, object detection, and segmentation solve different problems. Fuentes et al. demonstrated field object detection [2], whereas most reviewed studies classify leaf images [3]–[12]. Future research should avoid using “detection” as a generic label and should evaluate the metric appropriate to the actual task, such as accuracy/F1 for classification, mAP for object detection, and IoU/Dice for segmentation.

IX. DISCUSSION

Collectively, the reviewed studies show a clear methodological transition from handcrafted feature engineering to learned representations and, more recently, to efficiency-aware and explainable architectures. Classical SVM systems demonstrate that carefully designed features can be effective on constrained datasets [1]. CNNs reduce dependence on manual descriptors and support multi-class classification [3], [4]. Transfer learning and lightweight architectures then address the practical tension between representational capacity and computational cost [5], [6]. Attention, distillation, and multi-scale feature fusion further attempt to improve feature selectivity or reduce the cost of deployment [7], [10], [11].

The evidence also shows why accuracy alone is an incomplete basis for judging progress. A model may achieve a very high test accuracy on a balanced dataset but still face difficulties when leaves are photographed in the field. Conversely, a model evaluated on complex field images may have a lower numerical accuracy while addressing a substantially harder recognition problem. Therefore, the most informative future studies will report performance under multiple conditions rather than maximizing a single within-dataset accuracy.

The literature further suggests that practical deployment should be treated as an experimental question. Ahmed et al. quantified model size and FLOPs [5], Zhang et al. quantified parameters and FLOPs [7], and Paul et al. demonstrated integration into web/Android applications [8]. These forms of evidence are useful, but application integration and theoretical efficiency should not be conflated with measured end-to-end field latency. Reproducible hardware benchmarks remain an important need.

Explainability is another emerging direction. XSE-TomatoNet incorporated LIME, SHAP, Grad-CAM, and Grad-CAM++ [11]. Such tools can help researchers inspect whether the model uses plausible disease regions, but an explanation method does not itself establish clinical or agronomic correctness. Future evaluation should therefore combine predictive performance, localization/attention consistency, robustness, and expert assessment.

X. CONCLUSION

This narrative review examined the development of machine-learning and deep-learning approaches for tomato leaf disease image analysis. Handcrafted feature methods, represented by SVM-based approaches, can achieve strong results on small and constrained datasets but depend on feature design and may be sensitive to changes in image conditions [1]. CNN-based methods learn visual representations directly from images and have become the dominant approach for multi-class tomato leaf classification [3], [4]. Recent studies increasingly address practical limitations through transfer learning, lightweight networks, attention mechanisms, knowledge distillation, and explainability [5]–[12].

The principal conclusion from the reviewed evidence is that reported accuracy must always be interpreted together with dataset characteristics and evaluation conditions. Very high accuracies are frequently obtained on curated tomato datasets, while field-oriented studies remain less common. The literature therefore supports continued work on diverse field datasets, independent validation, illumination and background robustness, class imbalance, and measurable deployment efficiency. Future studies should report standardized evaluation metrics together with model size, computational cost, inference latency, and target-hardware conditions. Clear separation of classification, detection, and segmentation tasks will also improve the scientific validity of comparisons.

The review thus identifies a research direction centered on robust and computationally efficient tomato disease image analysis, with evaluation designs that bridge the gap between controlled benchmark performance and real agricultural deployment.

REFERENCES

- [1] U. Mokhtar, M. A. S. Ali, A. E. Hassanien, and H. A. Hefny, "Tomato leaves diseases detection approach based on Support Vector Machines," in *Proc. 11th Int. Comput. Eng. Conf. (ICENCO)*, Cairo, Egypt, 2015, pp. 246–250, doi: 10.1109/ICENCO.2015.7416356.
- [2] A. Fuentes, S. Yoon, S. C. Kim, and D. S. Park, "A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition," *Sensors*, vol. 17, no. 9, Art. no. 2022, 2017, doi: 10.3390/s17092022.
- [3] R. Karthik, M. Hariharan, S. Anand, P. Mathikshara, A. Johnson, and R. Menaka, "Attention embedded residual CNN for disease detection in tomato leaves," *Applied Soft Computing*, vol. 86, Art. no. 105933, 2020, doi: 10.1016/j.asoc.2019.105933.
- [4] N. K. Trivedi, V. Gautam, A. Anand, H. M. Aljahdali, S. G. Villar, D. Anand, N. Goyal, and S. Kadry, "Early detection and classification of tomato leaf disease using high-performance deep neural network," *Sensors*, vol. 21, no. 23, Art. no. 7987, 2021, doi: 10.3390/s21237987.
- [5] S. Ahmed, M. B. Hasan, T. Ahmed, M. R. K. Sony, and M. H. Kabir, "Less is more: Lighter and faster deep neural architecture for tomato leaf disease classification," *IEEE Access*, vol. 10, pp. 68868–68884, 2022, doi: 10.1109/ACCESS.2022.3187203.
- [6] N. Ullah, J. A. Khan, S. Almakdi, M. S. Alshehri, M. Al Qathrady, E. A. Aldakheel, and D. S. Khafaga, "A lightweight deep learning-based model for tomato leaf disease classification," *Computers, Materials & Continua*, vol. 77, no. 3, pp. 3969–3992, 2023, doi: 10.32604/cmc.2023.041819.
- [7] E. Zhang, N. Zhang, F. Li, and C. Lv, "A lightweight dual-attention network for tomato leaf disease identification," *Frontiers in Plant Science*, vol. 15, Art. no. 1420584, 2024, doi: 10.3389/fpls.2024.1420584.
- [8] S. Guha Paul, A. Biswas, A. Saha, M. S. Zulfiker, N. A. Ritu, I. Zahan, M. Rahman, and M. A. Islam, "A real-time application-based convolutional neural network approach for tomato leaf disease classification," *Array*, vol. 19, Art. no. 100313, 2023, doi: 10.1016/j.array.2023.100313.
- [9] J. Sharma, A. A. Al-Huqail, A. Almogren, H. Doshi, B. Jayaprakash, B. Bharathi, A. U. Rehman, and S. Hussien, "Deep learning based ensemble model for accurate tomato leaf disease classification by leveraging ResNet50 and MobileNetV2 architectures," *Scientific Reports*, vol. 15, Art. no. 13904, 2025, doi: 10.1038/s41598-025-98015-x.
- [10] B. Liu, S. Wei, F. Zhang, N. Guo, H. Fan, and W. Yao, "Tomato leaf disease recognition based on multi-task distillation learning," *Frontiers in Plant Science*, vol. 14, Art. no. 1330527, 2024, doi: 10.3389/fpls.2023.1330527.
- [11] M. Assaduzzaman, P. Bishshash, M. A. S. Nirob, A. A. Marouf, J. G. Rokne, and R. Alhajj, "XSE-TomatoNet: An explainable AI based tomato leaf disease classification method using EfficientNetB0 with squeeze-and-excitation blocks and multi-scale feature fusion," *MethodsX*, vol. 14, Art. no. 103159, 2025, doi: 10.1016/j.mex.2025.103159.
- [12] H. Ghosh, I. S. Rahat, M. M. R. Emon, et al., "Advanced neural network architectures for tomato leaf disease diagnosis in precision agriculture," *Discover Sustainability*, vol. 6, Art. no. 312, 2025, doi: 10.1007/s43621-025-01149-1.
- [13] M. J. Jelali, "Deep learning networks-based tomato disease and pest detection: A first review of research studies using real field datasets," *Frontiers in Plant Science*, vol. 15, Art. no. 1493322, 2024, doi: 10.3389/fpls.2024.1493322.