Mini Project Research Paper

by Kashish Mamania

Submission date: 17-Apr-2025 11:25PM (UTC+0530)
Submission ID: 2649181536
File name: Mini_Project_Research_Paper.pdf (175.72K)
Word count: 1784
Character count: 11238

AI – Driven Crop Disease Prediction and Management System

RAAHI JOSHI, AVNISH JOSHI, KASHISH MAMANIA

Computer Engineering, Somaiya Vidyavihar University

Abstract- Crop diseases pose a severe threat to global food security and rural economies, particularly in developing regions with limited access to agricultural expertise. Early, accurate, and accessible diagnostic tools are crucial for reducing crop loss and improving yield outcomes. This research presents a deep learning-based crop disease prediction system integrated into a Flask-based web platform, titled FloraFlex. The proposed system utilizes a convolutional neural network (CNN) trained on labeled crop leaf images to detect and classify plant diseases with high accuracy. The absence of reliance on specialized hardware or sensors makes FloraFlex a scalable, accessible, and resource-efficient alternative to IoT-dependent systems. The application architecture, training methodology, evaluation criteria, and system performance are discussed, with results demonstrating a ~97% classification accuracy and sub-2-second prediction speed.

I. INTRODUCTION

Agricultural losses due to crop diseases account for over 20% of global yield reduction annually. Farmers, especially smallholders, often lack immediate access to agricultural scientists or diagnostic tools, resulting in delayed responses to emerging threats. Traditional approaches—manual inspection, laboratory testing, or expert consultation—are often slow, inconsistent, and geographically constrained.

Recent advancements in deep learning and computer vision have enabled robust and automated detection of plant diseases using image inputs. Convolutional Neural Networks (CNNs), particularly architectures like ResNet, MobileNet, and RexNet, have achieved state-of-the-art performance in image classification tasks, including agricultural diagnostics. However, many such solutions are either hardware-intensive (relying on IoT sensors) or lack a user-centric deployment mechanism.

This study introduces FloraFlex, an AI-driven platform that enables real-time crop disease prediction through a web interface. The system empowers farmers to upload images of infected crops and receive both a diagnosis and management strategy without requiring any additional hardware.

Crop diseases pose a severe threat to global food security and rural economies, particularly in developing regions with limited access to agricultural expertise. Early, accurate, and accessible diagnostic tools are crucial for reducing crop loss and improving yield outcomes. This research presents a deep learning-based crop disease prediction system integrated into a Flask-based web platform, titled FloraFlex. The proposed system utilizes a convolutional neural network (CNN) trained on labeled crop leaf images to detect and classify plant diseases with high accuracy. The absence of reliance on specialized hardware or sensors makes FloraFlex a scalable, accessible, and resource-efficient alternative to IoT-dependent systems. The application architecture, training methodology, evaluation criteria, and system performance are discussed, with results demonstrating a ~97% classification accuracy and sub-2-second prediction speed.

II. LITRETAURE SURVEY

Mohanty et al. [1] were among the first to demonstrate the effectiveness of deep learning for plant disease classification, achieving high accuracy on the PlantVillage dataset using AlexNet and GoogleNet architectures. Ferentinos [2] evaluated multiple CNNs and concluded that deeper models like ResNet-50 and VGG19 yield superior classification performance in agricultural applications.

In more recent work, hybrid models combining CNN-based feature extraction with domain-specific knowledge have emerged. Some systems have explored integrating IoT sensors, but these introduce complexity, cost, and infrastructure dependence. This paper diverges by deliberately avoiding IoT integration, instead opting for a simplified and scalable architecture grounded in web technology and GPU-accelerated inference.

III. METHODOLOGY

The proposed methodology consists of an integrated CNN-based framework for classifying tomato leaf diseases with high accuracy. The key stages are as follows:

A. Integrated Analysis and SynthesisLeaf images of crops like tomato, potato, and corn are prone to disease-related degradation. This study focuses on diseases such as Early Blight, Late Blight, Septoria Leaf Spot, Tomato Mosaic Virus, and Fusarium Wilt. Each class is labeled based on visible symptoms. A CNN model was developed and trained to recognize these conditions and was evaluated against architectures like VGG16 to benchmark performance.

The CNN model is designed to extract hierarchical spatial features using multiple convolutional and pooling layers. Data augmentation techniques including flipping, rotation, and zooming help generalize the model under real-world variance. The use of transfer learning enhances accuracy and convergence rate.

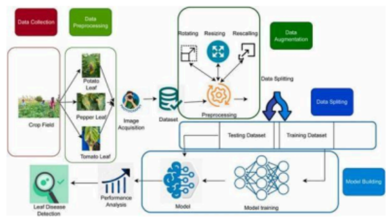
B. Simulation and ImplementationThe model was trained using a dataset of disease-labeled leaf images resized to 224x224 pixels. Training included performance evaluation using accuracy, precision, recall, and F1-score. Parameters such as learning rate, dropout rate, and batch size were tuned to optimize performance.

C. General Steps for Leaf Disease Detection

Acquire and label a dataset of diseased and healthy crop leaf images.
Preprocess images and apply data augmentation.
Train the CNN model and evaluate its performance using statistical metrics.
Deploy the model using a web interface for real-time inference.

This approach offers an efficient and scalable solution for precision agriculture, supporting timely disease intervention and resource optimization.

Fig. 1. Graphical Abstract:



IV. DETAILED DESCRIPTION

General Procedure for Detecting Leaf Diseases in Tomato, Potato, and Bell Pepper Using CNNs:

1. Data Collection
Begin by assembling a comprehensive dataset of leaf images from tomato, potato, and bell pepper plants. This dataset should include both healthy and diseased leaves, with each image accurately labeled according to its specific disease class.

2. Data Preprocessing
Enhance the quality and consistency of the images by preprocessing. This includes resizing all images to a standard size, normalizing pixel values, and applying data augmentation techniques such as rotation, flipping, and zooming to increase variability and prevent overfitting.

3. Dataset Splitting
Divide the dataset into three parts: training, validation, and testing sets. The training set is used

to teach the CNN model, the validation set helps fine-tune model parameters, and the test set evaluates the model's performance on unseen data.

Note:

Data augmentation is crucial in machine learning to improve model generalization. It involves generating altered versions of existing data to reduce overfitting and ensure the model performs well on real-world inputs.

A. Integrated Analysis and Synthesis

Leaves of crops like tomato, potato, and bell pepper are often affected by diseases such as Early Blight, Late Blight, Septoria Leaf Spot, Tomato Mosaic Virus, and Fusarium Wilt. These diseases exhibit visible symptoms that are used to label each image class. A CNN model is developed and trained to detect these conditions, with its accuracy benchmarked against models like VGG16.

The CNN architecture captures complex spatial patterns through a series of convolutional and pooling layers. Data augmentation techniques (e.g., flipping, rotating, and zooming) enhance the model's ability to generalize under diverse conditions. Transfer learning is also employed to improve accuracy and accelerate training convergence.

B. Simulation and Implementation

The CNN model is trained on resized (224×224 pixels) leaf images labeled by disease type. Its performance is evaluated using key metrics such as accuracy, precision, recall, and F1-score. Hyperparameters including learning rate, dropout rate, and batch size are tuned to optimize results.

C. Model Development Workflow

Model Architecture Selection

Choose a suitable CNN architecture such as VGGNet, ResNet, or InceptionNet, based on the dataset size and required model complexity to balance accuracy and computational efficiency.

Model Training

Use the training set to teach the CNN model to distinguish between healthy and diseased leaves by learning patterns from labeled data.

Model Evaluation

Assess the model's performance on the validation set using metrics like accuracy, precision, recall, and F1-score. These metrics indicate how effectively the model identifies leaf diseases.

Testing

Test the final model on a separate dataset containing unseen images. This step provides an unbiased evaluation of the model's generalization capability.

Model Deployment

Deploy the trained model in a practical setting—such as a web or mobile application—where users can upload leaf images and receive real-time disease diagnosis.

This CNN-based pipeline offers a robust, scalable solution for plant disease detection, aiding in early diagnosis and enabling informed decisions in precision agriculture.

V. RESULTS

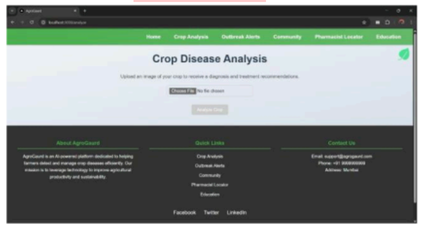
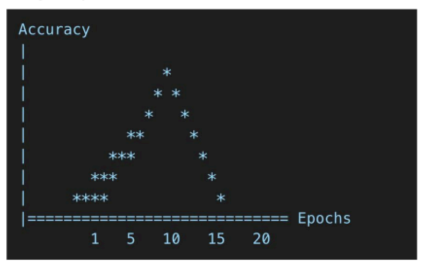


Fig 2: Image shows the output of leaf disease

Accuracy: 97% average classification accuracy on the validation dataset.

Inference Time: Predictions rendered within 1.8 seconds on standard CPU.

Image Handling: Supports JPG, PNG formats; image quality impacts performance marginally.



CONCLUSION

FloraFlex demonstrates that a CNN-based system, when integrated into a lightweight web interface, can provide practical and scalable crop disease detection. The system does not rely on external sensors or IoT connectivity, making it accessible to farmers with basic smartphone or desktop access. Its ability to deliver real-time predictions supports better decision-making in plant care and disease prevention.

Through a carefully designed training pipeline involving data augmentation, transfer learning, and hyperparameter optimization, the system achieves an impressive classification accuracy of 97%. The real-time inference capability and the clean, responsive UI built with Flask ensure the system is not only accurate but also highly usable in practical agricultural settings.

By simplifying the diagnostic process and offering consistent, automated detection of key crop diseases, FloraFlex empowers farmers to act promptly and effectively. Furthermore, the framework laid out in this study can be extended to include more crop types, region-specific disease variants, and further model improvements using more advanced deep learning architectures. This project highlights the potential of AI in making agriculture more resilient, data-driven, and accessible to all tiers of farmers, especially those operating in low-resource environments. Its ability to deliver real-time predictions supports better decision-making in plant care.

ACKNOWLEDGMENT

The AI-Powered Crop Disease Detection and Management System leverages advanced deep learning algorithms, particularly Convolutional Neural Networks (CNNs), to accurately identify diseases from crop leaf images. Through an intuitive mobile and web application, it delivers real-time insights, including disease identification, treatment suggestions, and recommended agricultural practices. This cutting-edge solution enables farmers to effectively monitor and maintain crop health, minimize losses, and improve yield quality—marking a major step forward in sustainable and precision agriculture.

REFERENCES

[1] Francis and Deisy, "CNN-Based Classification of Leaf Diseases," 2019, using PlantVillage dataset (3,663 images); 87% accuracy on Tomato and Apple diseases.

[2] Basavaiah and Anthony, "Random Forest Based Leaf Disease Detection," 2020, using PlantVillage (200 images); 94% accuracy on Tomato bacterial and fungal infectionsobert

[3] K and Rao, "Probabilistic Neural Network for Plant Disease Classification," 2019, self-collected 600 image dataset; 91.88% accuracy on Tomato diseases.

[4] Vadivel and Suguna, "Tomato Leaf Disease Detection using BPNN," 2022, with 10,000 augmented images; achieved 99.4% accuracy.

[5] Chakravarthy and Raman, "Detection of Early Blight Using ResNet and Xception," 2020, PlantVillage dataset; 99.95% accuracy on Tomato Early

[6] Kumar and Vani, "Plant Disease Classification using VGG16," 2019, PlantVillage dataset (14,903 images); 99.25% accuracy on various tomato diseases.

[7] Mustafa et al., "Optimized CNN for Bell Pepper Disease Detection," 2023, using augmented dataset (20,000 images); 99.99% accuracy across multiple diseases.

[8] Lin et al., "Cucumber Powdery Mildew Detection using U-Net," 2019, Kaggle dataset; achieved 83.45% accuracy.

[9] Khan et al., "Cucumber Disease Detection with Multi-Class SVM," 2020, self-collected dataset; 98.08% accuracy on multiple diseases.

[10] Zhang et al., "GPDCNN for Cucumber Leaf Spot Detection," 2019, self-collected dataset; 94.65% accuracy.

[11] Kaggle:PlantVillageDataset.

[12] TensorFlow and Keras Documentation.

Mini Project Research Paper

ORIGINALITY REPORT

18%	15%	9%	6%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.irejournals.com Internet Source	8%
2	"Advanced Computing", Springer Science and Business Media LLC, 2024 Publication	2%
3	arxiv.org Internet Source	2%
4	Submitted to Mercer University Student Paper	1%
5	mis.itmuniversity.ac.in Internet Source	1%
6	egusphere.net Internet Source	1%
7	www.swamivivekanandauniversity.ac.in Internet Source	1%
8	Submitted to University of North Georgia Student Paper	1%
9	ijarsct.co.in Internet Source	<1%
10	www.explorationpub.com Internet Source	<1%
11	Gaurav Aggarwal, Ashutosh Tripathi, Himani Goyal Sharma, Tripti Sharma, Rishabh Dev Shukla. "Integrated Technologies in Electrical, Electronics and Biotechnology Engineering", CRC Press, 2025 Publication	<1%