Interpreting potato disease classification using explainable artificial intelligence

Rakesh Kumar Gumasta, Ajay Somkuwar
Department of Electronics and Communication Engineering, Maulana Azad National Institute of Technology, Bhopal, India

Article Info

Article history:

Received Aug 11, 2025
Revised Mar 18, 2026
Accepted Aug 13, 2026

Keywords:

Explainable artificial intelligence
Fidelity
Model interpretability
Plant disease classification
Robustness

ABSTRACT

Timely and accurate crop disease detection is most important for ensuring global food security. For detecting diseases in crops, many different machine learning (ML) models were proposed. These models work as a black-box, and without proper explanation of these models' decisions, farmers may find it difficult to trust these systems. For this, many model explainability methods were also proposed. All these methods have been evaluated qualitatively, but quantitative and cross-method comparisons are lacking. This study addresses this gap by emphasizing quantitative validation of models' explanations. This study investigates the interpretability and performance of two approaches for potato leaf disease classification: i) manual feature engineering with relief-based feature selection and artificial neural network (ANN) classification with local interpretable model-agnostic explanations (LIME) and ii) deep convolutional neural networks (CNNs) based on mobile network version 2 (MobileNetV2) with gradient-weighted class activation mapping (Grad-CAM). Quantitative assessment of explanation quality was performed using fidelity and robustness metrics. While LIME achieved a lower average fidelity drop 10.90% and higher robustness 84.27% compared to Grad-CAM 18.68% fidelity drop and 79.38% robustness, Grad-CAM provided clearer visual explanations. These findings suggest that explanation effectiveness is influenced by model design and data characteristics, highlighting the need for careful selection of interpretability techniques in precision agriculture applications.

Corresponding Author:

Rakesh Kumar Gumasta
Department of Electronics and Communication Engineering, Maulana Azad National Institute of Technology
Bhopal 462003, India
Email: rakeshgumasta44@gmail.com

1. INTRODUCTION

Early detection of crop diseases is essential for timely intervention and reducing yield losses. Manual monitoring of large agricultural areas is difficult; therefore, intelligent and automated disease detection systems have gained increasing attention. Machine learning (ML) techniques have been widely investigated for plant disease classification. For identifying relevant features, relief-based feature selection has been explored for plant leaf disease classification [1]–[3]. Different ML classifiers have also been applied to distinguish healthy and diseased plants. Artificial neural networks (ANNs) and support vector machines (SVMs) have demonstrated their effectiveness for plant disease classification [4], [5]. Deep learning, particularly convolutional neural networks (CNNs), has been widely adopted for automatically extracting discriminative patterns from plant images [6], while lightweight architectures and transfer learning support efficient disease identification [7]. Data augmentation further improves model robustness and generalization under varying field conditions [8]. Although deeper and more complex deep learning models can improve predictive performance [9], [10], their black-box nature makes the reasoning behind predictions difficult to understand, potentially limiting trust and practical adoption.

Explainable artificial intelligence (XAI) methods address this limitation by identifying features or image regions influencing model decisions. Local interpretable model-agnostic explanations (LIME) provides local model-agnostic explanations [11], while gradient-weighted class activation mapping (Grad-CAM) highlights class-discriminative image regions in CNNs [12]. Both methods have been applied in plant disease classification and related agricultural applications [13]–[17]. However, existing studies have predominantly evaluated explanations qualitatively, with limited quantitative comparison of different model–explanation combinations [18]–[20]. Although advanced explanation techniques have also been explored [21], [22], the reliability of explanations from lightweight models such as mobile network version 2 (MobileNetV2) remains insufficiently quantified [23]. Simpler feature-based ANNs are seldom benchmarked against deep models. Recent studies highlight the need for quantitative validation of model explanations [24], [25], a direction that remains underexplored in plant disease research. Collectively, existing research lacks a systematic, quantitative framework to assess explanation fidelity and robustness across model types. Moreover, the integration of interpretable and lightweight models for mobile or field-level plant disease diagnosis remains underdeveloped, revealing a critical research gap in ensuring the reliability and comparability of explanation methods.

To address these gaps, this study provides a comparative analysis of two combinations of models and interpretation methods. In the first approach, this study combined manual feature engineering with a less complicated ANN model and then applied the LIME interpretation method to find which particular feature is most important in making a decision for a particular leaf image. In the second approach, this study used the MobileNetV2 architecture and applied the Grad-CAM interpretation method to directly visualize the portion of the image responsible for a particular decision. Extensive image augmentation was used to improve generalization to field conditions. This comparison evaluates both predictive performance and interpretability fidelity, providing insights into which XAI approaches best align with domain knowledge and supporting more transparent, reliable AI for precision agriculture.

The objective is to provide a framework for quantitative analysis of LIME and Grad-CAM explanation methods for the agricultural domain. The architecture supports practical mobile deployment through low inference cost and efficient processing. The relief-based ANN enables rapid on-device computation, while CNN models provide heatmap-based visual explanations rendered on smartphone screens. Farmers can capture leaf images via mobile cameras to obtain instant disease predictions and visualize affected regions, even offline. It enables farmers to interpret model reasoning and obtain actionable insights directly in the field. This transparency enhances user trust, supports localized disease diagnosis, and promotes wider adoption within smart agriculture ecosystems. The main contributions of this work include: i) development of two complementary explainable models: a relief-based ANN interpreted via LIME, and a MobileNetV2 model analyzed using Grad-CAM and ii) quantitative assessment of explanation quality based on fidelity and robustness metrics.

The remainder of this paper is organized as follows. Section 2 describes the proposed methodology. Section 3 presents the results and discussion. Finally, section 4 concludes the study with insights on future research directions.

2. METHOD

2.1. Dataset description

To address existing gaps in XAI for precision agriculture, this study presents a comparative analysis of two integrated frameworks that combine model architectures with interpretation methodologies. The comparison aims to evaluate the relationship between predictive performance and interpretability fidelity. Potato images from PlantVillage dataset were used. The selected dataset contains a total of 2,152 images belonging to three classes. The first class consists of 152 healthy leaves, while the second and third classes consist of 1,000 images of early blight and 1,000 images of late blight, respectively. The original dataset contains images with a resolution of 256 × 256 pixels, which were resized to 224 × 224 pixels. In this study, 80% of the images were used for training, while the remaining 20% were used for testing.

2.2. Data augmentation

This study employed image augmentation techniques to improve the generalization ability and robustness of the CNN model in the field [26]. During image augmentation, the dataset is synthetically expanded by applying random transformations. For this purpose, a rotation range of ±30 degrees, random horizontal flipping, and brightness adjustment in the range of 0.7 to 1.3, a zoom variation of up to 20%, and Gaussian noise with a small variance were applied. These augmentations replicate natural variability in field conditions such as changes in lighting, camera angles, and leaf appearance. Gaussian noise simulates sensor imperfections and minor environmental artifacts, encouraging the model to learn more robust and noise-tolerant representations. The augmented dataset was subsequently used to train the MobileNetV2-based CNN model.

2.3. Approach A: manual feature engineering + relief + artificial neural network

The first approach consists of four stages. The first stage is feature extraction, the second is feature selection, the third is training and testing of the classification model, and the fourth is model interpretation. In feature extraction stage, this study extracted a comprehensive set of handcrafted features from each leaf image to capture color, texture, and structural information. Color characteristics are quantified using histograms of the red, green, and blue channels, each computed with 32 bins to capture the distribution of intensities. Texture features, including contrast, homogeneity, dissimilarity, angular second moment (ASM), and correlation, are extracted, which describe spatial relationships of pixel intensities. In statistical descriptors, the mean and standard deviation of each color channel are calculated to summarize color variations. The entropy of the grayscale image is estimated to measure the complexity of intensity patterns. Histogram of oriented gradients (HOG) features are extracted to characterize shape and local gradients, using eight orientations and 32 × 32-pixel cells, providing a compact representation of edge information. These diverse features are concatenated into a single vector that serves as input for subsequent ML classifiers. In this way, a total of 501 features were extracted from each image. The names of these original features are stored in a list called "feature_names". In feature selection stage, the relief algorithm was applied to identify the 128 most relevant features for classification. Here, the purpose of using Relief is to identify features that globally help distinguish between classes and improve model classification accuracy. It is also used to reduce the dimensionality or complexity of the model. The indices of these top-ranked features, based on their importance scores, are stored in the array "selected_indices". These indices indicate the positions of the selected features in the original feature space.

For the classification model, an ANN is designed using a sequential architecture comprising four hidden layers. The hidden layers included 256, 128, 64, and 32 neurons each with rectified linear unit (ReLU) activation, followed by batch normalization and a dropout layer with different rate to reduce overfitting. The output layer employed a softmax activation function with a number of neurons equal to the number of target classes to produce class probabilities. A batch size of 32 is selected, and the Adam optimizer is used to train the ANN model with a learning rate of 0.001. To enhance interpretability, this study used LIME framework in tabular mode to interpret the ANN model predictions. LIME highlights the top 20 most influential features in the model's decision. However, LIME provides the explanation in terms of column indices (lime_column_index) corresponding to the reduced feature set of relief. To find out which original features these indices refer to, each lime_column_index is remapped to its original feature_index using selected_indices as shown in (1). The actual feature name is then retrieved using this index from the feature_names list as shown in (2).

$$feature\_Index = selected\ indices\ [lime\_column\_index] \tag{1}$$

$$feature\_name = feature\_names\ [feature\_Index] \tag{2}$$

This mapping ensures that the explanations produced by LIME are interpretable in terms of the original image-derived features, even though the model operates on a reduced feature subset after relief selection.

2.4. Approach B: convolutional neural network + gradient-weighted class activation mapping

In the second approach, two stages are used: the first is CNN model training, followed by the Grad-CAM interpretation method. Training deep learning models generally requires a large amount of training data. Since our dataset does not have large number of training images, a transfer learning approach is employed, leveraging the MobileNetV2 architecture as the base model. The MobileNetV2 model was initialized with weights learned on ImageNet and configured with include_top = False to remove its original classification head. During the first stage of training, only the top layers were retrained, while the pretrained MobileNetV2 layers remained frozen. This step enabled the model to adapt its high-level representations to the target dataset without disrupting the pretrained convolutional filters. The model was compiled using the Adam optimizer and the sparse categorical cross-entropy loss function, and trained the model for eight epochs using both training and validation sets. For fine-tuning, the base model MobileNetV2 was unfrozen by setting base_model.Trainable = true, enabling all convolutional layers to be updated during training. The model was then retrained for an additional seven epochs, starting from the last completed epoch of the initial training. This process allows the feature extractor to be refined jointly with the classifier, leading to improved discriminative performance. Overall, this two-stage training strategy, consisting of initial transfer learning with frozen weights, followed by end-to-end fine-tuning, balances the benefits of leveraging prior knowledge from large-scale datasets with the ability to specialize the model to the characteristics of the target domain. Grad-CAM is applied to visualize the portion of the image responsible for a particular decision.

2.5. Evaluation metrics

Both ANN and MobileNetV2 models were evaluated using accuracy, precision, recall, and F1-score as defined in (3) to (6).

$$Accuracy = \frac{TP + TN}{TP+TN+FP+FN} \quad (3)$$

$$Recall = \frac{TP}{TP+FN} \quad (4)$$

$$Precision = \frac{TP}{TP+FP} \quad (5)$$

$$F1 - score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (6)$$

Where TP is true positives, FP is false positives, TN is true negatives, and FN is false negatives.

Accuracy reflects overall classification performance and measures how well the model distinguishes between healthy and diseased leaves or identifies specific diseases. Precision measures the correctness of positive predictions, recall measures the detection of diseased samples, and F1-score balances both [27]. Explanation methods were systematically evaluated using fidelity and robustness. Fidelity measures how well an explanation reflects the model's decision-making process, while the fidelity drop quantifies the decrease in model confidence or accuracy when important features identified by the explanation are removed. Robustness assesses the consistency and reliability of explanations under varying input conditions. In agricultural disease detection, robust methods should maintain meaningful explanations even for noisy or degraded images, helping users identify disease-affected regions.

Hyperparameter settings for calculating fidelity and robustness of both methods are as follows. For LIME (on ANN model): number of perturbed samples = 5,000, Kernel width = 0.25, number of top features removed = 5, distance metric = Euclidean, perturbation type = Gaussian noise with $\sigma = 0.02$. For Grad-CAM (on CNN model): target layer = last convolutional block of MobileNetV2, gradient normalization = true, heatmap resolution = 224 × 224 pixels, and perturbation step for the robustness test = 5% pixel noise.

3. RESULTS AND DISCUSSION

The performance and interpretability of the proposed potato leaf disease classification framework were evaluated using both traditional ANN and CNN architectures. The CNN model clearly outperformed the ANN across all evaluation measures. Specifically, the CNN model based on MobileNetV2 with fine-tuning on the target dataset achieved a test accuracy of 99.76%, along with a precision of 99.82%, recall of 99.82%, and an F1-score of 99.82%. In comparison, the ANN model reached a lower test accuracy of 99.07%, with corresponding precision, recall, and F1-score values of 99.08%, 99.07%, and 99.07%, respectively. These results clearly demonstrate that the CNN's ability to learn hierarchical feature representations directly from images resulted in higher discriminative performance and better overall generalization. While the ANN combined with relief-selected features offered reasonably good accuracy, its reliance on manually engineered inputs limited its effectiveness relative to the deep learning approach.

For interpretability, LIME was applied to the ANN model, while Grad-CAM was applied to the CNN model. LIME provided feature-level explanations by highlighting the top 20 contributing handcrafted features responsible for each prediction. In an example instance, LIME identified features like CH_B_bin_22, CH_B_bin_21, CH_R_bin_32, and HOG_59 as most influential in correctly predicting the class "early blight". To quantitatively assess the quality of these explanations, fidelity and robustness scores were computed. The average fidelity drop for LIME on 30 test images was 10.90%, indicating that removing the top 5 features from the input reduced the model's prediction confidence, thus validating their importance. The robustness of LIME explanations, computed based on the consistency of top 5 features across similar images, was 84.27%, reflecting stable explanations.

Figure 1 shows the interpretability comparison of CNN and ANN models using Grad-CAM and LIME explanations for potato leaf disease classification. Figure 1(a) shows the original test images with true class labels (early blight, late blight, and healthy). Figure 1(b) presents Grad-CAM visualizations highlighting spatial regions that most influenced the CNN's predictions. Figure 1(c) depicts LIME feature contribution plots for the ANN model, showing the relative importance and polarity of features driving each classification decision. Together, these results illustrate complementary insights: Grad-CAM provides region-level visual focus, while LIME reveals feature-level attribution contributing to disease class identification. Grad-CAM provides spatial explanations directly on the input image by generating heatmaps that localized the most discriminative regions. The highlighted area (Figure 1) corresponds closely to the disease-affected region of the leaf, affirming the relevance of the model's attention. The average fidelity drop for Grad-CAM was observed to be 18.68%, which, although higher than LIME's, reflects the model's heavy reliance on spatial patterns. Grad-CAM robustness was measured at 79.38%, indicating generally consistent attention across similar inputs but slightly lower stability compared to LIME.

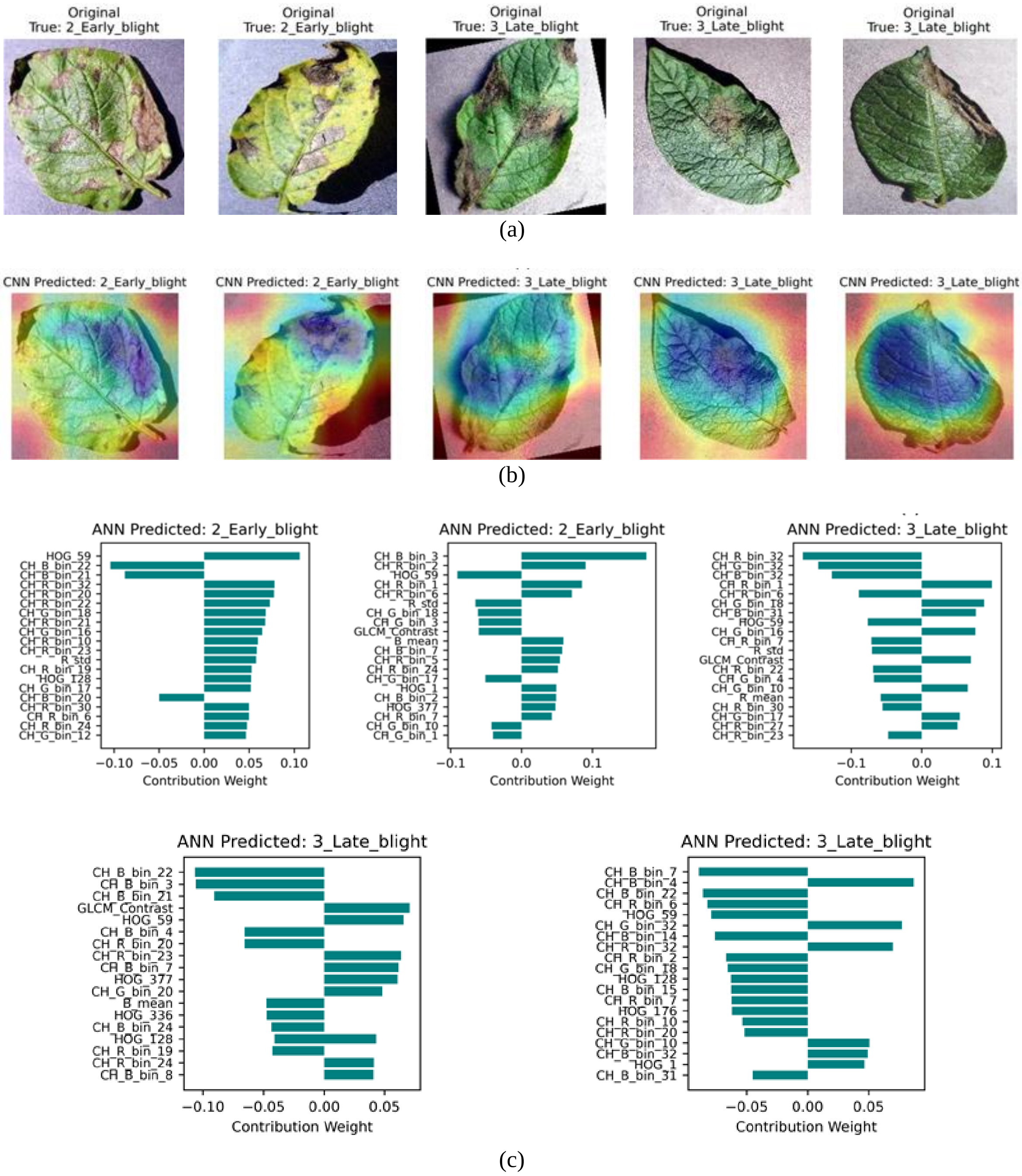

Figure 1. Interpretability comparison of CNN and ANN models using Grad-CAM and LIME explanations for potato leaf disease classification: (a) original test image, (b) Grad-CAM interpretation, and (c) LIME explanation

The ANN combined with LIME produced explanations in the form of numerical feature importances, providing a ranked list of the most influential features for each prediction. It explicitly quantifies the contribution of each input feature. In contrast, the CNN with Grad-CAM generates visual heatmaps highlighting spatial regions within the input images that most strongly influence the model's decision. It helps us to visually inspect the diseased areas of the leaf. Together, these methods provide a complementary understanding of model decision-making, from interpretable numerical features to intuitive visual evidence, ensuring transparency and trustworthiness in automated plant disease detection systems.

The results of this study demonstrate that both approaches achieved high classification performance, with the computationally efficient MobileNetV2 CNN model reaching an overall accuracy of 99.76%, confirming the effectiveness of deep convolutional architectures for plant disease detection [17], [26]. This is consistent with previous reports that fine-tuned CNNs can achieve near-perfect accuracy under controlled conditions [22], [23]. The study integrates a comprehensive augmentation pipeline including flips, rotations, brightness and contrast adjustments, and Gaussian noise. This multi-faceted strategy improves robustness to variable field imaging conditions. Earlier agricultural studies focused solely on classification accuracy without interpretability assessments [5], [7], [8]. In contrast, the approach produces class-discriminative heatmaps.

However, interpretability metrics revealed nuanced differences between methods. LIME explanations applied to the relief-selected ANN model achieved a lower average fidelity drop of 10.90% and higher robustness of 84.27%, while Grad-CAM explanations for the CNN model exhibited a higher fidelity drop of 18.68% and lower robustness of 79.38%. These findings partially contrast with earlier studies that have reported greater stability and fidelity for gradient-based visual explanations in plant disease classification. In the experiments, the simpler feature representations derived from manual feature engineering combined with relief selection may have contributed to more robust local explanations when perturbed by data augmentations such as rotation, brightness variation, and Gaussian noise.

Moreover, while Grad-CAM provided visually intuitive heatmaps highlighting disease regions, its quantitative robustness was somewhat lower, suggesting that explanation stability can depend strongly on model complexity and the homogeneity of image backgrounds. This divergence highlights the importance of evaluating interpretability methods not only qualitatively but also using objective metrics tailored to agricultural datasets. Overall, the results indicate that while CNNs offer superior classification performance, simpler models with feature selection can yield explanations with higher fidelity and stability under certain conditions, supporting the case for hybrid approaches that balance accuracy and interpretability in precision agriculture applications.

4. CONCLUSION

This study compared two approaches for plant disease classification and interpretability: an ANN leveraging manually engineered features selected by relief and interpreted using LIME, and a MobileNetV2-based CNN with Grad-CAM visual explanations. The MobileNetV2 model achieved a high classification accuracy of 99.76%, confirming the effectiveness of deep learning in plant disease detection. However, LIME explanations provided a lower fidelity drop (10.90%) and higher robustness (84.27%) compared to Grad-CAM (18.68% fidelity drop and 79.38% robustness), indicating that simpler models with explicit feature selection can yield more stable and faithful explanations under data augmentations simulating field conditions. These findings highlight the importance of evaluating interpretability methods quantitatively and contextually to ensure trustworthy deployment in precision agriculture. For future work, we plan to extend this comparative analysis to more complex and diverse datasets with real-world field images containing background variability and occlusions. Incorporating additional interpretability techniques, such as SHAP or integrated gradients, could provide further insights into the strengths and limitations of different explanation methods. Collaboration with agricultural experts can further tailor interpretability outputs to support informed decision-making in the field. Moreover, exploring hybrid models that integrate handcrafted and learned features may help balance interpretability with predictive performance.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Rakesh Kumar Gumasta	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓			
Ajay Somkuwar	✓	✓		✓		✓				✓		✓		

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

The authors declare that there are no known conflicts of interest associated with this publication. There are no financial or personal relationships that could inappropriately influence or bias the content of this work.

DATA AVAILABILITY

Data availability is not applicable to this paper as no new data were created or analyzed in this study.

REFERENCES

[1] A. Roman, M. M. Rahman, S. A. Haider, T. Akram, and S. R. Naqvi, "Integrating feature selection and deep learning: a hybrid approach for smart agriculture applications," *Algorithms*, vol. 18, no. 4, Apr. 2025, doi: 10.3390/a18040222.
[2] K. Syed, S. S. A. Begum, A. R. Palakayala, G. V. V. Lakshmi, and S. Gorikapudi, "Leaf disease detection and classification in food crops with efficient feature dimensionality reduction," *PLOS One*, vol. 20, no. 8, Aug. 2025, doi: 10.1371/journal.pone.0328349.
[3] K. Kira and L. A. Rendell, "A practical approach to feature selection," in *Machine Learning Proceedings 1992*, Burlington, United States: Morgan Kaufmann, 1992, pp. 249–256, doi: 10.1016/B978-1-55860-247-2.50037-1.
[4] S. C. Girones, S. Munera, M. M. Sober, J. Blasco, S. Cubero, and J. G. Sanchis, "Artificial neural networks in agriculture, the core of artificial intelligence: what, when, and why," *Computers and Electronics in Agriculture*, vol. 230, Mar. 2025, doi: 10.1016/j.compag.2025.109938.
[5] P. Sridhar and P. Angamuthu, "Enhancing image based classification for crop disease detection using a multiclass SVM approach with kernel comparison," *Scientific Reports*, vol. 15, no. 1, Nov. 2025, doi: 10.1038/s41598-025-23568-w.
[6] T. Nyawose, R. C. Maswanganyi, and P. Khumalo, "A review on the detection of plant disease using machine learning and deep learning approaches," *Journal of Imaging*, vol. 11, no. 10, Sep. 2025, doi: 10.3390/jimaging11100326.
[7] B. Sambana *et al.*, "An efficient plant disease detection using transfer learning approach," *Scientific Reports*, vol. 15, no. 1, May 2025, doi: 10.1038/s41598-025-02271-w.
[8] M. B. Diallo *et al.*, "Enhanced-RICAP: a novel data augmentation strategy for improved deep learning-based plant disease identification and mobile diagnosis," *Frontiers in Plant Science*, vol. 16, Sep. 2025, doi: 10.3389/fpls.2025.1646611.
[9] E.-M. Nikolados, A. Wongprommoon, O. M. Aodha, G. Cambray, and D. A. Oyarzún, "Accuracy and data efficiency in deep learning models of protein expression," *Nature Communications*, vol. 13, no. 1, Dec. 2022, doi: 10.1038/s41467-022-34902-5.
[10] Y. Bahri, E. Dyer, J. Kaplan, J. Lee, and U. Sharma, "Explaining neural scaling laws," *Proceedings of the National Academy of Sciences*, vol. 121, no. 27, Jul. 2024, doi: 10.1073/pnas.2311878121.
[11] M. T. Ribeiro, S. Singh, and C. Guestrin, "'Why should I trust you?': explaining the predictions of any classifier," in *22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Aug. 2016, pp. 1135–1144, doi: 10.1145/2939672.2939778.
[12] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: visual explanations from deep networks via gradient-based localization," in *2017 IEEE International Conference on Computer Vision*, Oct. 2017, pp. 618–626, doi: 10.1109/ICCV.2017.74.
[13] R. P. Porfirio, R. N. Madeira, and P. A. Santos, "Exploring explainable AI techniques for plant disease classification in digital agriculture," *Procedia Computer Science*, vol. 265, pp. 175–182, 2025, doi: 10.1016/j.procs.2025.07.170.
[14] R. Kusuma and R. Rajkumar, "Plant leaf disease detection and classification using artificial intelligence techniques: a review," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 38, no. 2, pp. 1308-1323, May 2025, doi: 10.11591/ijeecs.v38.i2.pp1308-1323.
[15] V. Boppana, "Plant disease detection using hybrid MobileNetV2-compact CNN architecture with LIME integration," *Journal of Information Systems Engineering and Management*, vol. 10, no. 13s, pp. 554–568, Feb. 2025, doi: 10.52783/jisem.v10i13s.2111.
[16] M. Nahiduzzaman *et al.*, "Explainable deep learning model for automatic mulberry leaf disease classification," *Frontiers in Plant Science*, vol. 14, Sep. 2023, doi: 10.3389/fpls.2023.1175515.
[17] V. M. Gutiérrez *et al.*, "AI-driven plant health assessment: a comparative analysis of InceptionV3, ResNet-50 and ViT with SHAP for accurate disease identification in taro," *Agronomy*, vol. 15, no. 1, Dec. 2024, doi: 10.3390/agronomy15010077.
[18] S. M. Alhammad, D. S. Khafaga, W. M. El-hady, F. M. Samy, and K. M. Hosny, "Deep learning and explainable AI for classification of potato leaf diseases," *Frontiers in Artificial Intelligence*, vol. 7, Feb. 2025, doi: 10.3389/frai.2024.1449329.
[19] K. Gopalan, S. Srinivasan, Pragya, M. Singh, S. K. Mathivanan, and U. Moorthy, "Corn leaf disease diagnosis: enhancing accuracy with resnet152 and grad-cam for explainable AI," *BMC Plant Biology*, vol. 25, no. 1, Apr. 2025, doi: 10.1186/s12870-025-06386-0.

[20] M. J. Karim, M. O. F. Goni, M. Nahiduzzaman, M. Ahsan, J. Haider, and M. Kowalski, "Enhancing agriculture through real-time grape leaf disease classification via an edge device with a lightweight CNN architecture and Grad-CAM," *Scientific Reports*, vol. 14, no. 1, Jul. 2024, doi: 10.1038/s41598-024-66989-9.
[21] S. Mostafa, D. Mondal, K. Panjvani, L. Kochian, and I. Stavness, "Explainable deep learning in plant phenotyping," *Frontiers in Artificial Intelligence*, vol. 6, Sep. 2023, doi: 10.3389/frai.2023.1203546.
[22] J. Amara, B. K. Ries, and S. Samuel, "Concept explainability for plant diseases classification," in *18th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, 2023, pp. 246–253, doi: 10.5220/0011667900003417.
[23] J. Lu, X. Liu, X. Ma, J. Tong, and J. Peng, "Improved MobileNetV2 crop disease identification model for intelligent agriculture," *PeerJ Computer Science*, vol. 9, 2023, doi: 10.7717/peerj-cs.1595.
[24] H. K. Kondaveeti and C. G. Simhadri, "Evaluation of deep learning models using explainable AI with qualitative and quantitative analysis for rice leaf disease detection," *Scientific Reports*, vol. 15, no. 1, 2025, doi: 10.1038/s41598-025-14306-3.
[25] A. Bhujel, N.-E. Kim, E. Arulmozhi, J. K. Basak, and H.-T. Kim, "A lightweight attention-based convolutional neural networks for tomato leaf disease classification," *Agriculture*, vol. 12, no. 2, Feb. 2022, doi: 10.3390/agriculture12020228.
[26] A. Mumuni and F. Mumuni, "Data augmentation: a comprehensive survey of modern approaches," *Array*, vol. 16, Dec. 2022, doi: 10.1016/j.array.2022.100258.
[27] A. R. Al-Shamasneh and R. W. Ibrahim, "Classification of tomato leaf images for detection of plant disease using conformable polynomials image features," *MethodsX*, vol. 13, Dec. 2024, doi: 10.1016/j.mex.2024.102844.

BIOGRAPHIES OF AUTHORS

Rakesh Kumar Gumasta is presently pursuing a Ph.D. in the Department of Electronics and Communication Engineering at Maulana Azad National Institute of Technology, Bhopal, India. He received an M.E. degree in Digital Communication Engineering from UIT-RGPV Bhopal. He has more than 12 years of experience in research, development, and teaching. He has conducted training for industries. His areas of expertise are MATLAB, VLSI, and AI. He can be contacted at email: rakeshgumasta44@gmail.com.

Dr. Ajay Somkuwar is presently working as a professor in the Department of Electronics and Communication at Maulana Azad National Institute of Technology. He received his Ph.D. from Indian Institute of Technology, New Delhi, India, in 2003. He has published more than 100 papers of national and international repute. His research area includes signal processing, speech signal processing, and biosignal engineering. He can be contacted at email: Asomkuwar@gmail.com.