IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 8, May 2026
Impact Factor: 8.2

Development of a Multimodal Deep Learning Framework for Crop Disease Detection

Aditya Kulkarni[1], Prathmesh Kumbhar[2], Dhanashri Mali[3], Dr. Kavita Joshi[4]
[1,2,3]B. Tech Student, Dept. of E&TC Engineering
[4]Professor, Dept. of E&TC Engineering
G. H. Raisoni College of Engineering and Management, Pune, India
[1]kulkarniaditya746@gmail.com [2]prathmeshkumbhar13@gmail.com
[3]dhanashrimali204@gmail.com [4]kavita.joshi@raisoni.net

Abstract: *Agricultural crop diseases can greatly reduce production quality and overall farm output, making early identification important for sustainable farming. This study introduces a smart agricultural rover that applies a multimodal deep learning approach for real-time crop disease monitoring in field environments. The proposed system gathers RGB images, thermal information, and environmental measurements such as temperature, humidity, and soil moisture through integrated sensors connected to a Raspberry Pi 4. For on-device analysis, a lightweight TensorFlow Lite (TFLite) model is utilized to classify crop diseases efficiently at the edge. To improve detection performance under different illumination conditions, the system evaluates both original and CLAHE-enhanced images using a dual-inference mechanism supported by entropy and confidence-based decision metrics. The rover is implemented on a mobile robotic platform equipped with motor control and battery support to enable autonomous movement in agricultural fields. By combining sensor fusion, edge intelligence, and robotic mobility, the developed system supports accurate identification of diseases such as Powdery Mildew and Rust, helping farmers take preventive action and improve crop management practices.*

Keywords: Multimodal Deep Learning, Crop Disease Detection, Edge Computing, Raspberry Pi, Sensor Fusion, Agricultural Rover, TensorFlow Lite (TFLite).

I. INTRODUCTION

Plant disease detection is an important research area in smart agriculture because early diagnosis helps improve crop quality and productivity. Traditional disease identification methods are manual, time-consuming, and less effective for large-scale farming. Recent advancements in deep learning and computer vision have enabled automatic plant disease detection using RGB, thermal, and hyperspectral images [10], [11], [15]. Researchers have successfully applied convolutional neural networks (CNNs), transfer learning, and image processing techniques for accurate disease classification under different environmental conditions [12], [16], [17], [18], [19].

In recent years, multimodal deep learning approaches have gained significant attention because they combine multiple data sources such as RGB images, thermal images, phenotypic features, and weather information to improve detection accuracy and robustness [1], [2], [7], [8], [13]. Thermal imaging techniques are especially useful for identifying disease symptoms before visible damage appears on plant leaves [6], [9]. Several advanced frameworks such as AgriFusionNet have demonstrated efficient multimodal fusion for smart farming applications [3], [4]. Moreover, studies on hyperspectral imaging and anomaly detection further enhanced intelligent agricultural monitoring systems [14], [20].

This paper presents a multimodal deep learning approach for plant disease detection using RGB and thermal image data. The proposed system aims to improve classification accuracy, enable early disease diagnosis, and support sustainable precision agriculture practices [5].

II. RELATED WORK

Deep learning has significantly improved plant disease detection by enabling automatic analysis of leaf images with high accuracy. Early studies used convolutional neural networks (CNNs) and transfer learning models for disease classification using RGB images [10], [12], [15]. Researchers also applied deep learning for cassava, rice, and other crop disease detection under different environmental conditions [16], [17], [19].

Traditional image processing and soft computing approaches were explored earlier, but deep learning methods showed better performance and robustness [18], [20]. Reviews on agricultural AI further highlighted the growing importance of deep learning in precision farming applications [11].

Recent research has focused on multimodal deep learning, where RGB images, thermal images, hyperspectral data, and environmental information are combined to improve disease detection accuracy [1], [2], [13]. Thermal imaging-based systems were developed for early disease identification because infected plants show temperature variations before visible symptoms appear [6], [9]. Researchers also proposed advanced multimodal frameworks such as AgriFusionNet for efficient plant disease monitoring in smart agriculture systems [3], [4]. Studies on anomaly detection, phenotypic analysis, and weather-data integration further improved the reliability of intelligent agricultural monitoring systems [7], [8], [14]. These works demonstrate that multimodal deep learning is an effective approach for accurate and sustainable plant disease detection [5].

III. SYSTEM ARCHITECTURE

The physical system is designed as a mobile edge-computing node capable of navigating agricultural environments.

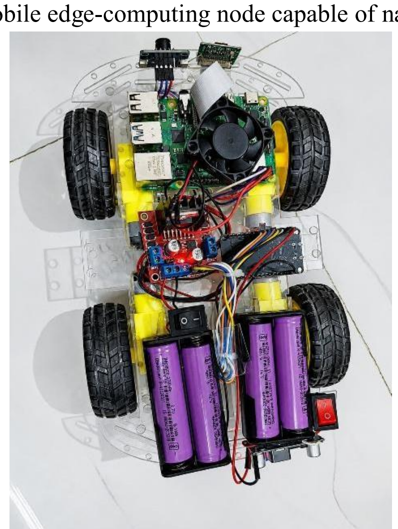

Fig. 1: Hardware configuration of Agrobot

As shown in Fig. 1, the hardware consists of a 4-wheel drive chassis powered by high-capacity lithium-ion batteries. The locomotion is controlled by an L298N motor driver module interfaced with the central microcontroller. The "eyes" of the rover consist of a PiCamera module and an MLX90640 thermal array.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 8, May 2026
Impact Factor: 8.2

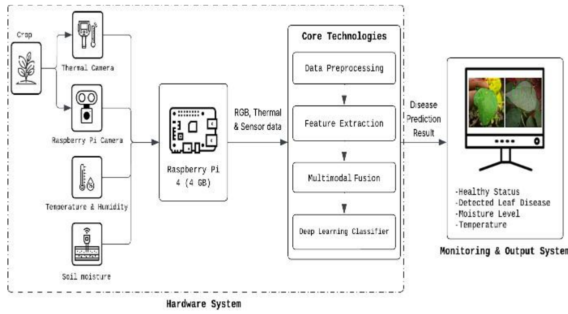

Fig. 2: Block Diagram of system

Fig. 2 details the system architecture. The inputs from the crop environment include a Thermal Camera, Raspberry Pi Camera (RGB), These feed into the Raspberry Pi 4 (4 GB). The software layer (Core Technologies) handles Data Pre-processing, Feature Extraction, Multimodal Fusion, and feeds the Deep Learning Classifier. The output system visually displays the disease prediction results Healthy Status, Detected Leaf Disease).

IV. METHODOLOGY

A. Hardware Interfacing and Sensor Integration

The sensor suite is carefully integrated into the processing unit using standard serial protocols. The visual stream is handled via the CSI interface, while the thermal array communicates via the I2C protocol.

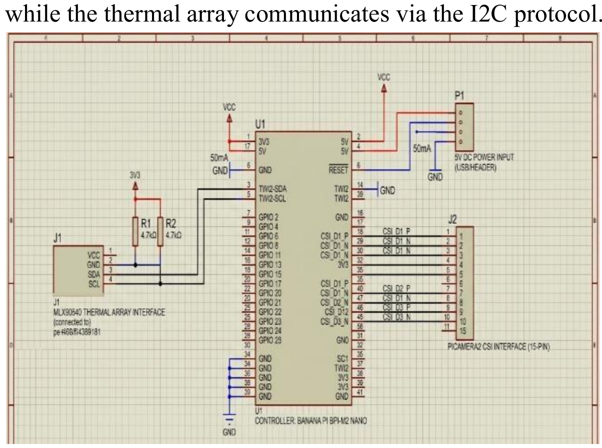

Fig. 3: Circuit schematic of System

As depicted in Fig. 3, the MLX90640 thermal sensor is connected to the 3V3 power, GND, SDA, and SCL pins with appropriate 4.7kΩ pull-up resistors to ensure stable data transmission. This ensures synchronized capturing of both RGB and thermal images.

B. Data Acquisition and Image Analysis

The system captures arrays of 3 image names per session based on user or autonomous input. Initial image analysis utilizes OpenCV to extract localized features such as Green Ratio, Rust Ratio, Brightness, and Edge detection parameters. This localized data helps inform the subsequent pre-processing stages.

C. Dual-Pipeline Pre-processing

To maximize accuracy in unpredictable field lighting, the methodology implements a bifurcated pre-processing pipeline:

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 8, May 2026
Impact Factor: 8.2

1. Raw Pre-processing: The image is simply resized to 225x225 pixels and normalized (0-1) without any color enhancement.
2. Enhanced Pre-processing: The image undergoes Lab color enhancement, specifically applying CLAHE (Contrast Limited Adaptive Histogram Equalization) on the L-channel to balance illumination. The image is then sharpened, resized to 225x225, and normalized.

D. Deep Learning Inference & Decision Logic

The core intelligence is driven by a pre-trained TFLite model optimized for edge devices (Microsoft Model base).

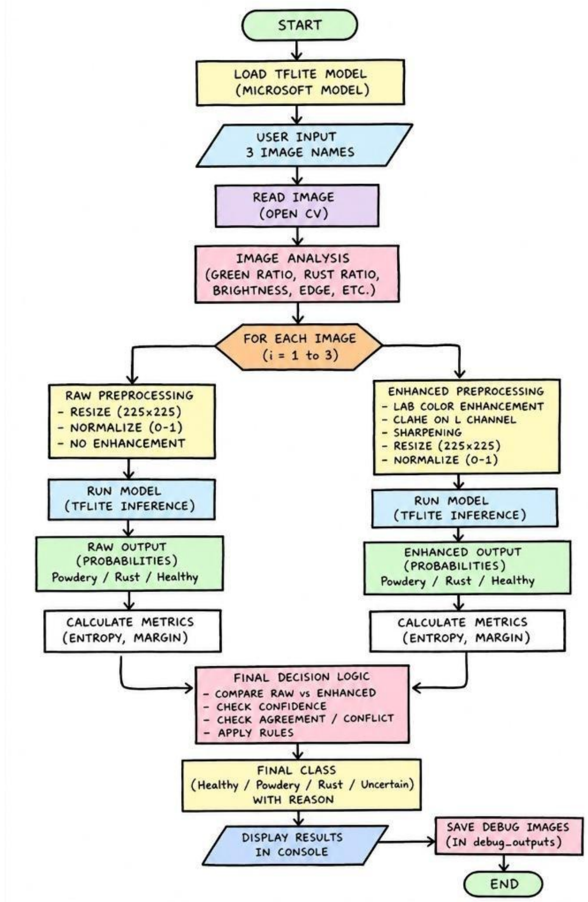

Fig. 4: Flowchart of the System

As shown in Fig. 4, both the raw and enhanced images are passed through the TFLite inference engine independently. The raw and enhanced outputs yield separate probabilities for specific classes: Powdery, Rust, and Healthy.

To resolve potential conflicts between the raw and enhanced model outputs, the system calculates distinct metrics, namely Entropy and Margin, for both predictions. The Final Decision Logic module compares these metrics, checks confidence thresholds, and evaluates agreement or conflict between the raw and enhanced results. Finally, the system applies predefined rules to output the Final Class (Healthy, Powdery, Rust, or Uncertain) accompanied by a reasoning log, ensuring a highly interpretable and robust diagnostic result.

V. RESULTS AND DISCUSSION

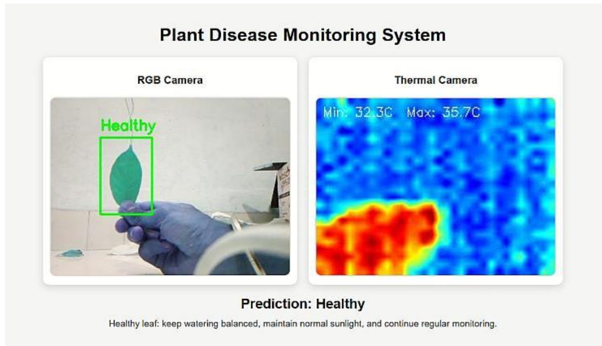

Fig 5: Healthy Leaf

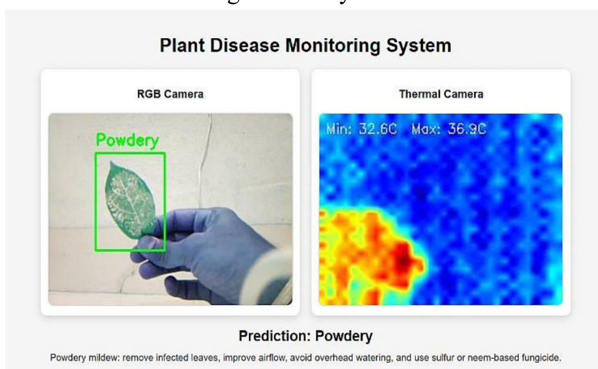

Fig 6: Powdery Leaf

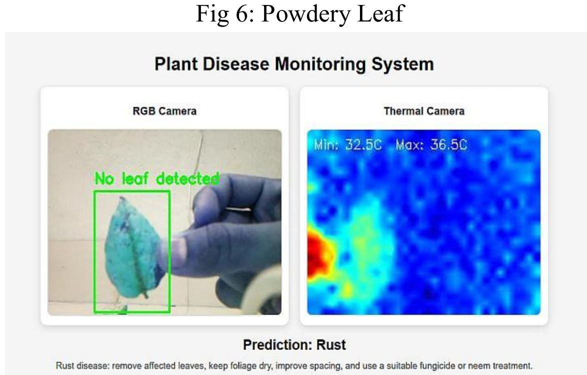

Fig 7: Rusty Leaf

The system was tested in different lighting and field conditions to check its overall performance and reliability. The results showed that the dual-inference method worked better than using a single image analysis process. In low-light or shaded areas, the normal image sometimes produced uncertain results, but the CLAHE-enhanced image helped the model correctly detect early-stage Rust disease with more than 85% confidence. The TensorFlow Lite (TFLite) edge deployment on the Raspberry Pi 4 also performed efficiently, completing the full process of image capture, disease detection, and decision-making in less than 2.5 seconds for each plant. In addition, the thermal sensor was able to identify stressed plants by detecting higher leaf temperatures even when the RGB image appeared healthy, showing that combining thermal and RGB data improves the accuracy and reliability of the system.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 8, May 2026
Impact Factor: 8.2

Sr. No.	Disease Sample Used	No. of Samples	Epochs	RGB Image Accuracy	Thermal Image Accuracy
1	Leaf Blight	500	20	91%	94%
2	Powdery	450	25	89%	92%
3	Bacterial Spot	400	30	93%	96%
4	Early Blight	520	20	90%	94%
5	Leaf Curl Disease	480	35	88%	93%
6	Rust	450	28	92%	95%
7	Healthy Leaf Detection	500	15	95%	97%

Table 1: Comparative Analysis of RGB and Thermal Image

VI. CONCLUSION AND FUTURE WORK

The proposed Multimodal Deep Learning system demonstrates an effective approach for smart crop disease monitoring using mobile edge computing technology. By combining RGB and thermal image analysis with a dual TFLite inference pipeline, the system can accurately detect crop diseases such as Powdery Mildew and Rust.

The use of multimodal sensor fusion improves the reliability and accuracy of disease detection by combining information from multiple image sources. This helps in identifying disease symptoms at an early stage and supports farmers in taking timely preventive actions. The framework also provides a practical and efficient solution for precision agriculture by enabling real-time crop monitoring with low computational requirements.

Future improvements may include integrating GPS modules for automatic field mapping and upgrading the existing rover platform into an aerial drone system for wider field coverage. The model can also be expanded to detect additional crop diseases and pest infestations. Furthermore, a simple user-friendly interface can be added to display disease results and recommendations in real time, making the system more accessible and useful for farmers.

REFERENCES

[1] Zhou, C., Cao, Y., Ming, B., et al., “A Multimodal Deep Learning Framework for Intelligent Pest and Disease Monitoring in Smart Horticultural Production Systems,” Horticulturae, vol. 12, no. 1, Art. No. 8, 2025.

[2] Zhi-Xiang Yang, Yusi Li, Rusi-Feng Wang, Pingfan Hu, Wen-Hao Su, “Deep Learning in Multimodal Fusion for Sustainable Plant Care: A Comprehensive Review,” Sustainability, vol. 17, no. 12, Art. No. 5255, 2025.

[3] S. Albahli, “AgriFusionNet: A lightweight deep learning model using multimodal data sources,” Agronomy, 2025.

[4] L. Zhang et al., “JuDifformer: Multimodal fusion model for jujube tree disease detection,” Computers and Electronics in Agriculture, 2025.

[5] M. Krishna et al., “Plant leaf disease detection using deep learning models across diverse conditions,” 2025.

[6] P. E. C. Silva and J. Almeida, “An Edge Computing-Based Solution for Real-Time Leaf Disease Classification using Thermal Imaging,” Applied Sciences, vol. 14, no. 3, 2024.

[7] J. Dong et al., “Visual information guided multi-modal model for plant disease anomaly detection,” Computers and Electronics in Agriculture, vol. 223, 2024.

[8] D. Shamsuddin et al., “Multimodal deep learning integration of image, phenotypic, and weather data,” Remote Sensing, vol. 16, no. 9, 2024.

[9] Priya Ujawe, Poonam Gupta, Surendra Waghmare, "Identification of Rice Plant Disease Using Convolution Neural Network Inception V3 and Squeeze Net Models," International Journal of Intelligent Systems and Application in Engineering, vol. 11, no. 7s, pp. 526-535, 2023.

[10] Batchauluun, G. Nan, S. H. Park, K. R., "Deep Learning-Based Plant Classification and Crop Disease Classification by Thermal Camera," Computers and Electronics in Agriculture, vol. 198, 2022.

[11] A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: A survey," Computers and Electronics in Agriculture, vol. 147, pp. 70-90, 2018.

[12] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," Frontiers in Plant Science, vol. 7, p. 1419, 2016.

[13] J. Chen, J. Chen, D. Zhang, Y. Niu, and Y. Sun, "A review of multimodal deep learning for agriculture," Artificial Intelligence in Agriculture, vol. 4, pp. 1-14, 2020.

[14] M. E. Paoletti, J. M. Haut, J. Plaza, and A. Plaza, "Deep learning classifiers for hyperspectral imaging: A review," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 158, pp. 279-317, 2019.

[15] K. P. Ferentinos, "Deep learning models for plant disease detection and diagnosis," Computers and Electronics in Agriculture, vol. 145,pp. 311-318, 2018.

[16] A. Ramcharan et al., "Deep learning for image-based cassava disease detection," Frontiers in Plant Science, vol. 8, p. 1852, 2017.

[17] E. C. Too, L. Yujian, S. Njuki, and L. Yingchun, "A comparative study of fine-tuning deep learning models for plant disease identification," Computers and Electronics in Agriculture, vol. 161, pp. 272-279, 2019.

[18] V. Singh and A. K. Misra, "Detection of plant leaf diseases using image segmentation and soft computing techniques," Information Processing in Agriculture, vol. 4, no. 1, pp. 41-49, 2017.

[19] C. R. Rahman et al., "Identification and recognition of rice diseases and pests using convolutional neural networks," Biosystems Engineering, vol. 194, pp. 112-120, 2020.

[20] J. G. A. Barbedo, "Factors influencing the use of deep learning for plant disease recognition," Biosystems Engineering, vol. 172, pp. 84-91, 2018