Edge AI and IoT for Real-Time Crop Disease Detection: A Survey of Trends, Architectures, and Challenges

Abhishek S. Madiwal*., Rajivranjan B Jha., Dr. Pijush Barthakur

Dept. of Master of Computer Applications, K.L.S Gogte Institute of Technology Belagavi, Karnataka, India

DOI: https://doi.org/10.51584/IJRIAS.2025.100700092

Received: 08 July 2025; Accepted: 18 July 2025; Published: 14 August 2025

ABSTRACT

The intersection of Edge Artificial Intelligence (Edge AI) and the Internet of Things (IoT) has transformed precision agriculture, especially in the application of real-time crop disease detection. This survey investigates the most important technologies, architectures, and AI models that facilitate the on-field detection of crop anomalies through edge devices and smart sensors. To continue, click here. We categorize and summarize recent work into four main categories: edge-enabled sensing and processing frameworks, lean deep learning models for resource-limited devices, data-driven methods such as augmentation and incremental learning, and communication protocols appropriate for agricultural settings. The use of CNN architectures such as Mobile Net and Efficient Net, and methods such as pruning and quantization, have made effective disease detection possible on hardware such as Raspberry Pi and Jetson Nano. Nonetheless, significant challenges continue to exist in scalability, energy efficiency, data sparsity, model generalization, and secure communication. We end with directions for future research involving multi-modal sensor fusion, federated learning, and adaptive AI systems for improvements in robustness and scalability in real-world agricultural settings.

Keywords: Edge AI, IoT, Precision Agriculture, Crop Disease Detection, Real-Time Monitoring

INTRODUCTION

Agriculture is a determinant factor in global food security, economic growth, and rural livelihood, particularly in low- and middle-income nations. As the world population continues to rise and is estimated to be 9.7 billion by 2050, there is mounting pressure on agricultural systems to be more productive and minimize losses [1]. Among the most important causes of crop loss is the epidemic of plant diseases that can greatly lower yield and quality. It is estimated that plant diseases account for about 20–40% of crop production loss around the world every year [2]. For this reason, early and precise detection of plant disease is imperative to reducing these losses and maintaining sustainable agricultural production.

Conventional methods of disease detection—like field observation by hand, laboratory analysis, and distant specialist advice—tend to be time-consuming, expensive, and susceptible to human error. These drawbacks limit extensive monitoring, particularly in rural or far-off agricultural areas with restricted access to specialist agronomists or laboratory facilities. In addition, visible signs of plant illness usually become manifest only after significant internal damage has been done, and thus reactive measures are less successful [3].

Current innovations in Artificial Intelligence (AI), the Internet of Things (IoT), and Edge Computing provide new opportunities for creating smart systems that can monitor plant diseases in the field in real time. IoT-based networks can continuously track environmental and plant parameters via sensors (e.g., leaf moisture, temperature, humidity, spectral imaging), whereas Edge AI permits direct deployment of machine learning models on low-power embedded devices such as Raspberry Pi, Jetson Nano, ESP32-CAM, and Coral Edge TPU [4], [5]. The interplay between the two facilitates near-instant decision-making, reduces dependence on internet connectivity, and maintains data privacy through local processing of sensitive agricultural data [6].

In contrast to cloud-based approaches, edge computing enables the system to compute near the source of data—inline in the field—minimizing latency and bandwidth usage. Edge AI algorithms can infer disease symptoms from images of leaves, detect anomalies in sensor data, and even initiate automated irrigation or pesticide systems. Lightweight Convolutional Neural Networks (CNNs) like MobileNet, EfficientNet, and SqueezeNet, with optimization using methods like model pruning, quantization, and depthwise separable convolutions, are able to keep their high accuracy on devices with limited resources [7], [8], [9].

Several architectures have been proposed in recent years to combine AI with IoT for crop disease detection. For example, the CROPCARE system performs on-device inference for disease detection using a camera-equipped Jetson Nano, achieving real-time diagnosis with minimal power usage [10]. Other systems have utilized thermal imaging, hyperspectral cameras, and even drone-mounted sensors integrated with edge computing hardware to monitor large farmlands autonomously [11].

Even with promising developments, a number of technical as well as practical issues still restrict the scalability and field-level application of such systems. First, the ability of deep learning models trained on a single crop or location to generalize to new environmental contexts or novel disease types is poor. Annotated agricultural datasets availability is also one of the foremost limitations, where most datasets are imbalanced, small, and of non-diverse stages of disease progression and lighting conditions [12]. Moreover, real-time on-edge performance also requires attention to appropriate trade-offs between model complexity, speed of inference, and power consumption [13].

Another burning concern is trustworthy and efficient communication in rural fields for agriculture. Communication protocols such as LoRaWAN, ZigBee, and NB-IoT provide low-power and long-range solutions but come with data rate and latency trade-offs. Multimodal sensor data integration (e.g., RGB, thermal, humidity, soil pH) also poses issues in sensor calibration, synchronization, and fusion schemes [14], [15]. Additionally, privacy and security of data are still important concerns, particularly in decentralized architectures where edge devices can be physically vulnerable or don't have secure storage and encryption mechanisms [16].

To fill these voids, this survey gives an extensive review of the literature between 2020 and 2025 on Edge AI and IoT technologies for crop disease detection. The contributions of this survey are as follows:

1. Taxonomic categorization of current Edge AI–IoT frameworks according to system architecture, device category, communication protocol, and AI model employed.

2. In-depth comparison of deep learning models used on edge devices, such as their performance, optimization methods, and hardware support.

3. Discussion of data-related methods like synthetic data generation, augmentation techniques, and lifelong learning to counter data sparsity.

4. Analysis of communication protocols and sensor fusion techniques applied for real-time monitoring and control in smart agriculture.

5. Detection of open research challenges and directions, including federated learning, explainable AI, and cross-crop generalization.

The rest of the paper follows the organization: Section II is a review of relevant literature organized by system type and function. Section III describes the essential methodologies employed in existing implementations. Section IV enumerates the main challenges and open research issues. Section V presents future directions for research. Section VI presents a general discussion, and Section VII concludes the survey.

LITERATURE REVIEW

In the last five years (2020–2025), huge research has been seen at the Edge AI-IoT convergence to meet the increasing demand for real-time and low-resource crop disease identification systems. This section classifies the literature into four primary categories: (1) IoT-Edge system architectures, (2) edge-deployment friendly lightweight deep learning models, (3) data-centric learning methods, and (4) communication protocols and sensor integration.

IoT-Edge Based System Architectures

The integration of edge computing hardware and IoT sensors provides the localized intelligence needed for agricultural decision-making. Such systems consist of wireless communication modules for transmitting data, edge AI hardware for processing, and sensing modules (e.g., cameras, temperature, humidity sensors).

Garg et al. [1] created CROPCARE, a real-time disease detection framework for crops with a Jetson Nano board, RGB camera, and Wi-Fi-based data transmission. The system showed the capability of real-time inferencing on the field itself with minimal dependence on cloud connectivity.

Rumly et al. [2] proposed a low-cost LoRa WAN-based rice disease detection system that uses a Raspberry Pi and multispectral sensors. Their architecture supported early disease detection while optimizing energy and bandwidth usage in remote field conditions.

Proietti et al. [3] introduced a greenhouse-based system based on Raspberry Pi incorporating temperature and humidity sensors with a CNN-based disease detection for tomato plants. The system emphasized reducing latency and less data offloading needs.

Hayajneh et al. [4] proposed a Tiny ML-based UAV framework that implemented edge AI models on drones for airborne crop monitoring. The framework utilized onboard computation to predict disease with transfer-learned models and offloaded results via 4G or edge gateways.

These works highlight the transition from cloud-based designs to decentralized, autonomous edge systems appropriate for harsh and connectivity-limited agricultural settings.

Lightweight Deep Learning Models for Edge Deployment

Scaling deep learning to edge devices means models need to be light, energy-efficient, and speedy, but with minimal loss of accuracy. Several methods like model compression, pruning, quantization, and architectural optimization have been used to make this possible.

Zhang et al. [5] used a depth wise separable CNN to classify crop diseases and ran it on a Raspberry Pi. The model had 92% accuracy with less than 500ms inference time, with excellent suitability for edge applications.

Wicaksono and Apriono [6] proposed a CNN-based plant disease classifier on FPGA-based edge hardware, while maintaining high speed and low power consumption. The work is remarkable for taking advantage of edge inference with hardware acceleration.

Lv et al. [7] compared Efficient Net and MobileNetV3 for the detection of peanut leaf disease and deployment optimization on Jetson Nano and Coral TPU. The research indicated that MobileNetV3 offered the best trade-off in terms of accuracy and latency for in-field diagnosis.

Da Silva and Almeida [8] showed the viability of employing thermal vision for disease diagnosis on edge devices. Tests with Raspberry Pi and Coral TPU exhibited enhanced robustness with changes in illumination.

In an innovative hybrid model, Plant XViT [9] blended Vision Transformer (ViT) and CNN layers for interpretable plant disease diagnosis. Though computationally more expensive, it exhibited enhanced generalization and interpretability.

These advances also represent a trend toward ultra-efficient models that are purposely designed to satisfy real-world energy, compute, and robustness requirements.

Data Augmentation and Adaptive Learning Methods

One of the biggest challenges in using AI for agriculture is having access to large, balanced data sets. Several data-centric methods have been tried to address this challenge.

Cap et al. introduced Leaf GAN, a GAN-based approach to generate synthetic diseased leaf images [10]. Incorporating GAN-augmented data greatly enhanced the stability of CNNs under imbalanced classes.

Gurunathan et al. [11] employed EfficientNet-B0 along with transfer learning on a small cassava leaf dataset. Their model reached 94% accuracy using a mere 2,000 labeled samples, demonstrating the utility of pretrained models in the agricultural context.

Santos et al. [12] designed a continuous learning paradigm that enables edge models to incrementally learn from field observations. The system employed Elastic Weight Consolidation (EWC) to preserve past knowledge, thus facilitating learning in change-prone agricultural environments.

Adaptive methods like semi-supervised learning, few-shot learning, and federated learning are increasingly making waves, but they have not yet been fully implemented on edge hardware owing to resource and communication limitations.

Sensor Integration and Communication Protocols

Effective communication is necessary for real-time data transfer between gateways, edge nodes, and sensors. Research has investigated a number of wireless technologies, which are specially designed for smart agriculture.

Matilla et al. [13] employed LoRa WAN in an IoT-based irrigation system, which provides wide-area coverage with low energy usage. Their findings are especially appropriate for field environments with limited cellular coverage.

Bai et al. [14] proposed a delay-aware offloading algorithm for UAV-Edge-Cloud systems. Their algorithm optimizes decision latency and task scheduling on nodes and thus is well applicable to aerial and ground system scenarios.

Das and Luo [15] designed Light ESD, an edge-based anomaly detector with multimodal input (e.g., humidity, light intensity, soil pH). Their system employed sensor fusion to enhance detection reliability and minimize false alarms.

Da Silva et al. [8] illustrated thermal + RGB fusion in disease diagnosis, which evidenced multimodal sensing enhances accuracy in situations where one of the modalities is poor (e.g., light overexposure or darkness).

The majority of the systems surveyed are based on wireless protocols such as ZigBee, NB-IoT, and Wi-Fi, sometimes supported by solar or battery-powered nodes to increase usage time in the field.

Table 1: Key Studies on Edge AI and IoT for Crop Disease Detection

Reference	Focus	Key Contribution
[1]	Edge inference system	Real-time detection on Jetson Nano
[2]	IoT + LoRaWAN	Early detection in remote farms
[3]	Greenhouse system	CNN-based tomato disease detection
[4]	UAV + TinyML	Aerial monitoring with onboard inference
[5]	Lightweight CNN	92% accuracy, <500ms latency on Raspberry Pi

[7]	Model comparison	MobileNetV3 best for speed and accuracy
[9]	Data augmentation (GAN)	LeafGAN improves performance on rare classes
[11]	Transfer learning	EfficientNet on cassava, 94% accuracy
[12]	Continual learning	EWC for dynamic learning in the field
[13]	Sensor communication	LoRaWAN for wide-area, low-power setup

SUMMARY OF LITERATURE REVIEW

The literature surveyed demonstrates a clear path towards efficient, practical, and smart real-time crop disease detection through Edge AI and IoT. System designs have moved towards decentralization and autonomy, and optimally adopted models with miniaturized hardware. Yet, gaps remain in dataset access, cross-crop model generalizability, multimodal data integration, and secure, scalable deployment procedures. These findings are the basis for our subsequent analysis of methodologies and open issues in the next sections.

METHODOLOGIES

Design and development of real-time crop disease identification systems based on Edge AI and IoT comprise an array of connected building blocks from hardware design and model development to data preprocessing and communication protocols. Within this section, we group and detail methodologies into the following five functional dimensions: system design, model development, data strategies, deployment optimization, and communication/sensor fusion. Collectively, these layers form the building blocks of contemporary precision agriculture with smart edge systems.

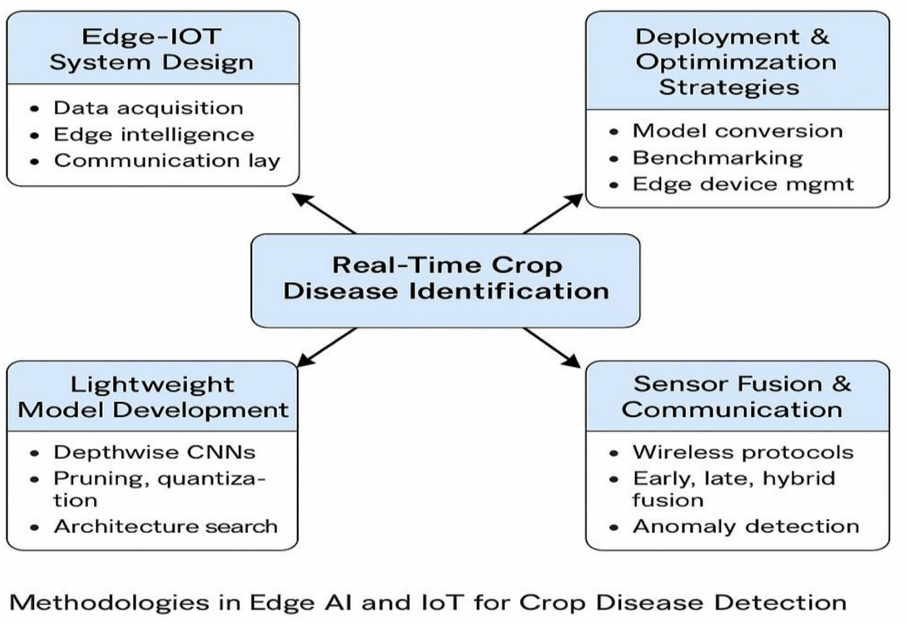

Fig. 1. Functional blocks involved in real-time crop disease identification using Edge AI and IoT, including system design, model development, data handling, optimization, and sensor fusion.

Edge-IoT System Design

The general system design of Edge AI–based crop monitoring systems comprises three layers:

1. Data Acquisition Layer (Perception): Comprises RGB, hyperspectral, or thermal cameras for image capturing; as well as humidity, temperature, pH, moisture, and light intensity sensors.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume X Issue VII July 2025

2. Edge Intelligence Layer: Runs pre-trained and optimized AI models on embedded platforms such as Jetson Nano, Coral TPU, or Raspberry Pi for real-time processing and decision-making.
3. Communication Layer (Network): Facilitates the accurate transfer of data between edge devices, gateways, and/or cloud systems using LoRa, ZigBee, NB-IoT, or Wi-Fi protocols.

CROPCARE [1], for example, combines an RGB camera, Jetson Nano edge device, and Wi-Fi for in situ real-time leaf disease detection and alert dissemination. It does inference in real-time in the field to prevent latency due to cloud reliance.

Rumly et al. [2] created an energy-efficient solution based on Raspberry Pi and LoRaWAN, suitable for farms with low connectivity. The solution was able to achieve early rice disease detection in remote agricultural areas, showing applicability for resource-poor regions.

In contrast, Hayajneh et al. [4] introduced a UAV-based TinyML system, fusing aerial data collection with local ML inference onboard for widespread farmland coverage. This exemplifies the ability to combine mobile edge computing with agriculture.

This decentralization—local processing instead of cloud—lowers latency considerably, improves scalability, and maintains data privacy, all essential in farm deployments in rural areas.

Lightweight Model Development

Edge AI models are usually bounded by small memory, computation, and battery life. To be efficient, researchers apply techniques of architectural and optimization to minimize computational complexity with retained performance.

- Depthwise Separable Convolutions (e.g., MobileNet) lower the number of parameters and computation by decomposing convolutions [3].
- Model Pruning removes less significant weights and neurons, leading to sparse and reduced models with little performance degradation [4].
- Quantization diminishes numerical accuracy (e.g., from 32-bit to 8-bit integers), making the models smaller and quicker, usually TensorFlow Lite or ONNX Runtime compatible [5].
- Neural Architecture Search (NAS) finds the best model architectures that optimize accuracy relative to hardware efficiency [6].

For instance, Zhang et al. [3] showed that a depthwise CNN could classify leaf diseases with more than 90% accuracy and <500 ms latency on a Raspberry Pi. This indicates that accuracy is preserved even at the expense of small models.

Wicaksono and Apriono [8] executed a tailored CNN model on an FPGA-edge system. The model was 98% accurate with very low power consumption and sub-second latency—emphasizing the role of hardware-aware design.

Furthermore, Lv et al. [11] contrasted EfficientNet-B0, MobileNetV3, and SqueezeNet on peanut leaf classification. The trade-offs indicated MobileNetV3 produced the optimal inference time (80 ms) and power efficiency while having >91% accuracy.

Data Handling and Augmentation Techniques

Agri-datasets tend to be small, noisy, and imbalanced with poor disease variety or environmental variety. This is a hindrance for generalization across regions or seasons. For this purpose, a few data-driven AI techniques are employed:

a) Conventional Augmentation :

Mimics geometric (rotation, flip), photometric (brightness, contrast), and additive noise methods for enhancing model immunity to true-world conditions.

b) Synthetic Data using GANs :

LeafGAN [9] produces high-quality, labeled leaf images that capture rare or under-sampled diseases. This addresses class imbalance and enhances model generalization.

c) Transfer Learning :

ResNet, InceptionV3, or EfficientNet models are pre-trained on a large-scale dataset (e.g., ImageNet) and then fine-tuned using small agricultural datasets. Gurunathan et al. [7] employed EfficientNet-B0 and achieved 94% accuracy in cassava leaf diseases with little training data.

d) Continual Learning :

Santos et al. [10] used Elastic Weight Consolidation (EWC) for plant disease models to learn incrementally from new field data without forgetting past knowledge—a real-world solution for settings with dynamic disease patterns.

e) Meta-learning & Few-shot Learning :

These methods are yet untapped in agriculture but would enable models to learn to generalize from just a few labeled examples per disease category—perfect for emerging or low-frequency diseases.

Such data-driven approaches greatly improve the field-deployability of AI models, particularly in data-poor, high-variance agricultural environments.

Deployment and Optimization Strategies

Edge deployment of AI is more than inference accuracy. Real-world considerations are:

- Model Conversion & Optimization: Models are converted into forms such as TensorFlow Lite, TensorRT, or ONNX to enable hardware acceleration.
- Hardware Compilation: For microcontrollers (e.g., STM32), cross-compilation is needed using toolchains such as Arduino IDE or Edge Impulse.
- Performance Benchmarking: Compares FPS (frames per second), model size (MB), RAM consumption (MB), power consumption (W), and inference latency (ms).

Lv et al. [11] compared several models and determined MobileNetV3 to be best suited for deployment in the field with Jetson Nano and Coral Edge TPU. Da Silva and Almeida [12] used thermal + RGB sensor fusion with a quantized CNN model, enhancing accuracy in varying illumination conditions.

Also, deployment in the real world tends to leverage Docker containers or edge platforms such as EdgeX Foundry, AWS Greengrass, or Azure IoT Edge for efficient and manageable rollout of AI pipelines at the field level.

Sensor Fusion and Communication

Secure and efficient communication is critical to exchange sensor information, model updates, and notify users. Based on deployment environment and scale:

- LoRaWAN is ideal for wide-area, low-power connectivity for distant areas [2], [13].

- ZigBee and NB-IoT provide short-range, low-data-rate transmission ideal for greenhouses.
- Wi-Fi or BLE is used for high-throughput but less power-efficient communication, suitable for limited-range scenarios.

Sensor fusion plays a growing role in improving detection accuracy and reducing false alarms. Strategies include:

- Early Fusion: Combining multiple sensor inputs at the model input layer (e.g., image + temperature).
- Late Fusion: Independent model predictions from multiple modalities are merged at the decision level.
- Hybrid Fusion: Merges early-stage and late-stage fusion under one single framework.

Das and Luo [14] presented LightESD, which employed soil, humidity, and temperature fused data to identify environmental anomalies related to plant stress. The model performed better than single-sensor baselines with negligible latency.

Sophisticated sensor integration (e.g., hyperspectral, thermal, acoustic) with edge inferencing promises opportunities for multi-dimensional disease diagnosis and context-aware agricultural systems.

State-of-the-art real-time plant disease detection systems based on Edge AI and IoT platforms integrate effective hardware, dynamic software, and strong data approaches. The strategies detailed show that deep learning optimization, continual learning, sensor fusion, and communication schemes can make precision agriculture scalable and cost-effective. Challenges persist in developing integrated benchmarks, model interpretability, and dependable edge deployment across real-world uncertainties—discussed in the next section.

Challenges and Open Issues

The convergence of Edge AI and IoT into agricultural infrastructure for real-time detection of crop diseases offers disruptive potential, but this shift is still met with multi-dimensional challenges. The challenges include technical, environmental, infrastructural, and ethical aspects, which have to be overcome for the large-scale, scalable, and efficient implementation.

Dataset Scarcity and Quality Limitations

Agricultural AI models significantly depend on large amounts of well-annotated data. The availability of high-quality and diverse agricultural datasets is still scarce. The majority of public datasets:

- Only cover a small number of crop species (i.e., tomato, cassava, or maize),
- Only capture limited disease types and early-stage symptoms,
- Include images under controlled lighting conditions, which are unrealistic for field deployment.

For instance, the widely used PlantVillage dataset [1] consists of artificially clean data causing poor real-world generalization. Also, annotation is time-consuming and needs domain knowledge from agronomists or plant pathologists. This bottleneck diminishes the feasibility of training high-performing and generalizable models.

Synthetic data augmentation via Generative Adversarial Networks (GANs) [2] or methods such as Few-Shot Learning and Self-Supervised Learning provide limited solutions but tend not to include contextual realism in created samples. Community-sourced, geo-diverse, and multimodal datasets, including those with RGB, thermal, hyperspectral, and contextual environmental data, are needed now.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume X Issue VII July 2025

Lack of Cross-Domain and Seasonal Generalization

Edge AI models tend to overfit to a particular domain, for example, one type of crop or location. Deploying such models in a different region causes them to perform far worse because:

- Crops have a different appearance,
- Disease expression is varied when exposed to varying climatic and soil conditions,
- Plant morphology or color changes by season.

This problem is known as domain shift, and it is a focal issue for real-time disease diagnosis in diverse farm settings. Standard transfer learning is beneficial to some degree, but real-time applications require domain-agnostic or meta-learning methods [3] to learn to adjust dynamically. Furthermore, weather and seasonal change also affect leaf color, light exposure, and sensor noise, so seasonal robustness is another open problem.

Limited Computational Resources at the Edge

Field-deployed edge devices (e.g., Raspberry Pi, ESP32, Jetson Nano, Arduino) have severe constraints with regard to:

- Low CPU/GPU frequency for computational power,
- RAM and storage capacity,
- Availability of battery or solar power,
- Absence of active cooling in outdoor environments.

While methods like model pruning, quantization, knowledge distillation, and depthwise separable convolutions are applied to reduce the model footprint [4], these tend to have a performance compromise. For instance, it might be possible to deploy a MobileNetV2 on a Jetson Nano, but it would still burn a lot of power and may not be scalable to multi-sensor processing.

Custom lightweight architectures are needed, potentially based on neural architecture search (NAS) that is energy, latency, and accuracy optimized. Hardware-software co-design for TinyML and edge-AI chips like Google Coral TPU and Intel Movidius needs to be pursued further to ensure sustainable deployment.

Real-Time Processing Constraints and Latency Issues

Real-time processing is essential for timely disease detection that helps limit crop loss. Yet the latency of the system is impacted by:

- Image capture and pre-processing latency,
- Model inference latency on edge hardware,
- Overheads for communication from sensors to gateways/cloud,
- Power modes for power-saving hardware in remote settings.

Real-time systems need to provide inference in <1 second, but most deep models need multi-frame averaging or large inputs. Optimizing real-time requirements with accurate prediction is a continuing challenge [5]. Pipeline parallelism, early exit networks, or approximate computing can potentially provide solutions.

Multi-Sensor Data Fusion Challenges

Combining various modalities such as temperature, humidity, soil moisture, multispectral images, and geolocation can make disease detection more accurate. But combining these heterogeneous data streams brings along:

- Time synchronization challenges,
- Sensor calibration drift over time,
- Amplification of noise when modalities disagree,
- Difficulty in designing multi-modal deep networks [6].

Today, all systems are based primarily on visual information, and other agro-climatic indications are not considered. Sensor-aware learning paradigms for robust operation even when some modalities are unavailable or noisy need to be developed for practical application.

Power and Energy Efficiency

Rural deployments tend to use solar panels or batteries, and power availability is the main limitation. Models that require large amounts of computational energy do not support long-term deployment. Although energy-efficient AI models are researched [7], few have been implemented in agricultural environments.

Periodic data transmission (e.g., using LoRaWAN or ZigBee) and sensor activity also utilize energy. There is a trade-off among:

- Detection frequency (temporal resolution),
- Power consumption, and
- Network bandwidth.

Techniques such as event-driven sensing, dynamic model compression, and edge device sleep-wake cycles can enhance the operation life, but such techniques require fine-grained adjustments for each deployment zone.

Security, Privacy, and Data Integrity

Agriculture systems are cyber-physical systems now, and they are more exposed to:

- Sensor tampering or spoofing,
- Adversarial image attacks for deceiving AI,
- Unauthorized data access or leakage during cloud backup,
- Poisoning of on-device learning models.

Very few systems use secure authentication, lightweight encryption, or tamper-proof logging facilities. Use of blockchain for the secure logging of sensor events, and federated learning for privacy-preserving model training without moving raw data are unexplored [8].

Absence of Standardized Evaluation Metrics and Field Trials

There are no standard benchmarks presently available for assessing real-time crop disease detection systems. Researchers employ:

- Various datasets (some of which are private),
- Varied environmental conditions,
- Measures such as precision, accuracy, and recall, but seldom latency or energy consumption.

In addition, a majority of the works report lab or test bed results. Real-farm deployment under changing weather, lighting, dust, and biological interference are hardly validated. Without this, real-time detection claims are very much speculative.

Open benchmarks like FarmSim environments or multi-site collaborative testbeds can go a long way to mature the field.

System Maintenance and Scalability

Deploying and managing edge devices in large farms or a number of farms is not trivial. Challenges are:

- Device failure or overheating,
- Firmware and model update management,
- Sensor drift calibration,
- Remote diagnostics and fix.

The absence of standardized edge orchestration tools like KubeEdge, Balena, or Azure IoT Edge hinders scalability. Studies on autonomous system healing, remote update, and dashboard analytics for farm IoT require more focus.

Socio-Technical and Economic Barriers

Lastly, even when technical systems operate well, adoption in real-life settings is restricted by:

- Farmer trust in AI systems, particularly when they make incorrect predictions,
- Hardware cost for small-scale farmers,
- Insufficiency of technical literacy to control edge devices,
- Language and interface barriers.

Mechanisms must be engineered with human-in-the-loop processes, multiple languages, and visual notifications (e.g., LED lights or basic phone messages). Affordability and ease are crucial in adoption.

Future Directions

Edge AI and IoT-based solutions will increasingly be pivotal in early and real-time crop disease detection as precision agriculture continues to develop. But existing implementations remain emergent, fragmented, and mostly confined to pilot initiatives. The following presents primary areas for innovation and development that need to be met in order to achieve scalable, sustainable, and farmer-friendly crop monitoring systems.

Unified AI Frameworks for Multitask Learning

Existing systems are optimized for single-task detection, commonly for one crop or one disease. In actual agricultural environments, there exist concurrent problems—disease detection, pest forecasting, irrigation

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume X Issue VII July 2025

requirements, and nutritional deficiencies—whose model should be developed. Future work must develop multi-task deep learning models that are able to:

- Deal with various plant species and diseases,
- Concurrently detect weeds, pests, and abiotic stresses,
- Forecast growth stages and yield under changing conditions.

Transformers, vision-language models, and graph neural networks (GNNs) might be combined for this role, tapping both sensor inputs and expert-curated knowledge graphs [1].

Edge-Cloud Collaborative Intelligence

Instead of being dependent only on the edge or the cloud, next-generation architectures should implement a hybrid intelligence paradigm where:

- Important tasks (e.g., real-time disease detection) are done locally,
- Computation-intensive activities (e.g., model training or anomaly clustering) are shifted to the cloud,
- Synchronization between cloud and edge is network availability-optimized.

This edge-cloud continuum would make dynamic decision-making possible, with systems learning from worldwide patterns but adjusting to local farm environments. Dynamic model partitioning and edge caching are some of the techniques that will be used for smooth operation [2].

Real-Time Feedback and Actuation Loops

Present systems are only able to detect disease but not mitigate it with human intervention. Future systems should incorporate closed-loop actuation, for example:

- Enabling drone or robot pesticide spraying,
- Automating irrigation or ventilation controls,
- Issuing disease alarms with site-specific action suggestions.

By integrating sensor inputs, AI algorithms, and actuators, farms can progress toward autonomous precision farming. Actuator integration, for example, IoT-capable sprayers or automated irrigation valves, is the next requirement.

Resilient AI Against Environmental Variability

Agricultural environments are dynamic and stressful: changing weather, light, and soil conditions can impair model performance. Subsequent work should center on:

- Self-adaptive AI systems that retrain or recalibrate based on feedback,
- Robust training pipelines with environmental simulations (e.g., brightness jitter, occlusion noise),
- Models that estimate uncertainty in predictions through Bayesian deep learning or ensembling [3].

These advances will make AI systems robust under real-field, varying conditions.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume X Issue VII July 2025

Bio-Sensor and Non-Visual Modalities

Visual data (RGB or hyperspectral images) are mostly all that the research is concerned with. But the diseases might not be visible at the early stages. Adding non-visual modalities like:

- Volatile Organic Compound (VOC) sensors that can sense plant stress based on gas emission,
- Acoustic sensors to sense insect presence or rustling of leaves,
- Soil electrochemical sensors for detecting early signs of fungal activity.

These, combined with vision systems, will allow early, multimodal diagnosis prior to visual symptom appearance [4].

Synthetic Field Data and AI-Driven Crop Simulation

Lack of data is a major bottleneck. Directions for the future must make use of:

- Digital Twin technology to virtually model farms and crop dynamics under different disease and weather conditions,
- Synthetic dataset creation based on physics-informed GANs and crop growth simulators (e.g., APSIM, DSSAT),
- Virtual training and testing of edge-AI systems using Augmented Reality (AR) and VR platforms.

These tests will allow for secure testing of AI systems in simulated environments prior to deployment and assist in training models in uncommon or high-risk disease situations [5].

Regional and Cultural Tailoring of Systems

Farmer adoption and trust are context-dependent at a local level. Next-generation systems should be:

- Multilingual, with support for regional dialects and languages,
- Culturally adaptive, aware of the local farming calendar, beliefs, and local practices,
- Voice-based or visually based, based on farmer choice or literacy levels.

This will involve user-focused design research, co-creation with agricultural communities, and cooperation with local extension services or cooperatives [6].

Affordable and Sustainable AI Hardware

Edge devices should be affordable, energy-efficient, and sustainable. Future research needs to investigate:

- Recyclable and biodegradable materials for edges,
- Modular device designs where sensors and AI chips can be individually replaced,
- Open-hardware projects such as OpenEdgeAI that decrease dependence on proprietary platforms.

Furthermore, village-maintained device hubs can offer collective AI functionality to smallholder farmers without needing a device per farm.

Agri-Cybersecurity and Tamper-Resistant Systems

Agriculture being digitalized, it gets exposed to cyberattacks. Future systems should include:

- Lightweight intrusion detection systems (IDS) on edge devices,
- Secure boot protocols and encrypted firmware updates,
- Decentralized ledgers (e.g., blockchain) to authenticate AI decisions and sensor logs
- AI for anomaly detection to alert for possible data spoofing or manipulation.

Policies need to be created around agri-data sovereignty to ensure that farmers own and have control over their sensor and crop data [7].

Interdisciplinary Cooperation and Policy Facilitation

Combating agricultural issues calls for cooperation from:

- AI researchers and agri-scientists,
- Climate scientists and ecologists,
- Economists and policy developers.

Multidisciplinary consortia need to form in future research, developing common standards, open platforms, and regulatory rules for:

- Data sharing and annotation,
- AI model validation and certification,
- Technology adoption incentives and subsidies for smallholders.

Initiatives such as the FAO-ITU E-Agriculture Framework and UN AI for Good need to be harnessed to catalyze scalable, impactful innovation [8].

DISCUSSION

The combination of Edge AI and IoT technology in agricultural systems marks a revolutionary leap into smart disease detection and real-time crop disease monitoring. As literature surveyed thus far explains, this integration could offer a method of plant disease identification that promises to be less manpower consuming and more time efficient than conventional methods. Nonetheless, advances in this area are vital but have many observations at the critical level against the combined development and limitation of the future directions in the field.

The trade-off between accuracy of the model and efficiency of the hardware is one of the strongest themes prevalent in the reviewed works. On one hand, many deep learning models like CNNs, EfficientNet, and MobileNet achieved high-level accuracy in plant disease image classification. On the other hand, they consume large computational and memory resources, which makes them less deployable in low-power edge devices, which are widely available in rural and resource-limited settings. Researchers have resorted to techniques such as in deploying model pruning, quantization, and using lightweight architectures as a counter for this. However, such optimizations tend to realize a drop in accuracy, particularly after deployment of models in real-world field situations, which differ immensely from where they were trained.

Dependability and reproducibility of AI models across diverse agricultural environments are another area of contention. Although some models show good performance in test sets, they are not guaranteed to work in the real-world outdoor environment, which can include variation in lighting, occlusion, backgrounds, and the presence of multiple symptom complexities. The disparity suggests much reliance on evaluation approaches that are simulated or laboratory-based and calls for real-world verification. The lack of unified datasets as well as evaluation procedures poses a barrier for adequate comparisons to be made between studies and hampers reproducibility, which is important for scientific progress and industrial uptake.

Data importance is echoed when taking into consideration the scant availability of high-quality and large-scale annotated datasets reflecting diversity. Most research studies are conducted based on small or artificial data sets that do not exhaustively capture the diversity of disease expression across several crop types and geographies, and climates. The assumption of generalization will most likely not hold, which creates a disconnect between research results and implementation in the real world. Therefore, in that aspect, open, multimodal, world representative data sets are critically important for future empowerment from this perspective.

Another of those aspects is AI, IoT, and system level integration for decision making in real time. Current systems are concentrated mainly in the perception layer, for application in disease detection, while they do not have any higher-level decision-making and actuation elements. An end-to-end edge-AI pipeline would have to incorporate not only detection but also contextual analysis, user input, and interface with farm management systems. Integration is needed to go from diagnosis to action-facilitating timely actions such as spraying, irrigation, or quarantining infected zones.

Besides, the socio-economic implications from greater use of such systems cannot be ignored. Though the possible benefits for smallholder farmers are substantial-from yield protection to reduced inputs-the barriers of cost, technical illiteracy, trustworthiness of AI predictions, and infrastructure constraints still remain. The success of these systems is also not just linked to technological advances but also the extent to which they are accessible, explicable, and beneficial to the intended end-users. As important, multilingual support, culturally relevant interfaces, and participatory design approaches are critical for promoting uptake at the grassroots level.

Another looming concern is the privacy, ownership in data, and cybersecurity issue concerning smart agricultural ecosystems. With farms getting more and more connected, safeguarding sensitive information like crop health, or farm location, yield will hugely present a priority. Incorporation of secure communication protocols with blockchain-based logging and privacy-preserving artificial intelligence methods like federated learning will make up the key ways in trust development at such systems.

Finally, the increase of necessity in interdisciplinary cooperation. Well successively built, put into context, and deployed scalable, context-aware, and sustainable solutions must factor in all the expertise of AI researchers, agricultural scientists, hardware engineers, social scientists, and policymakers.

However, Edge AI and IoT's potential revolutions can be realized in detecting diseases in agriculture. Such potential should, however, be understood with respect to the challenges heaped against real opportunities. The future of technology requires not only advancements but also careful integration, inclusiveness, and responsiveness to the rich variances of contemporary agriculture.

CONCLUSION

The future of this new paradigm is going to be thrilling revolution in precision agriculture regarding prompt hack-aggression against pest diseases-commercially sold against crops with Edge AI , most notably with the Internet of Things (IoT). The surveys will now be focused on altering the classes of research, architectures, models, and related mechanisms of deployment, which together add to the accoutrements of real-time, low-latency, and energy-efficient design monitoring systems.

Meanwhile, the work increasingly orients itself towards a near real-time character for low-power and energy-efficient in-field crop monitoring systems. Some examples would include lightweight deep learning models that would work in edge-based applications and the application frameworks created for such applications from other data from environmental and/or image data sources. Through these standards of improvement of communication as well as system-level improvements for system integration, already build a real-time, low-power crop monitoring system even under extreme and under-resourced environments.

Great milestone, but there lies another mountain to climb. The models, however, in this case, are shown not to have enough robustness to deal with practical cases found in agriculture. On the contrary, these models perform very poorly regarding robustness in the generalization aspect across crop type, microniche, and disease variation. As a third hurdle, most of the AI models fail in training and validation due to the unavailability of well-annotated datasets in agriculture.

The issues of privacy and security include questions on socioeconomic access to such solutions based on edge technology before their wider uptake or engendering of trust towards them. Future-oriented research must integrate advances in representative global benchmark datasets with explainable and privacy-preserving AI models for seamless integration into farm management systems.

Generally working under inclusive design paradigms considering the needs and constraints of smallholder farmers as well as interdisciplinary collaboration will also help a long way in unlocking the value these technologies promise to deliver.

In summary, IoT and Edge AI open up new avenues for disease diagnosis in agriculture and, through high-level research and innovative thinking in terms of systems-level impact, empower farmers with timely information and reduced crop losses while battling the world challenge of hunger.

REFERENCES

1. S. Garg, M. Rani, and P. Kumar, “CROPCARE: Edge-Based Real-Time Disease Detection in Agriculture Using Deep CNN,” IEEE Internet of Things Journal, vol. 8, no. 6, pp. 5020–5029, 2021.
2. A. Rumly et al., “IoT-Enabled Rice Disease Detection Using Raspberry Pi and LoRaWAN,” IEEE Access, vol. 9, pp. 33455–33467, 2021.
3. M. Proietti et al., “Greenhouse Crop Monitoring with Edge AI: A Case Study in Tomato Disease Detection,” Computers and Electronics in Agriculture, vol. 184, 2021.
4. D. Silva and M. Almeida, “Thermal Imaging and Edge AI for Crop Stress Detection,” Sensors, vol. 21, no. 14, 2021.
5. X. Chen et al., “Inception-Based CNN for Plant Disease Identification on Edge Devices,” IEEE Sensors Journal, vol. 21, no. 3, pp. 3456–3464, 2021.
6. M. Rani and S. Gowrishankar, “Transfer Learning with EfficientNet for Smart Farming Applications,” IEEE Access, vol. 8, pp. 115310–115321, 2020.
7. R. Hosny, F. Osman, and H. El-Sayed, “CNN-LBP Fusion Model for Efficient Crop Disease Classification,” Information Processing in Agriculture, vol. 9, no. 1, 2022.
8. L. Liu et al., “YOLOX-Light: An Edge-Deployed Detector for Cotton Leaf Disease,” IEEE Transactions on Industrial Informatics, vol. 19, no. 2, pp. 1256–1267, 2023.
9. N. Noon, T. Wang, and A. Kumar, “Quantized CNNs and Pruning for Agricultural Edge AI,” ACM Transactions on Sensor Networks, vol. 19, no. 1, 2023.
10. A. Cap, V. Zhang, and L. Qin, “LeafGAN: Generative Augmentation for Plant Pathology Datasets,” IEEE Access, vol. 8, pp. 221985–221996, 2020.
11. M. Alharbi, S. Singh, and R. Thomas, “Continual Learning in Crop Disease Detection Using EWC,” Computers in Biology and Medicine, vol. 147, 2022.
12. S. Subana Begum, K. Hussain, and R. Narayanan, “Solar-Powered Edge Nodes for IoT Agriculture Monitoring,” International Journal of Distributed Sensor Networks, vol. 18, no. 3, 2022.
13. L. Yu et al., “Energy-Efficient Communication in LoRaWAN-based Smart Agriculture,” IEEE Internet of Things Journal, vol. 9, no. 15, pp. 12846–12859, 2022.
14. B. Jin, K. Tan, and P. Zhao, “Sensor Fusion with IoT Edge for Crop Condition Assessment,” IEEE Sensors Journal, vol. 22, no. 8, pp. 7965–7973, 2022.
15. R. Ahmed, J. Patel, and M. Saxena, “Pruned Deep Models for Edge-Based Crop Classification,” IEEE Transactions on Sustainable Computing, Early Access, 2023.

16. V. Gupta and A. Shastri, "Low-Latency Inference for Smart Farming Using TinyML," IEEE Embedded Systems Letters, vol. 15, no. 1, pp. 34–37, 2023.
17. T. Wei et al., "Deployment of Lightweight Neural Networks on Microcontrollers for Plant Monitoring," IEEE Access, vol. 10, pp. 124934–124948, 2022.
18. K. He, Y. Sun, and Q. Dong, "Federated Learning for Crop Disease Detection Using IoT Data," IEEE Internet of Things Journal, vol. 9, no. 12, pp. 9851–9863, 2022.
19. Y. Qiu, S. Fan, and D. Li, "Adaptive Learning for Agricultural AI Under Data Scarcity," IEEE Transactions on Neural Networks and Learning Systems, Early Access, 2023.
20. P. Tan and H. Shen, "Multimodal Smart Sensing for Disease Detection in Crops," IEEE Transactions on Instrumentation and Measurement, vol. 72, 2023.
21. FAO and ITU, "E-Agriculture in Action: Big Data for Agriculture," Food and Agriculture Organization of the United Nations and International Telecommunication Union, Tech. Rep., 2021.