International Journal of Science and Research (IJSR)
ISSN: 2319-7064
Impact Factor 2025: 7.089

Potato Leaf Disease Detection: Classical ML with Future Quantum and RL Directions

Roshini N

CMR University, School of Science and Computer Studies, India

Abstract: ***Plant disease detection is essential for improving crop productivity and agricultural sustainability. Machine learning enables automated classification of leaf diseases, reducing reliance on manual visual inspection. This study experimentally evaluates K-Nearest Neighbors (KNN) and Random Forest (RF) for classifying potato leaf diseases using a balanced subset of the PlantVillage dataset. Images were preprocessed through resizing, normalization, and flattening prior to classification. Random Forest outperformed KNN across all metrics, achieving 88.00% accuracy versus 75.33% for KNN. Beyond the experimental study, this paper reviews recent advances in Quantum Machine Learning (QML) for agricultural disease detection and proposes a Reinforcement Learning (RL)-based framework as a conceptual future extension for adaptive disease diagnosis. The RL framework is not experimentally implemented in this work. The results establish reliable classical baselines while highlighting future directions for QML and RL in precision agriculture.***

Keywords: Plant Disease Recognition, Precision Agriculture, Random Forest, K-Nearest Neighbors, Quantum Machine Learning, Reinforcement Learning, PlantVillage Dataset

1. Introduction

Agriculture remains vital to food security and the economies of many developing nations, yet crops are routinely threatened by diseases caused by fungi, bacteria, viruses, and other pathogens, resulting in reduced yield and financial loss for farmers. Conventional disease diagnosis relies on visual inspection by experts, which is labour-intensive, time-consuming, and prone to subjective error- particularly when diseases share similar visual symptoms. Machine learning and deep learning techniques have therefore been increasingly adopted to automatically analyse leaf images and classify disease categories, improving diagnostic reliability while reducing dependence on expert availability.

K-Nearest Neighbors (KNN) and Random Forest (RF) are widely used classical algorithms due to their simplicity, efficiency, and low computational cost. RF benefits from ensemble learning, improving robustness and reducing overfitting, while KNN offers a straightforward distance-based baseline. However, classical methods can struggle when disease classes share highly similar visual characteristics. Quantum Machine Learning (QML) has emerged as a promising paradigm combining quantum computing with machine learning to improve computational efficiency for such complex classification tasks, though practical adoption remains limited by the early stage of quantum hardware.

Most existing literature evaluates either classical or quantum-enhanced approaches independently, with few studies establishing strong classical baselines while also discussing how emerging quantum techniques might enhance future diagnosis. Motivated by this gap, the present study experimentally evaluates KNN and RF for potato leaf disease classification, reviews recent QML developments relevant to agriculture, and proposes a conceptual RL-based framework for adaptive disease monitoring as a future research direction.

2. Literature Review

Recent studies have explored hybrid quantum-classical frameworks for crop disease prediction. One line of work combined quantum-enhanced techniques with CNN and IoT-based sensor data for early tomato disease prediction, while another integrated MobileNetV2 with quantum-inspired feature extraction to detect nematode-affected leaves alongside soil and environmental data. A comparative study on sugar beet Cercospora leaf disease benchmarked classical algorithms against Quantum Support Vector Classifiers and Variational Quantum Classifiers, finding that classical methods achieved slightly higher accuracy while quantum models showed encouraging potential. Further work has examined QML for broader crop-health monitoring using imaging, spectroscopy, and IoT sensors, and has shown that Variational Quantum Circuits combined with conventional classifiers can outperform standard deep learning architectures on noisy, highdimensional data.

Across this literature, most studies evaluate classical or quantum approaches in isolation rather than establishing dual classical baselines alongside a forward-looking discussion of quantum and reinforcementbased extensions. The present study addresses this gap by experimentally benchmarking two classical algorithms while contextualising QML and proposing RL as future directions.

3. Methodology

The methodology establishes an experimental framework for potato leaf disease classification comprising dataset preparation, image preprocessing, feature extraction, model training, and performance evaluation. The workflow (Fig. 1) contrasts the classical processing pipeline used in this study with the additional statepreparation and measurement stages required by a quantum pipeline, and produces classical baselines against which literature-reported QML and proposed RL directions are discussed.

International Journal of Science and Research (IJSR)
ISSN: 2319-7064
Impact Factor 2025: 7.089

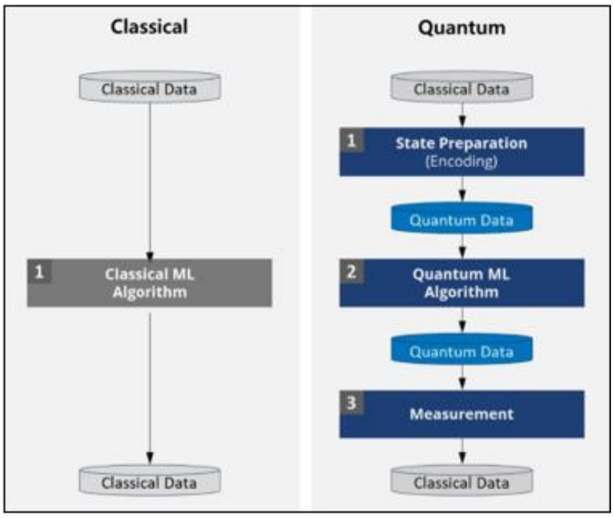

Figure 1: Classical machine learning pipeline compared with a quantum machine learning pipeline (state preparation, quantum processing, measurement).

To ensure unbiased evaluation, a perfectly balanced class distribution was maintained throughout. Raw images were standardised via three preprocessing steps: (i) dimensional modification- resizing to a uniform 64×64 resolution to limit computational complexity; (ii) intensity normalization- scaling pixel values to a continuous [0, 1] range to stabilise optimisation; and (iii) vector transformation- flattening each image into a 4,096-element one-dimensional feature vector.

Two supervised classifiers were benchmarked. KNN uses a distance-based approach in high-dimensional feature space and serves as a transparent, non-parametric baseline. Random Forest aggregates predictions from multiple independent decision trees, producing robust decision boundaries and reduced variance across the high-dimensional pixel space, directly addressing overfitting that distance-based methods are prone to. Model robustness was measured using accuracy, precision, recall, and F1-score, all computed with weighted averaging to account for class-specific performance differences- particularly the visual overlap between Early Blight and Late Blight.

4. Experimental Setup

Experiments were implemented in Python using Scikit-learn, NumPy, and Pandas, executed entirely on standard CPU infrastructure; QML and RL were not physically deployed and are discussed only via literature. A balanced subset of 1,500 PlantVillage potato leaf images (500 each of Early Blight, Late Blight, and Healthy) was partitioned with a fixed random seed into training (900 images), validation (300 images), and an independent test set (300 images), each maintaining equal class representation. After preprocessing (resizing, normalization, flattening to 4,096 features), KNN was configured with k=5 and Euclidean distance, and Random Forest was built as a 100-tree ensemble using Gini impurity for split selection. Final metrics were computed on the independent 300-image test set using weighted averaging.

5. Results and Discussion

Table I summarises overall classification performance on the 300-image test set.

Table I: Experimental Classification Performance on Potato Leaf Disease Dataset.

Algorithm	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
KNN (k=5)	75.33	80.12	75.33	75.54
Random Forest	88.00	88.09	88.00	87.97

Random Forest achieved the strongest performance overall- 88.00% accuracy, 88.09% precision, 88.00% recall, and 87.97% F1-score- outperforming KNN's 75.33% accuracy by 12.67 percentage points. This gap reflects Random Forest's ensemble aggregation, which reduces variance across the high-dimensional pixel space and generalises better than KNN's purely distance-based decision rule.

Table II breaks down performance per disease class.

Table II: Per-Class Performance- KNN vs. Random Forest

Algorithm	Class	Precision	Recall	F1-Score
KNN (k=5)	Early Blight	0.97	0.65	0.78
KNN (k=5)	Healthy	0.81	0.69	0.75
KNN (k=5)	Late Blight	0.62	0.92	0.74
Random Forest	Early Blight	0.91	0.95	0.93
Random Forest	Healthy	0.90	0.83	0.86
Random Forest	Late Blight	0.83	0.86	0.84

KNN showed a strong precision/recall imbalance: Early Blight had high precision (0.97) but low recall (0.65), indicating frequent misclassification into other classes, while Late Blight showed the inverse pattern (0.62 precision, 0.92 recall), indicating over-prediction of that class. This stems from the visual similarity between Early and Late Blight lesions, which the distance-based KNN classifier struggles to separate. Random Forest showed a far more balanced trade-off across all three classes, with F1-scores of 0.93 (Early Blight), 0.86 (Healthy), and 0.84 (Late Blight), as visualised in Fig. 3.

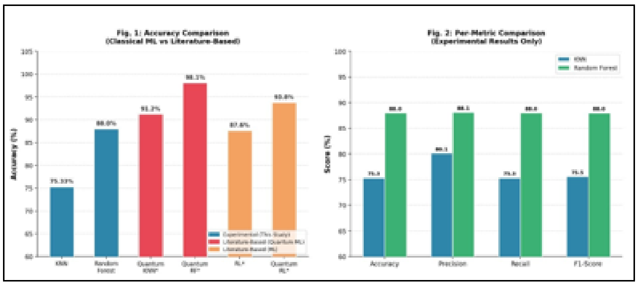

Figure 3: Per-class F1-score comparison between KNN and Random Forest across Early Blight, Healthy, and Late Blight.

Fig. 2 places these experimental results alongside literature-reported ranges for Quantum ML and the proposed RL direction. Since the underlying datasets and computational environments differ from those used here, these quantum and RL figures are shown only for qualitative context and are not a direct experimental comparison.

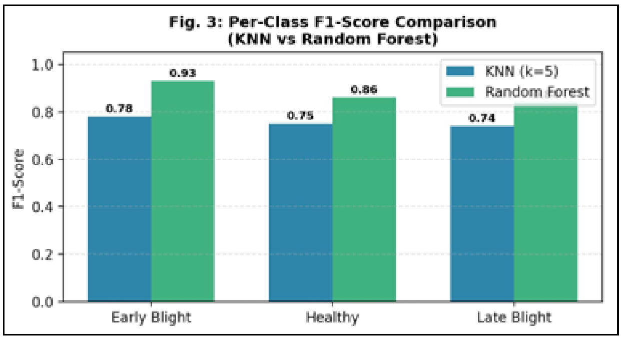

Figure 2: Experimental accuracy of KNN and Random Forest compared with literature-reported ranges for Quantum ML and RL approaches.

6. Conclusion

This study established reliable classical machine learning baselines for automated potato leaf disease classification using a balanced PlantVillage subset. Random Forest clearly outperformed KNN, achieving 88.00% accuracy, 88.09% precision, 88.00% recall, and an 87.97% F1-score, compared to KNN's 75.33% accuracy — with per-class analysis confirming KNN's particular sensitivity to the visual overlap between Early and Late Blight. Alongside these results, the study reviewed recent QML developments relevant to agricultural image analysis and proposed a conceptual RL-based framework for adaptive disease monitoring; neither QML nor RL was experimentally implemented here. Future work may implement Quantum KNN or quantum-enhanced ensembles on suitable simulators, and explore the proposed RL framework for adaptive, intelligent crop disease diagnosis, building on the classical baselines established in this work.

References

[1] N. Jhenkar, T. D. Mallikarjun, V. Madhuri, M. Kishan, and J. R. Devaru, "Quantum machine learning for crop disease prediction and forecasting," Dept. Electron. Commun. Eng., Dayananda Sagar Univ., Bengaluru, India, Tech. Rep., 2024.

[2] R. Borate, M. Kulkarni, A. Patil, A. Nikam, and S. Shaikh, "Fruit crop disease classification using quantum machine learning: A pilot study," Quantum J. Eng. Sci. Technol., vol. 6, no. 4, pp. 35-44, Dec. 2025.

[3] R. Katirci, K. Adem, M. Tatar, and F. Ölmez, "Comparison of the performance of classical and quantum machine-learning methods on the detection of sugar beet Cercospora leaf disease," Plant Pathol., vol. 74, no. 2, pp. 507-518, Feb. 2025.

[4] Y. Wu, A. Nagy, Z. Rajnai, B. Fregan, and K. Takács-György, "Quantum Machine Learning for Crop Disease Detection and Prevention," in Proc. 2025 IEEE 12th Int. Conf. on Computational Cybernetics and Cyber-Medical Systems (ICCC), Budapest, Hungary, Apr. 2025.

[5] S. Lakhani, "Revolutionizing Smart Farming: Quantum Computing Applications in Plant Disease Detection," Int. J. Computational Processing and Data Mining (IJCPDM), 2025.

[6] I. Attri, L. K. Awasthi, and T. P. Sharma, "Entangled Quantum Image Descriptor: An Approach for Early Plant Disease Detection," Crop Protection, vol. 188, Art. no. 107005, Feb. 2025.

[7] U. Naresh and T. B. Reddy, "A Review on Quantum-Enhanced Deep Learning Frameworks for Reliable Plant Disease Detection," Current Agriculture Research Journal, vol. 13, no. 3, 2025.

[8] H. N. Ngugi, A. A. Akinyelu, and A. E. Ezugwu, "Machine Learning and Deep Learning for Crop Disease Diagnosis: Performance Analysis and Review," 2024.

[9] A. Mukherjee and B. B. Mallik, "Transforming Precision Agriculture with Quantum Computing," EPJ Web of Conferences, vol. 325, Art. no. 01004, 2025.

[10] P. R. Naidu, D. Gowda V., D. Pathak, P. Meshram, S. D., and S. Devi, "Optimized Crop Disease Detection with Image Processing and Machine Learning," in Proc. 2025 3rd World Conf. on Communication & Computing (WCONF), 2025.

[11] J. Basit, H. Arshad, and A. Bibi, "Optimizing Crop Yield Forecasts Using Quantum Machine Learning Techniques," J. Computers and Intelligent Systems, vol. 3, no. 1, pp. 1-15, 2025.

[12] A. Nagy, Y. Wu, K. Takács-György, Z. Rajnai, and B. Fregan, "Glance at Quantum Innovations in Crop Pest and Disease Detection," in Proc. 2025 IEEE 23rd World Symposium (SAMI), 2025.

[13] S. Kumar, S. S. Singh, G. Bathla, S. Sharma, and M. Panjeta, "Fusion of Quantum Computing with Smart Agriculture: A Systematic Review," 2026.

[14] A. K. Gupta, A. Bharadwaj, and M. A. Haque, "Towards Accurate Crop Yield Forecasting with Quantum Machine Learning Models," 3rd Int. Online Conf. on Agriculture, MDPI, Oct. 2025.

[15] D. R. I. M. Setiadi, A. Susanto, K. Nugroho, A. R. Muslikh, A. A. Ojugo, and H.-S. Gan, "Rice Yield Forecasting Using Hybrid Quantum Deep Learning Model," 2024.