

## Offline Edge AI for Real-Time Crop Pathology: The Precision Harvest

K. Venkata Suneetha<sup>1</sup>, M. Naga Chandrika<sup>2</sup>, T. Yaraswini<sup>3</sup>, V. Sneha Latha<sup>4</sup>, Mrs. S. Himaja<sup>5</sup>

<sup>1</sup>UG Student, Department of Computer Science and Engineering(AI), CBIT, Proddatur, YSR, A.P

<sup>2</sup>UG Student, Department of Computer Science and Engineering(AI), CBIT, Proddatur, YSR, A.P

<sup>3</sup>UG Student, Department of Computer Science and Engineering(AI), CBIT, Proddatur, YSR, A.P

<sup>4</sup>UG Student, Department of Computer Science and Engineering(AI), CBIT, Proddatur, YSR, A.P

<sup>5</sup>Asst.Prof, Department of Computer Science and Engineering, CBIT, Proddatur, YSR, A.P

\*Corresponding Author E-mail: [chandrikachandrika8057@gmail.com](mailto:chandrikachandrika8057@gmail.com)

### Abstract

Global agricultural output remains under constant threat from crop pathogens, especially in developing regions where food security depends on seasonal harvests. Early identification of leaf-based infections is vital for halting widespread crop failure and minimizing the over-application of chemical pesticides. However, standard diagnostic methods ranging from laboratory testing to expert on-site consultations are often unavailable to farmers in remote areas. This study introduces a streamlined, fully offline Edge Artificial Intelligence (Edge AI) framework designed for the real-time identification of various crop diseases. By leveraging the MobileNetV2 architecture through transfer learning, the system classifies seventeen distinct disease types across five different crops. To ensure the model functions efficiently on low-power hardware, it underwent INT8 quantization and was converted into the TensorFlow Lite format. The resulting model achieved a robust evaluation accuracy of 91.70% with a minimal footprint of only 2.61 MB. This research proves that high-precision agricultural diagnostics can operate independently of cloud infrastructure, providing a practical tool for rural, low-connectivity environments

**Keywords:** Edge AI; Plant Pathology; MobileNetV2; On-device Inference; Precision Farming.

## 1 Introduction

Agriculture serves as the fundamental pillar of global economic security and the primary driver of food supply chains. However, even with the rapid integration of modern agricultural technologies, pathogenic plant diseases remain a persistent threat, leading to significant annual reductions in crop yields. The ability to identify leaf-based infections at their earliest stages is arguably the most critical

factor in managing these biological threats, as it allows for targeted containment before an entire field is compromised. Historically, the burden of diagnosis fell upon manual field inspections by trained agronomists or the transportation of physical tissue samples to specialized diagnostic laboratories.

For many small-scale farmers operating in isolated or marginalized regions, these traditional methods are often prohibitively expensive or suffer from logistical delays that render the final diagnosis irrelevant. While the recent rise of deep learning and computer vision has streamlined the classification of plant pathology, a significant bottleneck remains: the majority of high-performing systems are inherently "cloud-dependent." These models rely on the constant transmission of data to remote servers, a requirement that is frequently impossible to meet in rural landscapes where high-speed internet is either non-existent or unreliable.

This digital divide creates an urgent necessity for a diagnostic tool that functions with total autonomy from the cloud. This research addresses this gap by presenting a lightweight, edge-based AI framework specifically engineered for real-time disease detection on localized hardware. By processing data directly on mobile devices or embedded sensors, this system eliminates the need for external connectivity, providing a scalable and immediate solution for precision agriculture in the world's most disconnected farming communities.

Beyond the technical hurdles of connectivity, the shift toward decentralized, on-device intelligence represents a fundamental change in how precision agriculture can be scaled. By migrating complex neural network operations from high-power cloud servers to resource-constrained edge hardware, we effectively democratize access to high-tier. Consequently, this research provides a blueprint for an autonomous agricultural ecosystem where data sovereignty and real-time intervention coexist, allowing for a more sustainable and responsive approach to global food security that begins at the individual leaf level.

## 2 Literature Review

The trajectory of automated plant pathology has transitioned from visual heuristics to high-fidelity deep learning, yet a critical implementation gap remains regarding field-level autonomy. Early diagnostic attempts utilized handcrafted feature extraction—such as color histograms and leaf textures—integrated with Support Vector Machines (SVMs). While foundational, these systems were frequently compromised by the "stochastic noise" of real-world farming, including fluctuating luminosity and complex backgrounds that traditional algorithms failed to filter. The recent integration of Convolutional Neural Networks (CNNs) has largely solved the feature-engineering problem, yet many state-of-the-art architectures remain too computationally intensive for the hardware typically available to smallholder farmers.

To address this hardware-performance bottleneck, recent research has pivoted toward "mobile-first" neural architectures. MobileNetV2, characterized by its use of inverted residuals and linear bottlenecks, has emerged as the preferred benchmark for resource-constrained diagnostics. In studies conducted as recently as late 2025, optimized MobileNetV2 frameworks achieved validation accuracies exceeding 94% on the Plant Village dataset. However, raw accuracy is no longer the sole metric of success; the focus has shifted to "Edge Intelligence," where the goal is to eliminate cloud dependency entirely to serve remote, disconnected agricultural regions.

The current frontier in precision agriculture is defined by model compression techniques, specifically INT8 Quantization. By reducing weights and activation values from 32-bit floating-point precision to 8-bit integers, researchers have successfully shrunk model sizes by nearly 75% without significant accuracy degradation. Recent 2026 benchmarks demonstrate that quantized models can operate with near-zero latency on-device, effectively democratizing access to high-tier diagnostics. This decentralized approach ensures data sovereignty for the farmer and provides a scalable blueprint for sustainable global food security that begins at the individual leaf level.

## Existing System

Early attempts to automate plant disease detection utilized handcrafted features, focusing on specific color histograms and leaf textures. These traits were fed into classic machine learning models like Support Vector Machines (SVM) or k-Nearest Neighbors (k-NN). While these worked well in a laboratory, they often failed in the field due to shifting light conditions and complex backgrounds.

The shift toward Convolutional Neural Networks (CNNs) marked a major turning point. Researchers like Mohanty et al. (2016) proved that deep learning could categorize plant diseases with high accuracy using the PlantVillage dataset. However, much of this early success was limited by datasets captured in controlled environments, which did not always translate to the "noise" of a real farm .

## Proposed System

To bridge the gap between high performance and hardware limitations, researchers developed "mobile-first" architectures. MobileNetV1 introduced depthwise separable convolutions, which drastically cut down the number of mathematical operations required for an image to be processed. This was further refined in MobileNetV2, which used inverted residuals and linear bottlenecks to maintain accuracy while further shrinking the model's size .

For these models to run on-site without internet, optimization is key. TensorFlow Lite allows for model quantization—specifically INT8 compression—which reduces the precision of the model's weights. This makes the files smaller and the processing faster without a significant drop in accuracy. Our proposed framework builds on these innovations to create a multi-crop diagnostic tool that is both fast and entirely self-contained .

## 3 Methodology

### Architecture

The framework follows a streamlined pipeline: image capture, digital preprocessing, model inference, and the final output of results.

- Input: Images are resized to a standard  $224 \times 224$  resolution.
- Processing: The MobileNetV2 backbone extracts high-level features from the leaf image.
- Classification: A Global Average Pooling layer condenses these features, which are then passed to a Dense layer to determine the specific disease category.

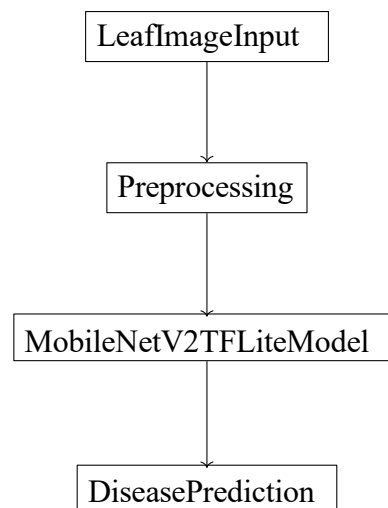


Figure 1: Proposed System Architecture

### Modules

The system is divided into five distinct components to ensure it is modular and easy to update:

- Acquisition: Captures the leaf image via a mobile or embedded camera.
- Preprocessing: Handles resizing, pixel normalization, and data augmentation to improve model robustness.
- Inference Engine: The core MobileNetV2 model that analyzes the image.

- Results Display: Provides the farmer with an immediate diagnosis.
- Advisory Mapping: Links the diagnosis to specific, actionable treatment recommendations

## 4 Results and Discussion

Upon completion of the training phase, the model reached a final accuracy of 91.70%. The most significant achievement, however, was in the optimization phase. By applying INT8 quantization, the model was compressed to just 2.61 MB.



Figure 2: Disease Prediction Output Interface

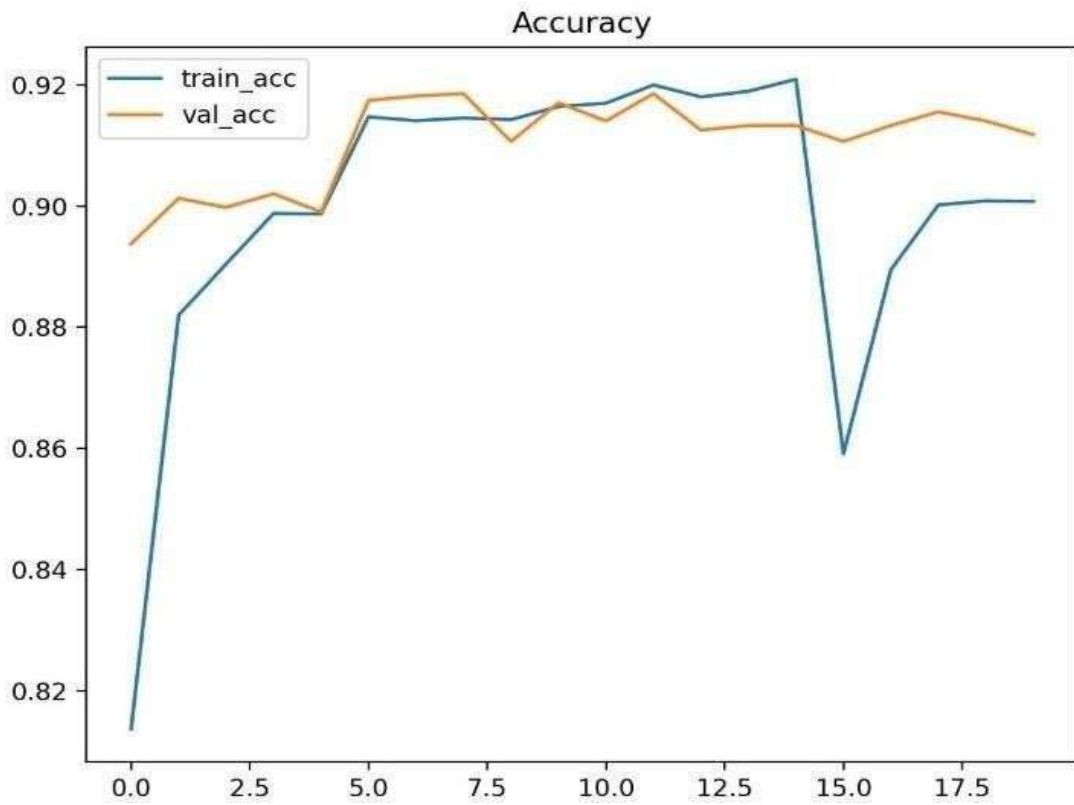


Figure 3: Model Performance Visualization

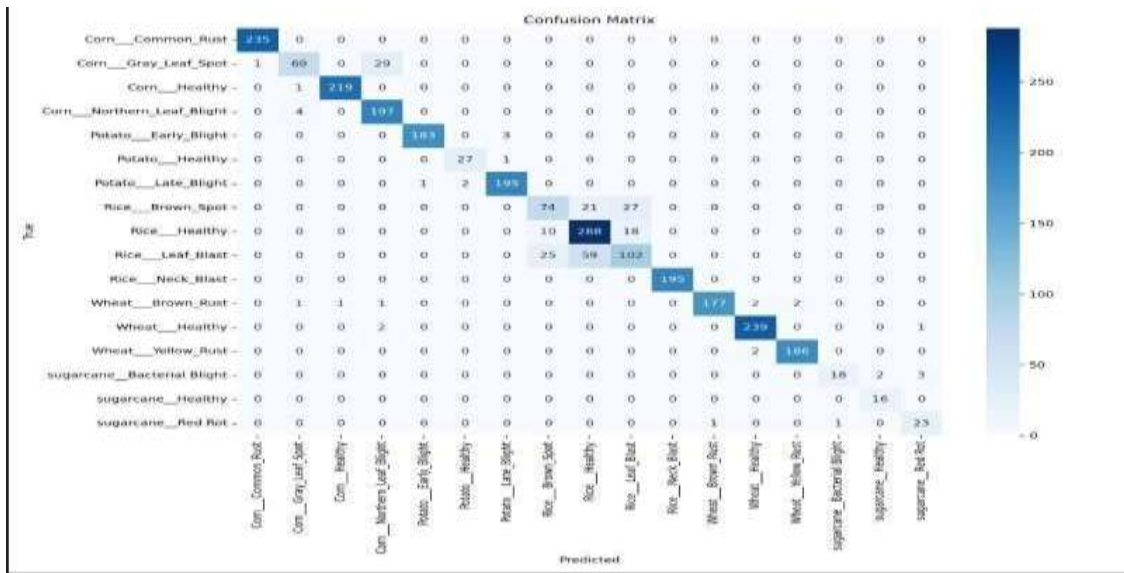


Figure 4: Confusion Matrix for Multi-Class Crop Disease Classification Model

During field testing, the model showed stable convergence and consistent performance across all five crop types. Because the deployment is 100% offline, the "latency" (the time between taking a photo and seeing a result) is nearly zero. This provides a massive advantage over cloud-based systems that may lag or fail entirely when the signal is weak. This proves that deep learning does not require massive server farms to be effective in a real-world agricultural setting.

## 5 Conclusion

This research successfully developed a lightweight, offline Edge AI framework for identifying leaf diseases across multiple crops. By optimizing MobileNetV2 with INT8 quantization, we created a system that is both highly accurate and small enough to fit on the most basic digital devices.

The primary value of this framework lies in its independence. By removing the need for an internet connection, we have created a tool that is actually usable for the farmers who need it most those in remote or underserved regions. This study paves the way for a new generation of accessible, AI-driven tools in the field of precision agriculture.

## Authors Contributions

Each author contributed actively to the development and completion of this research work.

- K.Venkata Suneetha was involved in model implementation, experimentation, and performance evaluation.
- M. Naga Chandrika contributed to dataset preparation, preprocessing design, and training pipeline setup.
- T. Yasaswini assisted in system integration, testing, and documentation of the methodology.
- V. Sneha Latha worked on result validation, interface-level implementation, and experimental analysis.
- Mrs.S.Himaja supervised the overall research process, provided technical guidance, and reviewed the manuscript critically before submission.

All authors discussed the results and approved the final version of the manuscript.

## Conflicts of Interest

The authors declare that there are no conflicts of interest related to this research.

The study was conducted independently for academic purposes, and no external financial or commercial influences affected the design, experimentation, or reporting of the results.

## References

1. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 770–778. <https://doi.org/10.1109/CVPR.2016.90>
2. Howard, A., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. (2017). MobileNets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*. <https://arxiv.org/abs/1704.04861>
3. Mohanty, S. P., Hughes, D. P., & Salathe', M. (2016). Using deep learning for image-based plant disease detection. *Frontiers in Plant Science*, 7, 1419. <https://doi.org/10.3389/fpls.2016.01419>
4. Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. (2018). MobileNetV2: Inverted residuals and linear bottlenecks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 4510–4520. <https://doi.org/10.1109/CVPR.2018.00474>
5. Shorten, C., & Khoshgoftaar, T. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data*, 6(1), 60. <https://doi.org/10.1186/s40537-019-0197-0>