



Deep Learning-Based Plant Disease Detection Using Convolutional Neural Networks (CNN)

Vijay S

Assistant Professor, Sacred Heart College(Autonomous), Tirupattur District, Tirupattur and 635601, Tamil Nadu

ABSTRACT :

Plant diseases are a major challenge in agriculture because they can reduce crop productivity, affect food quality, and cause economic losses. Conventional plant disease identification depends largely on visual inspection by farmers and agricultural experts, which can be time-consuming and may result in inaccurate classification. Recent advances in deep learning and computer vision provide opportunities to automate disease identification from plant leaf images. This paper proposes a deep learning-based plant disease detection system using Convolutional Neural Networks (CNN). The proposed framework consists of image acquisition, pre processing, feature extraction, model training, disease classification, and result visualization. The Plant Village dataset is proposed as the initial dataset for training and evaluating the model. The system aims to classify leaf images into healthy and disease-specific categories and provide preliminary disease information to users. Performance will be assessed using accuracy, precision, recall, F1-score, and a confusion matrix. A transfer-learning model such as MobileNetV2 may also be evaluated to compare classification performance and computational efficiency. The proposed system is intended to support early disease screening and improve access to plant health information. Further validation using images collected under real agricultural conditions is necessary before practical deployment.

Keywords: Artificial Intelligence, Deep Learning, Plant Disease Detection, Convolutional Neural Network, Image Classification, Computer Vision, Precision Agriculture.

Introduction

Agriculture plays an essential role in economic development and food security. However, plant diseases caused by fungi, bacteria, viruses, and other pathogens can significantly affect crop production. Early identification of plant diseases is important for reducing crop losses and supporting appropriate agricultural management.

Traditional disease identification methods generally involve visual examination of plant leaves by farmers or agricultural specialists. Although these methods are widely used, they can be affected by limited expertise, environmental conditions, symptom similarity, and the time required to inspect large numbers of plants.

Artificial Intelligence (AI) and deep learning have introduced new approaches to automated plant disease recognition. Convolutional Neural Networks are particularly suitable for image classification because they automatically learn visual features such as leaf texture, colour variation, spots, and lesion patterns.

The proposed research develops a CNN-based plant disease classification framework that accepts a leaf image and predicts its corresponding class. The system is intended to provide an accessible preliminary screening tool for agricultural applications.

1.1. Problem Statement

Manual plant disease identification can be time-consuming and inconsistent, particularly when several diseases exhibit similar symptoms. An automated image-classification system is therefore needed to support preliminary disease identification using deep learning techniques.

1.2. Research Objectives

The primary objectives of this research are:

1. To develop a CNN-based model for plant leaf disease classification.
2. To preprocess and classify plant leaf images using deep learning.
3. To evaluate the model using standard classification metrics.
4. To develop a user-friendly interface for image submission and prediction.

5. To investigate transfer learning as an approach to improve efficiency.
6. To identify limitations associated with applying laboratory-trained models to field conditions.

Literature Review

Deep learning has become an important research direction in automated plant disease recognition.

Mohanty et al. demonstrated the use of deep convolutional neural networks for image-based plant disease detection using the PlantVillage dataset. Their work established a widely cited benchmark for classifying plant health conditions from leaf images. However, high performance on controlled images does not automatically establish equivalent performance under field conditions.

Liu and Wang reviewed deep-learning approaches for detecting plant diseases and pests. Their review discussed image classification, detection techniques, datasets, and challenges associated with practical deployment.

Abade et al. conducted a systematic review of CNN-based plant disease recognition. Their work examined the development of deep-learning methods, the datasets used in research, and the challenges that affect model generalization.

Shoaib et al. reviewed advanced deep-learning approaches for plant disease detection and discussed issues relating to dataset diversity, model selection, and practical agricultural applications.

The literature indicates that CNNs are promising for plant disease image classification. Nevertheless, differences in illumination, background, camera quality, crop varieties, and disease severity can affect classification performance. The proposed study therefore emphasizes independent testing and recommends validation using field-collected images.

Proposed Methodology

The proposed system follows a sequence of image acquisition, preprocessing, model training, classification, and result presentation.

3.1 Dataset Collection

The Plant Village dataset is proposed as the initial dataset because it contains labelled images of healthy and diseased plant leaves across multiple crop types. The initial implementation may focus on tomato, potato, or pepper leaves. The selected classes should be documented clearly, including the number of images available in each class. The dataset should be divided into training, validation, and testing subsets. An example split is 70% for training, 15% for validation, and 15% for testing. The test set must remain separate from model training and hyper parameter selection.

3.2 Image Preprocessing

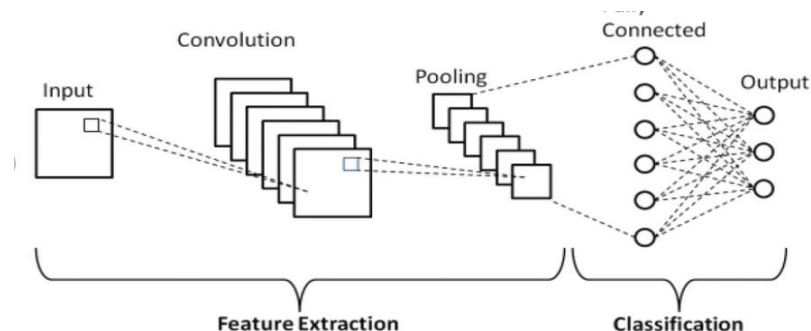
Image preprocessing prepares input images for consistent model training and prediction. The proposed preprocessing operations include:

- Image loading and format conversion.
- Resizing images to a standard resolution, such as 224×224 pixels.
- Pixel-value normalization according to the selected model.
- Data augmentation during training, including modest rotation, flipping, and zooming.
- Removal or handling of unreadable images and invalid labels.

Data augmentation should be applied only to the training subset. The validation and test sets should not leak information into training.

3.3 Convolutional Neural Network Architecture

A CNN learns hierarchical features directly from image pixels. The proposed model includes the following layers:



The model can be trained using the Adam optimizer and categorical or sparse categorical cross-entropy loss, depending on the label representation.

3.4 Transfer Learning

Transfer learning may be implemented using MobileNetV2 pretrained on ImageNet. The pretrained feature extractor can be adapted to the selected plant disease classes by replacing the classification head. Initially, the classification head can be trained while the pretrained layers remain frozen. Selected layers may subsequently be fine-tuned using a suitably small learning rate.

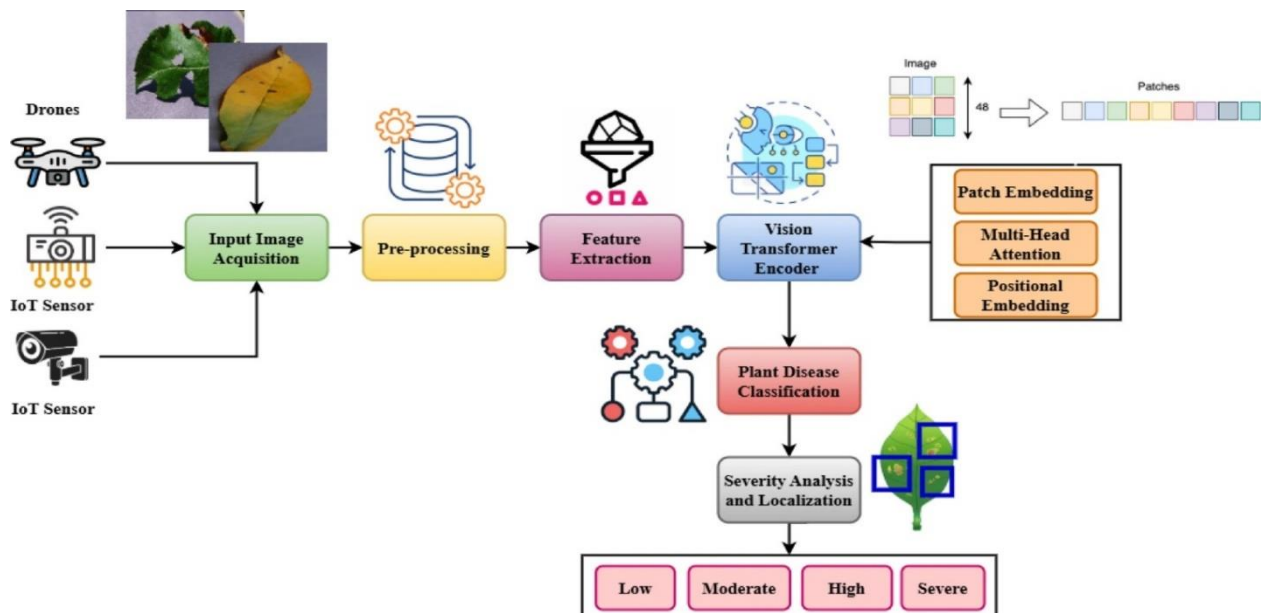
The custom CNN and transfer-learning model should be evaluated using the same data partitions to support a fair comparison.

3.5 Disease Classification

When a user uploads a leaf image, the application preprocesses it and passes it to the trained model. The model generates class probabilities and selects the class with the highest predicted probability.

The system displays the predicted class, the model's confidence score, and relevant general disease information. The output should be presented as a preliminary prediction rather than a guaranteed diagnosis.

4. System Architecture



5. Experimental Design and Evaluation

The model should be trained and evaluated using a documented experimental procedure. The following metrics are recommended.

5.1 Accuracy

Accuracy measures the proportion of correctly classified test samples.

Accuracy = Number of Correct Predictions / Total Number of Test Samples.

5.2 Precision

Precision measures the proportion of positive predictions that are correct.

Precision = $TP / (TP + FP)$.

5.3 Recall

Recall measures the proportion of actual positive samples correctly identified.

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN}).$$

5.4 F1-Score

The F1-score combines precision and recall using their harmonic mean.

$$\text{F1-score} = 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall}).$$

For multiclass classification, these metrics should be calculated for each class and reported using macro averages as well as class-level results.

5.5 Confusion Matrix

A confusion matrix helps identify which disease categories are frequently misclassified. It is particularly useful when different diseases have similar visual symptoms.

5.6 Proposed Comparison

The following table should be completed using the actual results of the experiments.

Evaluation Metric	Custom CNN	MobileNetV2
Test accuracy	To be measured	To be measured
Macro precision	To be measured	To be measured
Macro recall	To be measured	To be measured
Macro F1-score	To be measured	To be measured
Model size	To be measured	To be measured
Inference time	To be measured	To be measured

No numerical results are reported here because the models have not yet been experimentally evaluated as part of this manuscript.

6. Results and Discussion

The experimental phase should assess whether the proposed model can distinguish healthy leaves from the disease classes represented in the training dataset. The final results should include the training and validation curves, test-set metrics, confusion matrix, and representative correct and incorrect predictions. The custom CNN should be compared with MobileNetV2 under the same testing conditions.

If the model performs well on the benchmark dataset but less effectively on field images, possible explanations include differences in background, lighting, image quality, symptom severity, and crop variety. Such findings should be discussed using evidence from the actual experiments.

The model should also be assessed for its ability to recognize unsupported inputs. A classifier trained on a fixed set of classes may incorrectly assign an unfamiliar disease or non-leaf image to one of its known categories. Future implementations should consider input validation and uncertainty handling.

7. Conclusion

This paper proposes a deep learning-based plant disease detection framework using Convolutional Neural Networks. The proposed approach includes dataset preparation, image preprocessing, CNN training, disease classification, performance evaluation, and user-interface development. The framework can be implemented using TensorFlow or Keras, with a custom CNN serving as the baseline and MobileNetV2 providing a transfer-learning comparison. Evaluation should include accuracy, precision, recall, F1-score, and a confusion matrix.

Future work may investigate field-image validation, explainable AI techniques such as Grad-CAM, lightweight mobile deployment, disease severity estimation, and support for regional languages. Actual experimental validation is essential before drawing conclusions about model effectiveness or suitability for agricultural use.

REFERENCES

1. Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. *Frontiers in Plant Science*, 7, 1419. <https://doi.org/10.3389/fpls.2016.01419>
2. Liu, J., & Wang, X. (2021). Plant diseases and pests detection based on deep learning: A review. *Plant Methods*, 17, 22. <https://doi.org/10.1186/s13007-021-00722-9>
3. Abade, A., Ferreira, P. A., & de Barros Vidal, F. (2021). Plant diseases recognition on images using convolutional neural networks: A systematic review. *Computers and Electronics in Agriculture*, 185, 106125. <https://doi.org/10.1016/j.compag.2021.106125>
4. 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>
5. Hughes, D. P., & Salathé, M. (2015). An open access repository of images on plant health to enable the development of mobile disease diagnostics through machine learning and crowdsourcing. *arXiv*. <https://arxiv.org/abs/1511.08060>