WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

An Explainable Deep Learning Framework for Automated Plant Disease Detection and Classification Using Leaf Images

Dhanush Kumar. P[1], Yashema. N[2], Balaji. S[3]

[1234]Department of computer science and Engineering Kingston Engineering College, Vellore, Tamil Nadu, India
[2]AP/CSE, [3]HOD/CSE
Dhanushkumar2k22@gmail.com, yash.engineering@kingston.ac.in, balajicse.engineering@kingstion.ac.in

A B S T R A C T :

Crop diseases remain a major obstacle to global food security since early detection can avoid serious losses in crop yield. We have developed a deep learning model which is able to identify and classify plant diseases from leaves in a simple and clear manner, the image quality being improved through preprocessing and augmentation. Convolutional neural networks (CNNs) and transfer learning are then used to carry out the classification of the diseases. Moreover, Grad-CAM (Gradient-weighted Class Activation Mapping) has been included in order to help explain the reasons why the deep learning model makes the decisions it does; this is intended to reduce some of the "black box" aspect associated with deep learning. The model was evaluated using accuracy, precision, recall, the F1 score and a confusion matrix. The results we obtained indicate that the new deep learning framework we have created can detect and classify crop diseases accurately, transparently, and efficiently, allowing farmers to carry out precision agriculture effectively.

2. Keywords
Detection of plant diseases; deep learning; explainable AI; classification of leaf images; CNN; transfer learning; Grad-CAM; precision agriculture.

1. Introduction

Agriculture plays a vital role in supporting human life and promoting economic development, but plant diseases frequently endanger both crop yields and their quality. Such diseases, which are caused by pathogens like fungi, bacteria, and viruses, usually show themselves by means of visible alterations in the leaves, such as discolouration, the appearance of lesions, and odd patterns. In the past, farmers have depended on manual observation to detect these symptoms, a method which is slow, subjective, and requires a great deal of expert knowledge. Since global food demand is increasing, there is now a greater urgency in needing efficient, scalable, and reliable ways of detecting disease.

New developments in computer vision and deep learning have opened up the possibility of automating the diagnosis of plant diseases. Convolutional Neural Networks (CNNs) and transfer learning models are able to examine leaf images and pick up on visual features that enable them to tell the difference between healthy and diseased plants. Because they provide a level of consistency and speed that is superior to manual inspection, such models are useful in the field of agriculture. Yet, even though their predictive ability is good, the reasons for their predictions are generally not clear, which in turn reduces user confidence.

2. Problem Statement

One of the greatest dangers facing agriculture is plant diseases, since they lead to lower yields and alterations in both quality and economic value. The problems involved include the need to manually look for symptoms on the leaves, a process which is time-consuming, subjective and calls for expert knowledge. Moreover, many diseases produce similar symptoms on the leaves, which makes it hard for farmers to diagnose them in the field.

Even though deep learning models have shown promising results, they have a number of limitations. Since many of them act as black boxes, they provide no explanation of how their decisions are reached, which in turn affects users' confidence. Moreover, most of the research makes use of data gathered in controlled environments, so the models cannot easily be applied to real-world situations in which lighting, backgrounds, and the orientations of the leaves differ greatly. For these reasons, there is a need for a framework that can predict outcomes both accurately and transparently and thus give a basis for confidence in practical crop management.

3. Objectives

1. 1. Create a system for automatic disease detection and classification from plant leaf images.
2. Carry out preprocessing techniques such as image resizing, normalization, and noise reduction in order to ensure uniformity and consistency among the input images.

3. 3. Perform data augmentation, such as rotation, flipping, or brightness adjustments, to increase diversity and improve model robustness.
4. Carry out the implementation of the CNN and other transfer learning models in order to automatically extract visual features such as colour, texture, lesions, etc.
5. 5. Classify plant leaves into healthy and diseased categories with high accuracy and reliability.
6. Apply explainable AI methods by implementing the Grad-CAM technique in order to provide a visual explanation of the areas that are affecting the predictions.
7. Evaluate the performance using criteria like the F1 score in order to determine the efficiency of the model.
8. Contribute to precision agriculture by providing the farming community with predictions that are interpretable.

4. Literature Review

As Mohanty and others pointed out in 2016, a pioneering application of deep learning in the area of plant disease detection made use of the Plant Village database. This study applied convolutional neural networks (CNNs) in order to classify leaves that were either healthy or affected in different types of crops. The results showed that CNNs are able to achieve correct classification and thus pave the way for the use of deep learning in agriculture; nevertheless, the limitations associated with the controlled images were clear, as was the absence of an explanation for the predictions.

Yao and others in 2024 looked into the effect of deep learning on the identification of plants and the recognition of plant diseases. They used a number of different kinds of CNNs and created a Generalised Stacking Multi-output CNN (GSMo CNN) that enabled the simultaneous prediction of both the type of plant and the type of disease. Their results on databases including PlantVillage, Plant Leaves, and PlantDoc indicate a higher level of classification accuracy; yet they did not deal with the interpretability of the results.

Amara and others in 2023 dealt with transparency by using a concept-based method; they employed Testing with Concept Activation Vectors (TCAV), which offered explanations that were understandable to humans. This approach increased the transparency of automated diagnosis by associating the model's outputs with relevant concepts. The study, however, concentrated solely on concept-level reasoning and thus was not able to develop a full methodology that links automated classification with practical visual explanations for the agricultural community.

5. Research Gap

Even though deep learning technologies have developed tremendously, some issues remain unresolved. Most of the current techniques focus on maximizing classification accuracy without giving any explanations on how their results can be interpreted, which reduces users' trust.

Another limitation of the existing related works is using datasets that usually consist of images of leaves taken under controlled conditions, while in reality, there are many complications which make classification more difficult. Also, many studies concentrate on a particular crop or diseases, thus the general efficiency of their methods is limited. Lastly, many symptoms of various diseases may be similar visually.

The application of Explainable Artificial Intelligence (XAI) is not widespread in this field at the moment, but the proposed study aims to change this tendency by proposing an improved approach that will allow combining deep learning-based classification and XAI techniques.

6. Proposed System

Within the proposed framework called "An Explainable Deep Learning System for Automated Plant Disease Detection and Classification Using Leaf Images," there is a clear emphasis on accurately detecting plant diseases from leaf images in addition to providing better interpretations of the predictions. By using a combination of deep learning and Explainable Artificial Intelligence (XAI), the presented approach looks to solve the issues associated with human diagnosis and complicated machine learning models.

Firstly, unsorted datasets containing healthy and very ill leaf pictures are collected. This will involve preprocessing of the leaf images which includes resizing, normalizing, filtering the noise, and generating more data. After obtaining the improved versions of the leaf pictures, the pictures are passed over to a Convolutional Neural Network (CNN) or transfer learning model, which will automatically recognize the colour differences, the presence of lesions, spotted areas, various textures, and other important features which characterize pests and plant diseases.

Using the received features, the classifier will make a prediction regarding whether the leaf is sick or not. Meanwhile, the performance of the model will be measured amid several metrics such as classification accuracy, precision, recall, F1 score, or a confusion matrix. To ensure good levels of interpretability, the model includes Grad CAM, which is one of the techniques used in XAI. This means that during the evaluation, heatmap images would be generated showing which parts of the image were most significant in making predictions about the wellbeing of the plant, which is important for farmers and agronomists working in the field.

All in all, the framework will generate a disease name, confidence level, as well as a heatmap based on which farmers would be able to find out the reasons for the diagnosis.

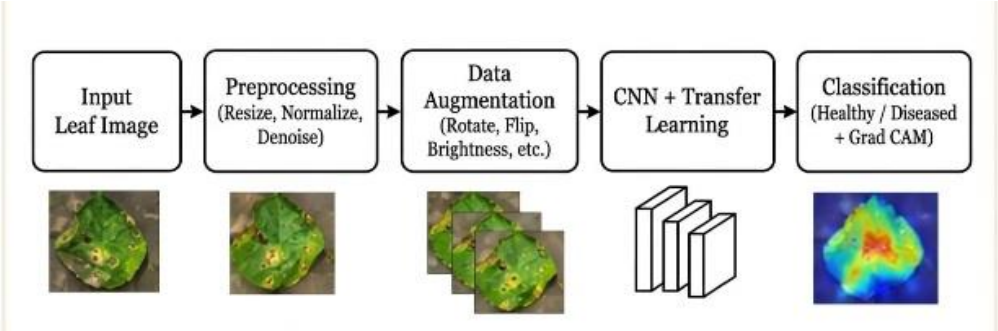
7. Methodology

The method described consists of an automated procedure for identifying and classifying the various diseases from leaf images. This is carried out through the gathering of images followed by preprocessing steps like size consistency, normalization, and reduction of the noise present in the images. The method uses augmentation techniques that increase data variability and consequently enhance the model's generalization.

Deep learning techniques, particularly the use of convolutional neural networks (CNNs) or transfer learning techniques are used next to capture the distinguishing features from the images that were processed in the previous step. These features make it possible for the system to differentiate the healthy

leaves and the leaves that have been infected by the diseases.

Explainable AI through the use of Grad CAM is integrated into the process to enhance transparency. In addition, this technique uses heat maps to show which parts of the leaves had an influence on the outcome of identification. Evaluation of accuracy, precision, recall, and F1 score is done to verify the effectiveness of the method.



1.Dataset Collection

Pictures of healthy and sick plant leaves were obtained from openly available data bases such as PlantVillage and PlantDoc. The datasets make use of a variety of plant species and diseases. The images are then divided into training, validation, and testing datasets depending on how efficient model development can be done.

2. Image preprocessing

Because of the varying resolution, illumination, and background complexities of the original images, they need to be preprocessed prior to their use for deep learning. This consists of making the images of a particular size, changing pixel values, and using methods to reduce noise.

3. Data Augmentation

In order to boost the performance of the model in terms of its generalization capability and to ensure minimal overfitting, several augmentation techniques were used on the training dataset. The processes include rotation, flipping, zoom, shifting, cropping, and changing brightness of the pictures. By producing different variations of the original images, the model gets access to more diverse situations.

4. Deep Learning Model Development

Regarding the classification of diseases, the framework utilizes Convolutional Neural Networks (CNNs) or advanced transfer learning models such as ResNet, EfficientNet and MobileNet. These architectures automatically learn and extract important visual fields from leaves including details about color, texture, spots, lesions, and disease patterns. Based on these deep insights, the system is capable of differentiating between healthy and sick leaves among different plant species.

5. Model Training

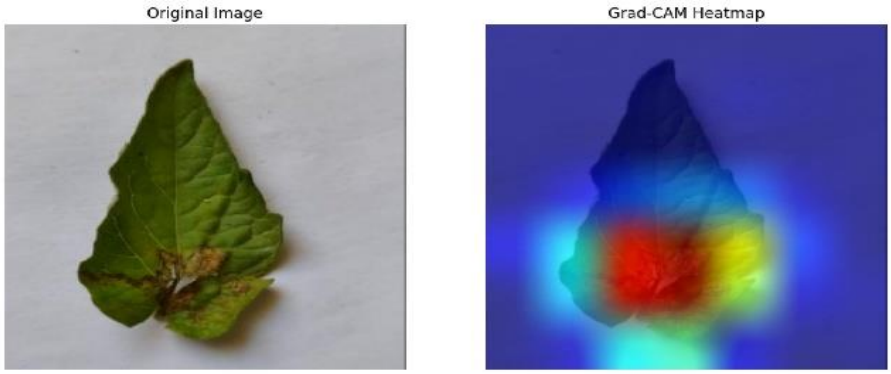
This type of training takes the training data that has been prepared and puts it into the deep learning architecture that has been selected. The model learns to connect image characteristics to various disease categories using the correct loss functions and optimization algorithms in order to minimize classification mistakes. A validation data set is used throughout the training process in order to check the model and decrease the possibility of overfitting.

6. Disease Classification

After the model has been trained, it is able to process new leaf images and classify them accordingly. This classification involves providing the class predicted by the model (whether healthy or diseased) as well as the degree of confidence of the prediction. The efficiency of the system will depend on the variety of plant species and diseases present in the dataset.

7. Explainable AI Integration

In order to overcome the problem of transparency in deep learning systems, Explainable AI technologies are used in the project. Grad CAM (Gradient-weighted Class Activation Mapping) helps produce “heat maps”, which are images showing which parts of the leaf played an important role in making the decision. Such visualization will help farmers and agronomists understand how the model works and make it more trustworthy.



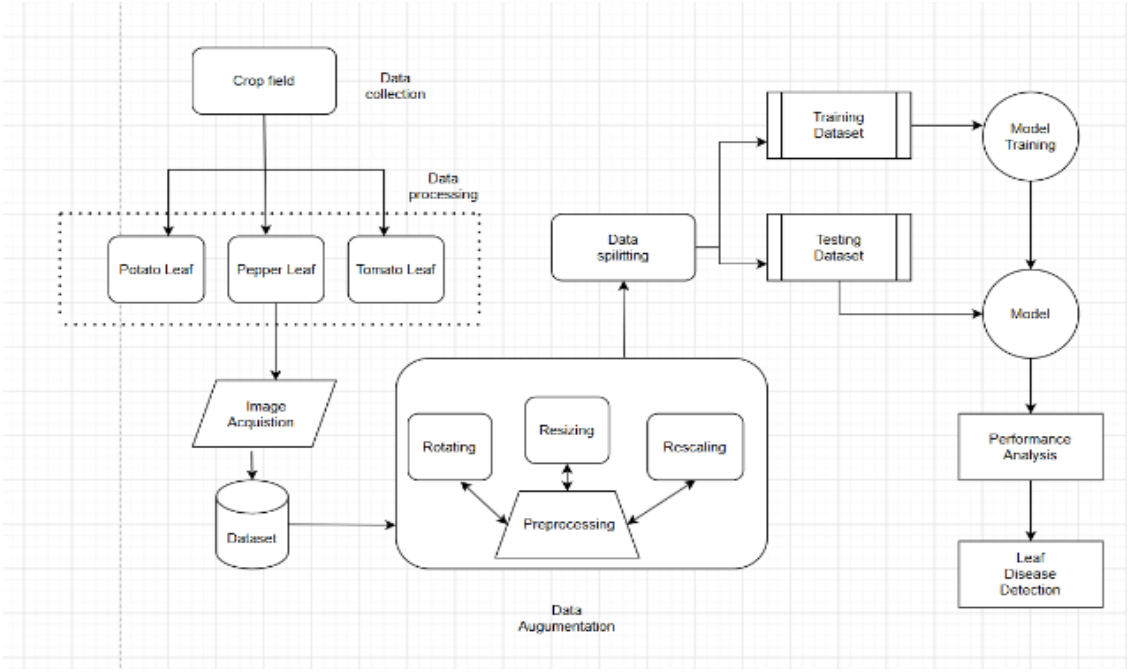
8. Performance Evaluation

The performance of the trained machine model is evaluated based on well-known performance indicators. These performance indicators include accuracy, precision, recall, F1 score, and confusion matrix. All these indicators provide a detailed evaluation of how well this technique distinguishes different types of plant diseases from healthy leaves.

9. Prediction and Visualization

When a prediction is made by the system, the input image of the leaf is presented including its disease category, confidence score in the form of a Grad CAM heat map. The Grad CAM heat map shows the area of the leaf based on which the decision was made, making it easier for farmers and agricultural specialists to understand the basis on which the prediction was made, thus increasing their trust in the system.

8.Architecture Diagram



This flowchart illustrates the methodology of the Plant Leaf Disease Detection Framework using the methodology of Convolutional Neural Network (CNN). The first validation or input in this methodology is the preparation of an image data set consisting of the images of healthy and diseased leaves, which is then used to train CNN. This images are first prepared by transforming and improving their quality by means of resizing, normalization, and noise reduction As soon as the images are ready, they are passed into the CNN network that is trained to recognize significant visual features that include any variations in color, spots, lesions, or specific patterns of diseases. After the model has been trained, it has the ability to classify any new leaf images by making the comparison between the images features and those learned ones while the model was being trained. In the end, this automated framework makes the classification of the leaves, whether healthy or infected the specific diseases it suffers from. This classification allows to provide quick and effective detection of the diseases and conduct timely measures in order to save the plants.

9. System Design

The Plant Disease Prediction Framework is a classification system that utilizes images, along with Convolutional Neural Networks (CNNs), to identify diseases on plant leaves. The framework processes images of leaves, identifies morphological characteristics that differentiate healthy leaves from the diseased leaves, and provides a diagnosis early in the process and classifies the disease automatically.

9.1 Image Collection Module

Images of leaves for the application are sourced from various datasets that provide images of healthy as well as diseased crops. Each of the images is tagged according to the disease it is infected by and the diseases are successfully identified with the help of supervised learning.

9.2 Image Preprocessing Module

Before the training process, the collected image data undergoes significant preprocessing to ensure that the data undergoes similar transformations. The preprocessing includes resizing images to a standard size, normalizing pixel values, and applying image augmentation techniques such as rotation, flipping, and zoom. All these procedures are meant to produce high-quality images and enhance the dataset diversity.

9.3 Dataset Splitting Module

The processed dataset is split into three parts to create train, test, and validation datasets. The training dataset is used to train the final model, the validation dataset is used to tune model parameters, and the test dataset is used to assess the final performance.

9.4 CNN Model Module

The primary classifier in this framework is built in the form of a Convolutional Neural Network that extracts visual features of images on its own without relying on any manual work. It is capable of identifying the color, texture, lesions, and features of the leaf showing signs of diseases.

9.5 Feature Extraction and Classification Module

The convolution and pooling components of the model allow meaningful characteristics to be understood from the input images. These features are processed through the layers of the connection paths. In the end, the leaves are classified into proper categories.

9.6 Disease Prediction Module

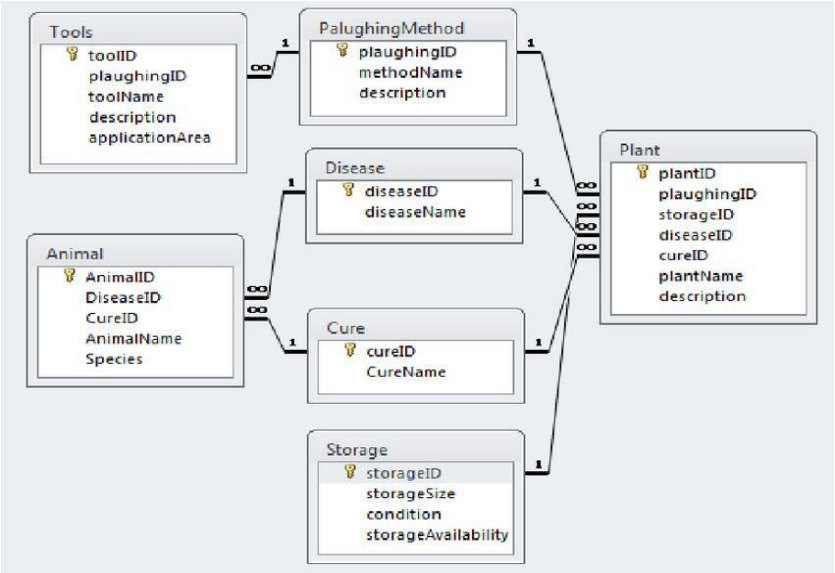
When a leaf image enters the system, the image is ready for future processing. The model will then analyze the image using the gained knowledge.

9.7 Result Display Module

The system produces the final output to its user. The output may involve such labels as “Healthy Leaf”, “Early Blight”, “Late Blight”, etc., characterized in the dataset. Along with the result, the system may also present a score of reliability.

10.Database Design

The design of the database within the Plant Disease Prediction System is achieved in the correct manner to hold information regarding plants, leaf images, disease records, and outcomes of prediction. The system consists of the tables connected with each other. Therefore, there are tables such as Users where the details regarding users are held, Plant Images where the metadata and references of the uploaded leaf images appear, Diseases with the information about diseases, symptoms, and others, and Predictions which represent the results given by the classifier.



10.1 The User Table

The User table keeps records for those individuals who use the Plant Disease Prediction System. Besides the identification, the table preserves the details for login. The fields are:

- User_ID
- Name
- Email
- Password
- Role
- Contact_Number

10.2 Plant Images Table

The Plant Images Table manages the information concerning leaf images uploaded for the purposes of disease detection. The uploaded image obtains a unique identifier per each image. The fields are:

- Image_ID
- User_ID
- Image_Path
- Plant_Name
- Upload_Date
- Image_Status

10.3 The Diseases Table

The Diseases Table contains the information of plant diseases known by the classifier including their symptoms as well as descriptions. The fields are:

- Disease_ID
- Disease_Name
- Plant_Name
- Symptoms
- Description
- Treatment

10.4 Predictions Table

The Predictions Table keeps the information regarding the results of each image uploaded including the disease predicted. The fields are:

- Prediction_ID
- Image_ID
- Disease_ID
- Predicted_Result
- Confidence_Score
- Prediction_Date

10.5 Feedback Table

This table keeps the users' opinion as well as the score given according to the results received which can assist in the improvement of the system. The fields are:

- Feedback_ID
- User_ID
- Prediction_ID
- Feedback_Text
- Rating
- Feedback_Date

10.6 Reports Table

This table summarizes the activity of the system for the monitoring and analysis purposes. The fields are:

- Report_ID
- Report_Type
- Generated_Date
- Total_Predictions
- Healthy_Count
- Diseased_Count
- Most_Detected_Disease

11. Result and Discussion

The Plant Disease Recognition System is capable of detecting diseases in plants through the examination of leaf images by utilizing machine learning/deep learning methods. The system looks into the visual features of the leaf, e.g., color, appearance, texture, and disease symptoms, and based on that, the appropriate disease category is diagnosed. The results confirm the great predictive potential of the model, as the system can efficiently recognize diseased leaves from healthy ones.

The present finding suggests that the system can minimize the time and work made for manual disease detection. Nevertheless, the quality of the predicted results could be impacted by image quality, illumination, background, and similarities between diseases. A better performance of the developed model could be achieved through a bigger and more varied dataset.

Feature	Existing	Proposed
Disease Detection	Manual	Automated
Detection Method	Visual inspection	ML/DL
Prediction Speed	Slow	Fast
Accuracy	Variable	Improved
Disease Identification	Difficult	Pattern-based
Image Processing	Not available	Automatic
Early Detection	Limited	Supported
Accessibility	Expert needed	Easy
Data Analysis	Manual	Automated

12. Conclusion

The Plant Disease Prediction System provides an effective method for detecting diseases in plants through image analysis using advanced machine learning and deep learning technologies. This automated process makes it possible to identify crop diseases from leaf pictures much faster than traditional manual methods. Since diseases can now be detected earlier, the need to hire special experts to oversee the process is minimized. Furthermore, it leads to better crop management and fewer losses in agriculture as farmers will be aware of certain problems in advance.

The potential of the system shows that the use of modern technologies in agriculture can lead to agriculture improvements since it makes it possible to use automation and innovative prediction solutions that can protect crops, increase productivity levels, and contribute to sustainable agriculture.

13. Future Scope

In the foreseeable future, the Plant Disease Prediction System aims to develop a mobile application that provides farmers with an effective and simple method of disease detection in the field. The extension of the existing system is critical in its versatility and applicability in agriculture as it will allow tracking diseases not only in the selected group of plants but also enabling the system to detect various other diseases affecting different kinds of plants. The usage of advanced data mining tools will guarantee not only accurate but also timely disease identification.

The integration with various IoT solutions, such as IoT sensors and weather information systems, will help to make the system more robust and provide it with some kind of disease surveillance capability. This improvement will enable early identification of the potential risks and preventive action based on information provided by IoT solutions.

14. References

1. Shoaib, M., et al. (2023). An advanced deep learning models-based plant disease detection: A review of recent research. Frontiers in Plant Science, 14, 1158933. https://doi.org/10.3389/fpls.2023.1158933
2. Plant Disease Detection and Classification: A Systematic Literature Review. (2023). Sensors, 23(10), 4769. https://doi.org/10.3390/s23104769
3. Demilie, W. B. (2024). Plant disease detection and classification techniques: A comparative study of the performances. Journal of Big Data, 11, Article 5. https://doi.org/10.1186/S40537-023-00863-9
4. A survey on using deep learning techniques for plant disease diagnosis and recommendations for development of appropriate tools. (2023). Smart Agricultural Technology, 3, 100083. https://doi.org/10.1016/j.atech.2022.100083
5. Das, P. K., Rupa, S. S., Pumrin, S., Das, U. C., & Hossen, M. K. (2024). Deep learning for plant disease detection and classification: A systematic analysis and review. Current Applied Science and Technology, 24(4), Article e0259016. https://doi.org/10.55003/cast.2024.0259016
6. 1. Wei, K., Chen, B., Zhang, J., Fan, S., Wu, K., Liu, G., & Chen, D. (2022). Explainable deep learning study for leaf disease classification. Agronomy, 12(5), 1035. https://doi.org/10.3390/agronomy12051035
7. 4. Prince, R. H., Mamun, A. A., Peyal, H. I., Miraz, S., Nahiduzzaman, M., Khandakar, A., & Ayari, M. A. (2024). CSXAI: A lightweight 2D CNN-SVM model for detection and classification of various crop diseases with explainable AI visualization. Frontiers in Plant Science, 15, 1412988. https://doi.org/10.3389/fpls.2024.1412988
8. Improving plant disease classification with deep learning based prediction model using explainable artificial intelligence. (2024). IEEE Access, 12, 100005–100014. https://doi.org/10.1109/ACCESS.2024.3428553
9. Richter, D. J., Bappi, M. I., Kolekar, S. S., & Kim, K. (2025). A systematic review of the current state of transfer learning accelerated CNN-based plant leaf disease classification. IEEE Access, 13, 116262–116303. https://doi.org/10.1109/ACCESS.2025.3584404
10. Ismail, U. I., Chua, H. N., Nordin, R., & Lee, H. L. (2026). Explainable artificial intelligence for plant disease detection: A scoping review of methods, datasets, and evaluation practices. Computers and Electronics in Agriculture. https://doi.org/10.1016/j.compag.2026.112338

11. A survey on using deep learning techniques for plant disease diagnosis and recommendations for development of appropriate tools. (2023). Smart Agricultural Technology, 3, 100083. https://doi.org/10.1016/j.atech.2022.100083

12. Natarajan, S., Chakrabarti, P., & Margala, M. (2024). Robust diagnosis and meta visualizations of plant diseases through deep neural architecture with explainable AI. Scientific Reports, 14, Article 13695. https://doi.org/10.1038/s41598-024-64601-8

13. Porfírio, R. P., Madeira, R. N., & Santos, P. A. (2025). Exploring explainable AI techniques for plant disease classification in digital agriculture. Procedia Computer Science, 265, 175–182. https://doi.org/10.1016/j.procs.2025.07.170

14. Gopalan, K., Srinivasan, S., Agarwal, P., Singh, M., Mathivanan, S. K., et al. (2025). Corn leaf disease diagnosis: Enhancing accuracy with ResNet152 and Grad-CAM for explainable AI. BMC Plant Biology, 25, Article 440. https://doi.org/10.1186/s12870-025-06386-0

15. Hasan, M. A., Haque, F., Sarker, H., Abdullah, R., Roy, T., Taaha, N., et al. (2025). Mulberry leaf disease detection by CNN-ViT with XAI integration. PLOS ONE, 20(6), e0325188. https://doi.org/10.1371/journal.pone.0325188