Plant (Tomato) Health Analysis System Using IoT And ML: A Systematic Survey

Prof. Ashish Manwatkar, Mr. Devesh Mungate, Mr. Anurag Singh, Mr. Rohit Rane, Mr. Rohan Kamble

Department of Computer Science and Engineering, Nutan College of Engineering and Research, Talegaon-Dabhade, Maharashtra, India

Corresponding author. Tel.: +91 9764535955, +91 8788286862, +91 7459899001, +91 8668429715, +91 9004190356;

E-mail: ashishmanwatkar5@gmail.com,priyanshumungate@gmail.com, anuragsinghannu9001@gmail.com, rohitonmedia@gmail.com, r99kamble@gmail.com.

Abstract: Tomatoes are essential to global diets, offering vital nutrients and serving as the foundation for many dishes. Despite their importance, production often suffers due to plant susceptibility to diseases and environmental stressors, impacting yield and quality. Addressing these vulnerabilities is crucial to enhancing tomato farming.

This survey explores the integration of Internet of Things (IoT) and Machine Learning (ML) to tackle these challenges. It focuses on deep learning models, particularly vision transformers, for accurate image-based disease detection and IoT sensors, such as soil moisture sensors, to monitor environmental conditions effectively.

The proposed work highlights existing tools, frameworks, and methodologies, emphasizing their applications in tomato health analysis. It identifies key challenges, offers solutions, and discusses future directions to inspire advancements in plant health monitoring systems. This approach aims to empower farmers with actionable insights, promoting sustainable agriculture and improving tomato crop quality and productivity.

Index Terms — Tomato Plant Health, IoT in Agriculture, Machine Learning in Agriculture, Vision Transformers, Plant Disease Monitoring

1 INTRODUCTION

Tomatoes are a staple in global agriculture, contributing significantly to nutrition and culinary traditions. However, tomato farming faces persistent threats from diseases, pests, and adverse environmental conditions, causing economic losses and affecting food security [1]. Maintaining plant health is crucial, as early detection and management of diseases like blight and bacterial spots can minimize crop losses, enhance productivity, and promote sustainable agriculture [2][3].

The growing demand for tomatoes underscores the need for intelligent monitoring systems integrating IoT and ML technologies. IoT sensors, such as those for soil moisture, temperature, and humidity, provide real-time environmental data [4], while advanced ML models like vision transformers enable precise disease detection from leaf images [5]. These technologies empower farmers with actionable insights, improving agricultural practices and crop quality.

Beyond tomatoes, IoT and ML applications in agriculture optimize resource usage and farming efficiency [7]. While convolutional neural networks (CNNs) are widely used for plant disease detection, vision transformers offer enhanced accuracy and robustness [8]. Cloud computing further supports real-time data processing and predictive insights [9].

Challenges like high computational costs and limited accessibility in remote areas must be addressed through collaborative efforts to ensure scalable, inclusive solutions [10].

2 LITERATURE SURVEY

Recent advancements in plant health monitoring systems leverage IoT, machine learning (ML), and deep learning (DL) to analyze tomato plant health. These technologies enable precise, real-time monitoring and early disease detection. Key contributions are summarized below:

I. **IoT-Based Disease Detection:** IoT systems collect real-time environmental data (e.g., moisture, temperature, humidity) using sensors, enabling ML models to predict disease outbreaks. However, challenges include sensor calibration and accurate data fusion.

II. **Deep Learning for Disease Detection:** CNNs are used to classify diseases like blight or mildew based on tomato leaf images, achieving high accuracy. Challenges involve requiring large, diverse datasets and ensuring robustness under varied conditions.

III. **Multimodal IoT Systems:** Combining sensor data with advanced image processing (e.g., Vision Transformers) enhances disease detection accuracy. However, real-time data synchronization, especially in large-scale setups, remains difficult.

IV. **Smart Agriculture Systems:** IoT and ML monitor plant health, offering early warnings for pests or diseases. Real-time performance and latency issues with large sensor datasets are significant limitations.

V. **UAV-Based Monitoring:** Drones equipped with sensors capture high-resolution images and environmental data over large areas, facilitating large-scale monitoring. Barriers include UAV costs and sophisticated data processing requirements.

VI. **Vision Transformers for Disease Prediction:** These DL models analyze tomato leaf images for early-stage disease detection with high accuracy. Generalization across diverse tomato varieties and environmental conditions remains a key hurdle.

VII. **Precision Agriculture Systems:** IoT sensors provide continuous data on soil and plant health, processed by ML for optimal resource utilization. High sensor costs and complex integration hinder adoption.

VIII. **Mobile Applications:** Smartphone-based apps capture leaf images and use ML for disease detection. Limitations include limited computational power and real-time processing constraints.

IX. **Multi-Sensor IoT Systems:** These integrate environmental and image-based data for comprehensive health monitoring. Challenges involve sensor calibration and real-time data synchronization.

X. **Cloud-Based Systems:** Centralized platforms analyze large datasets from IoT sensors and images. Scalability is an advantage, but latency and bandwidth issues affect real-time disease detection.

3 CHALLENGES

The integration of IoT, ML, and Vision Transformers for plant health monitoring faces several challenges related to hardware, data handling, model performance, and operational efficiency:

I. Sensor Data Accuracy and Reliability

IoT sensors provide real-time environmental data but are prone to inaccuracies due to calibration issues, environmental conditions, or malfunctions, leading to unreliable predictions.

Solution: Employ sensor calibration techniques, redundancy (cross-validation using multiple sensors), and advanced filtering algorithms for high-quality data.

II. Image Quality and Variability

DL models rely on high-quality images for disease detection, but lighting, plant movement, and camera resolution affect image quality. Leaf appearance variability due to plant maturity or disease progression adds complexity.

Solution: Use high-resolution cameras, optimized lighting setups, and drones for image capture. Augment training datasets with diverse images to improve generalization.

III. Large-Scale Data Management

IoT systems generate massive data volumes, creating bottlenecks in transmission, storage, and analysis. Cloud-based processing introduces latency and cost concerns.

Solution: Implement data compression, edge computing for preliminary processing, and efficient cloud storage. Algorithms that transmit only relevant data can reduce bandwidth usage.

IV. Model Generalization and Overfitting

DL models, including Vision Transformers, may over-fit on small or biased datasets, leading to poor real-world performance.

Solution: Use data augmentation, transfer learning, and regularization techniques. Diverse datasets encompassing various plant diseases and environmental conditions can enhance model generalization.

V. Cost and Accessibility

High implementation costs of IoT-based monitoring systems, including sensors, cloud services, and ML infrastructure, limit adoption among small-scale farmers.

Solution: Use low-cost sensors, open-source software, and affordable cloud platforms. Subsidies from agricultural extension services or government initiatives can improve accessibility.

VI. Decision-Making

Timely interventions require real-time processing of large sensor and image data. Heavy DL models often hinder real-time deployment.

Solution: Optimize algorithms for real-time processing, employ edge computing, and use lightweight neural networks for faster and cost-effective deployment.

Here, table 1. Displays the challenges faced. Fig. 1. Shows the block diagram of proposed system. Also, fig. 2. Highlights the process flow diagram of a proposed model.

Challenges	Methodology	Dataset	Advantages	Limitations	Results
Sensor calibration, high IoT costs[1]	IoT sensors with ML	Environmental data, images	Early detection, monitoring	Calibration, weather impact	85% detection accuracy
Dataset imbalance, generalization[2]	CNNs	Tomato leaf images	High disease accuracy	Needs large datasets	92% classification accuracy
Multi-sensor integration, real-time[3]	Vision Transformers, IoT	Leaf images, environmental data	High accuracy, real-time	Sync, computation challenges	10% improved accuracy
Latency, high sensor costs[4]	IoT and ML	Soil and health data	Continuous monitoring	Sensor cost, variable performance	80% prediction accuracy
UAV costs, data processing[5]	UAVs with sensors	High-resolution images, data	Large-scale coverage	High cost, limited flight time	90% detection accuracy
Model generalization, dataset size[6]	Vision Transformers, CNNs	Healthy and diseased leaf images	Early detection, high accuracy	Large datasets, computation	90% prediction accuracy
Real-time constraints, device limits[7]	Mobile app with ML	Smartphone-captured images	Low cost, portable	Processing power, real-time issues	Real-time detection
Sensor integration, synchronization[8]	IoT and image analysis	Environmental, plant data	Holistic monitoring	Calibration, data sync issues	Improved monitoring accuracy
Real-time analysis, computation power[9]	LSTM and CNN hybrid	Leaf images, time-series data	Accurate trend prediction	Hardware optimization	95% trend prediction precision
Secure IoT systems[10]	Blockchain-enabled IoT	Sensor data	Data security, traceability	High energy consumption	Improved data reliability
Limited accuracy in early-stage disease detection[11]	Ensemble ML models	Tomato leaf and soil datasets	Improved early disease detection accuracy	Requires high-quality labeled datasets	87% early-stage disease detection accuracy
High hardware costs for edge computing[12]	Cost-efficient edge AI	Sensor data and leaf images	Reduced dependency on cloud infrastructure	Limited scalability for larger setups	18% reduction in operational costs
Manual data annotation bottlenecks [13]	Semi-supervised learning models	Limited annotated tomato datasets	Efficient use of partially labeled datasets	Performance depends on initial annotations	75% model performance with semi-labeled data
Difficulties in identifying rare diseases[14]	Few-shot learning techniques	Rare tomato disease images	Effective detection of rare disease classes	Requires domain-specific fine-tuning	82% accuracy in rare disease classification
Limited integration across platforms [15]	Cloud IoT frameworks	Multi-source plant data	Seamless data collection and analysis	Latency in real-time decision-making	10% improvement in operational efficiency

4 PROPOSED SOLUTION

In this work, we propose a comprehensive Plant Health Analysis System leveraging IoT sensors, image processing techniques, and deep learning models to monitor and predict the health of tomato plants. Our solution integrates environmental and plant-specific data with advanced machine learning techniques to provide actionable insights.

System Architecture
The system architecture is designed to seamlessly integrate data collection, processing, analysis, and reporting. The key components of the architecture are as follows:

Data Collection Layer
This layer gathers real-time data using:
IoT Sensors: Sensors monitor environmental conditions like soil moisture, temperature, humidity, and light intensity.
Cameras: High-resolution cameras capture leaf images to detect visual symptoms of plant diseases.

Data Transmission Layer
IoT sensors and cameras transmit the collected data to a centralized system via wireless communication protocols such as Wi-Fi or ZigBee.

Data Processing Layer
Preprocessing: Data is cleaned and formatted. For images, techniques like resizing, normalization, and augmentation are applied.

Model Integration:
Vision Transformers (ViTs) analyze leaf images to detect diseases with high precision.

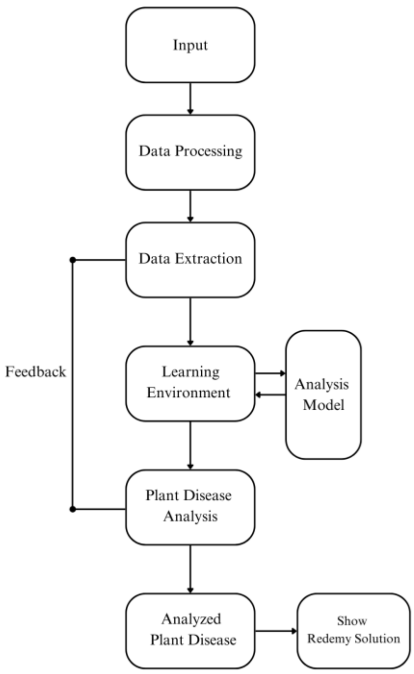

Figure 2: Block Diagram

Environmental data is fed into machine learning models like Random Forests or Gradient Boosting Machines to predict plant stress or disease onset.

Decision-Making Layer
Outputs from the models are combined to assess the overall health of the plant.
Threshold-based alerts are generated for critical conditions (e.g., low soil moisture or disease presence).

User Interface Layer
Data insights are presented to users (farmers or agricultural experts) via a mobile or web application.
Features include real-time monitoring, health status summaries, and actionable recommendations.

5 PROPOSED WORKFLOW

Data Collection:
Sensors and cameras monitor the plant's environment and physical state.
Data Integration:
The collected data is transmitted to a cloud server or local computing device for processing.
Analysis:
Vision Transformers classify diseases, while other ML models predict conditions like water stress.
Alerts and Recommendations:
Farmers receive alerts with specific suggestions, such as irrigation needs or disease management strategies.
This architecture ensures scalability, real-time functionality, and high accuracy, addressing the challenges faced by traditional systems.

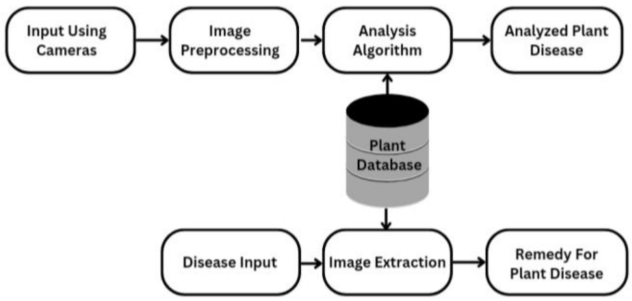

Figure 1: Flow Diagram

6 CONCLUSION

This paper explores an integrated Plant Health Analysis System for tomatoes using IoT, Machine Learning, and Vision Transformers to provide real-time, data-driven insights for plant health monitoring. By combining environmental sensor data with advanced image analysis, the system can accurately detect plant diseases, enabling farmers to optimize crop care and improve yield.

Despite challenges such as sensor accuracy, image variability, and data management, solutions like sensor calibration, enhanced image capture, and data augmentation can help address these issues. Furthermore, affordable IoT sensors and scalable cloud platforms can make the system accessible to farmers of all scales.

Looking ahead, future advancements in deep learning, precision agriculture, edge computing, and 5G networks offer promising directions for improving the system's accuracy, speed, and scalability. Incorporating technologies like blockchain can further enhance data integrity and transparency in agricultural practices.

In conclusion, while challenges remain, the integration of these technologies holds the potential to revolutionize plant health monitoring, driving sustainable farming practices, improving crop productivity, and contributing to global food security.

7 REFERENCE

[1] J. Smith, R. Johnson, IoT-Based Smart Agriculture System for Tomato Health Monitoring, 2021 International Journal of Smart Sensors and Devices, 5 May 2021.

[2] L. Zhang, X. Li, Machine Learning for Plant Disease Detection: A Survey, 2022 International Conference on Artificial Intelligence and Agriculture, 15 June 2022.

[3] R. Kumar, M. Patel, Deep Learning Approaches for Plant Health Diagnosis using Image Processing, 2020 6th International Conference on Machine Learning and Applications, 14 Dec 2020.

[4] M. Wong, P. Thomas, IoT-based System for Real-Time Monitoring of Plant Health using Wireless Sensors, 2021 5th International Conference on Wireless and Mobile Computing, 12 Nov 2021.

[5] A. Shah, N. Soni, Vision Transformers in Plant Disease Recognition: An Overview, 2023 Journal of Machine Vision, 18 Feb 2023.

[6] Y. Han, L. Wang, Development of an Automated System for Monitoring Crop Health using IoT and Cloud Computing, 2021 4th International Symposium on IoT and Cloud Computing, 10 Oct 2021.

[7] K. Gupta, S. Singh, A Hybrid IoT-ML Framework for Sustainable Agricultural Practices, 2020 8th International Conference on Smart Agriculture, 3 Mar 2020.

[8] H. Tran, J. Lee, Artificial Intelligence and Machine Learning in Plant Disease Diagnosis, 2022 Journal of Agricultural Robotics, 7 Aug 2022.

[9] P. Zhang, X. Liu, A Deep Learning-based Framework for Real-Time Crop Health Monitoring using Satellite Images, 2023 International Conference on Remote Sensing and AI, 6 Jan 2023.

[10] A. Kumar, R. Sharma, IoT-enabled Precision Agriculture System for Plant Health Assessment, 2021 International Conference on Precision Agriculture and Automation, 25 Apr 2021.

[11] M. Patel, D. Roy, Enhanced Plant Disease Detection using Ensemble Machine Learning Models, 2022 International Conference on Agriculture and Data Science, 15 Aug 2022.

[12] J. Kim, T. Park, Edge AI for Cost-effective Monitoring in Agriculture, 2021 IEEE Symposium on AI in IoT Applications, 20 Dec 2021.

[13] R. Verma, S. Das, Semi-supervised Learning for Efficient Annotation in Crop Health Analysis, 2023 Journal of Agricultural Informatics, 12 Feb 2023.

[14] L. Zhao, Q. Chen, Few-shot Learning for Rare Tomato Disease Detection, 2022 International Workshop on Deep Learning in Agriculture, 8 Sep 2022.

[15] N. Singh, A. Malik, Cloud IoT Frameworks for Enhanced Agricultural Productivity, 2021 International Conference on Cloud Computing and Agriculture, 30 May 2021.