A Comprehensive Survey on Deep Learning-based Techniques for Tomato Leaves Disease Detection and Classification

Jagisha

DAV Institute of Engineering & Technology, Jalandhar, 144408, Punjab, India
Contributing author: jagishaluther77@gmail.com
DOI: 10.64823/ijter.2506006

Abstract: Tomato is one of the most cultivated and consumed vegetable crops globally, but its yield is significantly threatened by a variety of diseases, primarily manifesting on the leaves. Early and accurate detection of these diseases is crucial for effective pest management and preventing substantial economic losses. Traditional methods, which rely on manual inspection by experts, are often slow, labor-intensive, and prone to human error. This survey paper provides a systematic and comprehensive review of the rapidly evolving field of automated tomato leaf disease detection, with a primary focus on deep learning (DL) techniques. We catalog a wide range of methodologies, from classical image processing and machine learning to state-of-the-art convolutional neural networks (CNNs) and vision transformers. The paper details publicly available datasets, discusses key technical challenges such as limited data, complex backgrounds, and real-time deployment, and analyzes the performance metrics of various approaches. Finally, we outline promising future research directions, including the integration of multimodal data, explainable AI (XAI), and the development of lightweight models for mobile and edge computing. This survey serves as a valuable resource for researchers and agricultural technologists aiming to understand the current landscape and contribute to advancing this critical application domain.

Keywords: Tomato Leaf Disease, Plant Pathology, Deep Learning, Convolutional Neural Networks (CNN), Image Classification, Object Detection, Vision Transformers, Precision Agriculture.

I. Introduction

The tomato (*Solanum lycopersicum*) is a cornerstone of global agriculture and nutrition, serving as a vital source of vitamins, minerals, and antioxidants. However, tomato plants are susceptible to over 200 diseases, with a significant number causing visible symptoms on the leaves [1]. Common and devastating diseases include Early Blight (*Alternaria solani*), Late Blight (*Phytophthora infestans*), Septoria Leaf Spot (*Septoria lycopersici*), and various bacterial and viral infections like Bacterial Spot and Tomato Yellow Leaf Curl Virus. These diseases can lead to severe defoliation, reduced photosynthetic capacity, and ultimately, catastrophic yield loss if not identified and managed promptly.

Historically, disease identification has been a manual process dependent on the expertise of farmers and agronomists. This method is not scalable, suffers from subjectivity, and is inefficient for monitoring large fields. The advent of digital imaging and artificial intelligence has paved the way for automated, rapid, and accurate diagnostic systems. The paradigm has shifted from manual inspection to computer-assisted diagnosis, and more recently, to fully autonomous systems powered by deep learning.

This survey aims to synthesize the extensive research conducted over the past decade on tomato leaf disease detection. We focus particularly on the era dominated by deep learning, which has demonstrated remarkable success in image-based classification and detection tasks. The contributions of this paper are fourfold:

1. To provide a structured taxonomy of the technical approaches, from traditional image processing to advanced DL models.
2. To review and compare the performance of prominent methodologies on standard benchmarks.
3. To identify and discuss the critical challenges and limitations currently faced by the research community.
4. To propose insightful future directions that can drive the field toward practical, real-world deployment.

II. Methodology of the Survey

To ensure a comprehensive and unbiased review, a systematic literature search was conducted. Primary sources included academic databases such as IEEE Xplore, ScienceDirect, ACM Digital Library, and SpringerLink. Keywords used in the search were: "tomato leaf disease detection," "deep learning," "CNN," "plant disease classification," "precision agriculture," and "image processing." The scope was limited to peer-reviewed journal articles and conference proceedings published predominantly between 2015 and 2024. Over 80 relevant studies were shortlisted, from which the most influential and representative works were selected for in-depth analysis in this survey.

III. Technical Approaches to Tomato Disease Detection

The evolution of automated disease detection systems can be broadly categorized into three stages: (1) Image Processing and Classical Machine Learning, (2) Deep Learning-based Classification, and (3) Object Detection and Segmentation.

3.1. Image Processing and Classical Machine Learning

Before the dominance of DL, approaches primarily involved a pipeline of image preprocessing, handcrafted feature extraction, and classification using classical ML algorithms.

- **Preprocessing:** This step aims to enhance image quality and standardize the input. Techniques include color space conversion (RGB to HSV, L*a*b*), noise removal using Gaussian or median filters, and image segmentation to separate the leaf from the background using methods like Otsu's thresholding or K-means clustering [2].
- **Feature Extraction:** Experts manually designed features to capture disease characteristics.
  - **Color Features:** Statistical measures (mean, standard deviation) from color histograms in different color spaces to capture discoloration.
  - **Texture Features:** Methods like Gray-Level Co-occurrence Matrix (GLCM) were used to quantify patterns of spots, blights, and mildew [3].
  - **Shape and Morphological Features:** Parameters like lesion area, perimeter, and eccentricity were extracted from segmented diseased regions.

- **Classification:** The extracted feature vectors were fed into classifiers such as Support Vector Machines (SVM), k-Nearest Neighbors (k-NN), and Random Forests. While these methods achieved moderate success, their performance was heavily reliant on the quality and discriminative power of the handcrafted features, which often failed to generalize to varied field conditions.

3.2. Deep Learning-based Classification

The breakthrough in this field came with the application of Convolutional Neural Networks (CNNs), which automate the feature extraction and classification process, learning hierarchical representations directly from raw pixel data.

- **Custom CNN Models:** Early DL works designed and trained CNNs from scratch. These architectures typically consisted of several convolutional, pooling, and fully connected layers. While effective, they required large datasets and significant computational resources to train robust models [4].
- **Transfer Learning with Pre-trained Models:** This has become the de facto standard due to its efficiency and high performance. Models pre-trained on massive datasets like ImageNet (e.g., VGGNet, ResNet, Inception, DenseNet, MobileNet) are fine-tuned on tomato disease datasets [5]. This approach leverages learned generic features and adapts them to the specific task, achieving state-of-the-art results even with relatively smaller datasets. Table 1 summarizes the performance of some popular architectures.

Table 1: Performance Comparison of Selected Pre-trained Models on Tomato Disease Classification (Example from Literature)

Model	Accuracy (%)	Key Strengths	Key Weaknesses
VGG16	94.5 - 97.0	Simple, proven architecture	Computationally heavy, many parameters
ResNet50	96.8 - 98.5	Solves vanishing gradient, deep	Slower inference than lighter models
InceptionV3	97.2 - 98.8	Efficient, uses multi-scale filters	Complex architecture
DenseNet121	97.5 - 99.1	Feature reuse, parameter efficient	High memory consumption for dense connections
MobileNetV2	95.0 - 97.5	Fast, lightweight, for mobile devices	Slight trade-off in accuracy for speed

- **Attention Mechanisms and Vision Transformers (ViTs):** Recently, attention mechanisms, and specifically Vision Transformers, have shown great promise. ViTs treat images as sequences of patches and use self-attention to model global dependencies, often outperforming CNNs on large-scale datasets [6]. Their application in plant disease detection is a nascent but rapidly growing area.

3.3. Object Detection and Segmentation

While classification models identify the disease present in an image, they often fail when multiple diseases co-exist on a single leaf or when the goal is to localize the specific infected regions. For this, object detection and instance segmentation models are employed.

- **Object Detection Models:** Frameworks like Faster R-CNN, YOLO (You Only Look Once), and SSD (Single Shot MultiBox Detector) are used to draw bounding boxes around diseased spots or leaves within a complex image of a plant [7]. YOLO variants, in particular, are favored for real-time applications due to their high inference speed.
- **Segmentation Models:** For pixel-level precision, models like U-Net, Mask R-CNN, and DeepLabv3+ are used to segment the exact shape and boundaries of the lesions [8]. This provides detailed information about disease severity, which is crucial for deciding the intensity of treatment.

IV. PUBLICLY AVAILABLE DATASETS

The development of robust models is contingent upon the availability of high-quality, annotated datasets. Several key datasets have been instrumental in propelling research forward:

- **PlantVillage Dataset:** The most widely used benchmark dataset. It contains over 50,000 curated, lab-quality images of healthy and diseased leaves across 14 crop species, with 10 classes dedicated to tomato [9]. Its clean background, however, is a limitation for testing robustness in field conditions.
- **PlantDoc:** A dataset designed to address the lab-field gap. It contains 2,598 images from the internet with complex backgrounds, making it a more challenging and realistic benchmark for object detection [10].
- **Taiwan Tomato Disease Dataset:** A dataset focusing on six common tomato diseases, providing images with varying resolutions and conditions.
- **AI Challenger 2018 Dataset:** A large-scale dataset with fine-grained annotations for crop diseases, including tomato.

The creation of larger, more diverse datasets capturing various growth stages, lighting conditions, and complex backgrounds remains an active need.

V. CHALLENGES AND LIMITATIONS

Despite significant progress, several formidable challenges impede the widespread adoption of these systems:

- **The Lab-Field Gap:** Models trained on clean, lab-style images (like PlantVillage) often experience a significant performance drop when deployed in real fields due to complex backgrounds, occlusions, varying lighting, and different camera angles [11].
- **Data Scarcity and Imbalance:** Collecting and annotating thousands of field images is expensive and time-consuming. Furthermore, datasets are often imbalanced, with some disease classes being underrepresented, leading to model bias.
- **Inter-class Similarity and Intra-class Variation:** Different diseases can exhibit visually similar symptoms (e.g., Early Blight vs. Septoria Leaf Spot), while the same disease can look different at various stages of progression or under different environmental conditions.

- **Real-time Processing and Model Efficiency:** For deployment on mobile devices or drones, models must be lightweight and fast without compromising accuracy. Balancing this trade-off is a key research focus.
- **Multiple Infections and Early Detection:** Distinguishing between co-occurring diseases on a single leaf and detecting infections at their very early, subtle stages are extremely challenging tasks.
- **Explainability and Trust:** DL models are often seen as "black boxes." For farmers to trust an AI's diagnosis, the system must be able to explain *why* it reached a certain conclusion, highlighting the visual evidence it used.

VI. Future Research Directions

To overcome the existing challenges and move toward robust, field-deployable systems, future research should focus on the following directions:

- **Data-Centric Approaches:** Leveraging data augmentation (e.g., Generative Adversarial Networks - GANs), transfer learning, and self-supervised learning to mitigate data scarcity and the lab-field gap.
- **Multimodal Fusion:** Integrating multiple data sources, such as hyperspectral imagery, thermal imaging, and environmental sensor data (temperature, humidity), can provide complementary information that improves diagnostic accuracy beyond visible spectrum images [12].
- **Explainable AI (XAI):** Incorporating techniques like Grad-CAM, LIME, or SHAP to generate visual explanations and build trust with end-users by showing the regions of the leaf that were most influential in the model's decision.
- **Lightweight and Efficient Architectures:** Designing and optimizing models based on MobileNet, EfficientNet, and NAS (Neural Architecture Search) for seamless deployment on edge devices and smartphones.
- **Web-based and Mobile Decision Support Systems (DSS):** Developing integrated platforms where farmers can upload images via a smartphone app and receive instant diagnosis, severity analysis, and management recommendations.

VII. Conclusion

The automation of tomato leaf disease detection through deep learning has transitioned from a theoretical concept to a highly active and promising research domain. This survey has chronicled this journey, highlighting the superior capabilities of DL models, especially CNNs and Vision Transformers, over traditional methods. We have discussed the importance of public datasets, the critical challenges of real-world deployment, and the architectural evolution from simple classification to sophisticated detection and segmentation. The path forward lies in creating more robust, efficient, and transparent systems that can function reliably in the unstructured environment of a farm. By focusing on data-centric learning, multimodal fusion, and explainability, the next generation of AI-powered tools can become indispensable allies to farmers, safeguarding global tomato production and enhancing food security.

References

[1] Agrios, G. N. (2005). *Plant Pathology* (5th ed.). Elsevier Academic Press.

[2] Khirade, S. D., & Patil, A. B. (2015). Plant Disease Detection Using Image Processing. *International Conference on Computing Communication Control and Automation*.

[3] Ariwazhagan, S., et al. (2013). Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features. *Agricultural Engineering International: CIGR Journal*, 15(1), 211-217.

[4] Sladojevic, S., et al. (2016). Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification. *Computational Intelligence and Neuroscience*.

[5] Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using Deep Learning for Image-Based Plant Disease Detection. *Frontiers in Plant Science*.

[6] Dosovitskiy, A., et al. (2020). An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. *International Conference on Learning Representations (ICLR)*.

[7] Fuentes, A., et al. (2017). A Robust Deep-Learning-Based Detector for Real-Time Tomato Plant Diseases and Pests Recognition. *Sensors*.

[8] Too, E. C., et al. (2019). A comparative study of fine-tuning deep learning models for plant disease identification. *Computers and Electronics in Agriculture*.

[9] Hughes, D. P., & Salathé, M. (2015). An open access repository of images on plant health to enable the development of mobile disease diagnostics. *arXiv preprint arXiv:1511.08060*.

[10] Singh, D., et al. (2020). PlantDoc: A Dataset for Visual Plant Disease Detection. *Proceedings of the 7th ACM IKDD CoDS and 25th COMAD*.

[11] Thapa, R., et al. (2020). The Plant Disease Diagnosis Pipeline: A Review of Challenges and Opportunities. *Plant Phenome Journal*.

[12] Zhang, J., et al. (2022). Fusion of deep learning and hyperspectral imaging for early detection of tomato diseases. *Computers and Electronics in Agriculture*.