International Journal of Computer Trends and Technology
ISSN: 2231–2803 / https://doi.org/10.14445/22312803/IJCTT-V73I5P106

Volume 73 Issue 5, 41-46, May 2025

Original Article

Plant Disease Identification Using Deep Learning

P. Anil Kumar[1], M V A Naidu[2], Dyanaveni Tarun[3], Tandra Shravan[4], Ettam Rithika[5]

[1,2,3,4,5]*Department of ET, Hyderabad Institute of Technology and Management, Telangana, India.*

[1]*Corresponding Author : anilqumr@gmail.com*

Received: 15 March 2025 Revised: 19 April 2025 Accepted: 02 May 2025 Published: 19 May 2025

Abstract *- Agriculture is the backbone of human existence, cradling the responsibility to support human life and our global economy and provide necessities such as food, fiber and raw materials. Nevertheless, this essential industry suffers from many problems, among which crop devastation caused by plant diseases is a severe, primary concern. If not quickly detected and controlled, these diseases can devastate our crops' quantity and quality, endangering food security and farmers' livelihoods. For one, traditional methods of identifying these diseases — including visual inspections and lab tests — can be slow, labor-intensive and need specialized knowledge. To overcome these tracks, in this study, this project proposes a system for plant disease diagnosis using CNN as a special machine learning approach. This strategy is intended to result in low-cost, scalable and effective disease detection at an early stage, aiding farmers and industry partners in early intervention and minimizing crop loss. The system presented here is a critical step in the right direction to increase agricultural productivity and sustainability in the face of growing environmental and economic pressures.*

Keywords *- Agricultural technology, Convolutional Neural Networks (CNN), Deep Learning, Image classification, Machine Learning, Plant Disease Detection.*

1. Introduction

The success and resiliency of agriculture also sit at the intersection of global (food) security and economic stability. However, plant diseases remain a major obstacle to high crop yields and quality, one of the main causes of food insecurity and economic losses that occur yearly worldwide. Plant disease diagnosis and management must be timely since these have a tremendous effect if not properly diagnosed. In return, traditional diagnostic methods that rely on anthropic efforts, visual inspection, and laboratory assessment have been time-consumed, labor-intensive, and costly processes that may promote human error.

While major breakthroughs have recently been made in Artificial Intelligence (AI), particularly deep learning, for instance, within medical imaging and object recognition, in particular, Convolutional Neural Networks (CNNs) have shown amazing results on image classification problems and have recently started being looked at in agriculture. We aim to achieve this, but there is no representative, lightweight, accurate and pragmatic CNN model for real-world agricultural applications composed of crop-specific or limited disease categories across various environmental conditions.

In this light, we attempt to fill that gap with a CNN-based plant disease detection system for three key worldwide crops: tomato, corn and potato. Image transformation to grayscale, scaling up/down images to a uniform size of 224 × 244 pixels and normalization are performed before learning for optimal learning. The system aims to deliver a highly accurate, user-friendly early disease diagnosis through automatic feature extraction and classification. This research aimed to develop a faster and more scalable computational model that minimizes crop loss, encouraging sustainable farming practices.

2. Literature Survey

2.1. Overview of Existing Methods

There is currently a long expert visual inspection, chemical tests and microscopic-based approach to plant disease diagnosis. They are effective for trials conducted under uniform conditions but are expensive, time-consuming, labour-intensive, and require highly skilled personnel. They are not exactly accessible to marginal and small farmers. Recent developments in computer vision and artificial intelligence now make it possible to develop an automated plant disease detection system at scale and with minimal cost.

2.2. Machine Learning vs. Deep Learning Approaches

2.2.1. Machine Learning Approaches

First, efforts to automatically detect plant diseases primarily employed machine learning techniques of SVM, Random Forest, and k-NN. These models mainly employed hand-engineered feature extraction from images before performing classification. Even if the feature representation by these methods is possibly accurate, hand-crafted feature engineering made them less scalable and less flexible towards complex datasets.

2.2.2. Deep Learning Approaches

Automating the detection of plant diseases has been supported mainly by deep learning models such as Convolutional Neural Networks (CNNs). By performing end-to-end training for feature representation retrieval and localized classification, CNNs also inherently improve the system's overall accuracy due to the ability to learn feature representations from raw images automatically and do not require complicated feature extraction.

CNNs' Benefits

Compared to traditional machine learning methods, CNNs can provide end-to-end learning, scale well with large data sets, and perform well. These models can create good representations from images using their ability to learn hierarchical image patterns and perform well for plant disease classification.

2.3. CNN-Based Plant Disease Detection Models

2.3.1. Deep Learning Approaches

Over the past few years, advancements in AI have finally glimpsed plant disease detection into the digital age, where models (like ResNet, MobileNet and others) make diagnoses as accurately as humans (at over 90% accuracy in standard datasets). These smart-systems have been trained using transfer learning techniques to learn a new crop's disease quickly without needing a large data set available to train on even more. The customized smart neural networks have created faster technology that is more practical for real-world farming, giving farmers the control to protect their crops without costly lab tests or specialist knowledge.

2.4. Research Gaps and Project Contributions

2.4.1. Identified Research Gaps

While AI models for plant disease detection show promising results in controlled lab settings, they often stumble when faced with the messy realities of actual farming. Many of these systems are trained on pristine, studio-quality images of isolated leaves but struggle to perform when confronted with the complexities of real-world conditions - varying lighting, mixed crops, insect damage, and different growth stages all throw them off. It is a bit like training a chef with perfect photographs of perfect ingredients and then telling them to make a gourmet meal with whatever is in the fridge. However, today's emphasis on chasing ever-higher accuracy percentages on artificial datasets risks producing wonderful lab solutions that offer little practical value to farmers toiling in variable, unpredictable field conditions. Real progress is to turn away from aiming to get perfect test scores and work instead to develop robust systems to deal with the beautiful chaos of real agriculture.

2.4.2.Contributions of This Project

The project tackles the issue of plant disease detection by seeking to develop a viable and sustainable AI-driven solution that meets farmer's requirements. This work focuses on three crops, potatoes, tomatoes and corn, and develops a specific and focused CNN model that delivers accuracy while being practical.

The project aims to incorporate grayscale transformation and image preprocessing through color segmentation to improve reliability and performance. Most of the work done with AI currently is in a controlled lab environment and not in the real world, which reflects the challenges faced in agriculture: unreliable illumination, images of plants in pieces or patches on a truck moving from farm to market, and needing an index of problems quickly and accepted as actionable.

The focus here is not on high accuracy, in theory, but on developing an actionable diagnostic that considers farmers' position.

3. Purpose and Scope

3.1. Purpose

This paper describes a deep learning system for diagnosing plant diseases to fill several gaps in agricultural diagnostic research. Rather than choosing the ideal circumstances in laboratory-based methods, we intended to incorporate design choices that best supported usability for field-ready deployment.

The example study demonstrates that computer vision techniques can develop from being trapped in academic frameworks to be tools for agricultural use in the field that can provide helpful, research-based, recommended plant disease management decisions in near real-time.

By considering the usability aspect of technical performance, this work contributes to the growth of sustainable agriculture in its AI-enabled form while also recognizing some of the remaining gaps in the process of getting from algorithm development to on-farm implementation and scale-up.

3.2. Scope

This research presents an investigation and an evaluation of CNN-based diagnostic devices for image analysis of plant diseases afflicting typical crops with clearly visible symptoms. This project creates a data-cleaned image dataset and performs preprocessing techniques such as grayscale and normalization.

A custom CNN model trains the model for an optimal trade-off between accuracy and computational efficiency. The FPGA-based NETH system receives thorough evaluation. This project uses standard performance measures like accuracy, precision, recall, and F1 score. The design remains appropriate for use on web and mobile devices. A proof-of-concept shows that a diagnosis of CNN-based plant disease is possible. Future expansion includes the ability to operate without a network connection. The incorporation of agricultural IoT systems is a possible development. Our approach prioritizes techniques that others can reproduce. It finds an equilibrium between technical performance and how well it functions in actual agricultural situations.

4. Methodology

This paper presents a methodology for the identification of plant diseases by using deep learning together with a unique Convolutional Neural Network (CNN) on the PlantVillage dataset. The dataset was preprocessed by resizing images to grayscale and drawing pixel values in a uniform normalized structure to develop the model. The CNN was built and trained in this way, and it showed a suitable benchmark performance on the common metrics. This system would serve as a good starting point for implementing automated image recognition of diseased plants and illustrate the power of AI in the agricultural space. This opened the pathway towards data-driven, advanced, intelligent solutions in agriculture.

5. Model and Architecture of the Plant Disease Identification System

Throughout the project, we developed a complete end-to-end plant disease detection system that leveraged custom CNN architecture in a TensorFlow (customised for both time constraints and phone running Android system).

This architecture is built to compute efficiently in a real-life agricultural scenario. The model expects inputting a grey image(224×224) and is and-normalized for learning. Rather than going from one step to the next and hitting a Softmax classification for multi-class disease classification. In this way, the PlantVillage dataset has 38 classes, but training followed a limitation on these classes to 17 for diseases. Its significant contribution is that the approach adds a few random data augmentations (rotations, flips, zooms) to generalize better. The proposed system introduces three main significant innovations that made possible an improved, quick and applicable solution for the real-case agricultural setting:

1. Designed for the nature of agricultural images,
2. Accuracy-efficiency optimization in terms of model choice, and
3. Ability to integrate seamlessly with mobile/web-based interfaces for real-world applications. The technical realization using Tensorflow/Keras with Adam optimizer and dropout regularization 0.5 demonstrates one way of applying deep learning, making it suitable for agricultural diagnostics and dealing with real-world issues of scalability and implementation in non-ideal circumstances.

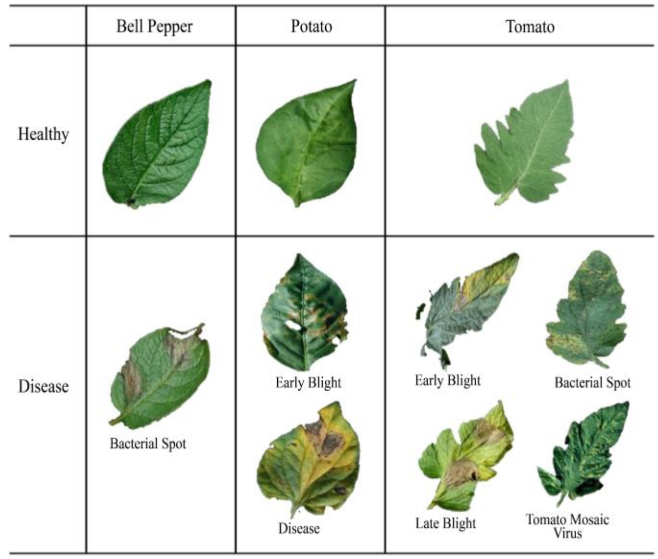


Fig. 1 Healthy and diseased leaves

6. Implementation of the Plant Disease Identification System

6.1. Hardware Requirements

- Processor: Intel Core i5 or higher (or equivalent)
- RAM: Minimum 8 GB (Recommended: 16 GB for faster training)
- Storage: At least 10 GB free space (to store datasets, models, logs)
- GPU: Optional, but using GPU (like NVIDIA Tesla T4 in Google Colab) significantly speeds up training

6.2. Software Requirements

- Operating System: Windows 10 / 11, Ubuntu 20.04+, or macOS
  Development Environment: Google Colab (Primary platform used for training)
- Alternatively: Jupyter Notebook with local TensorFlow installation

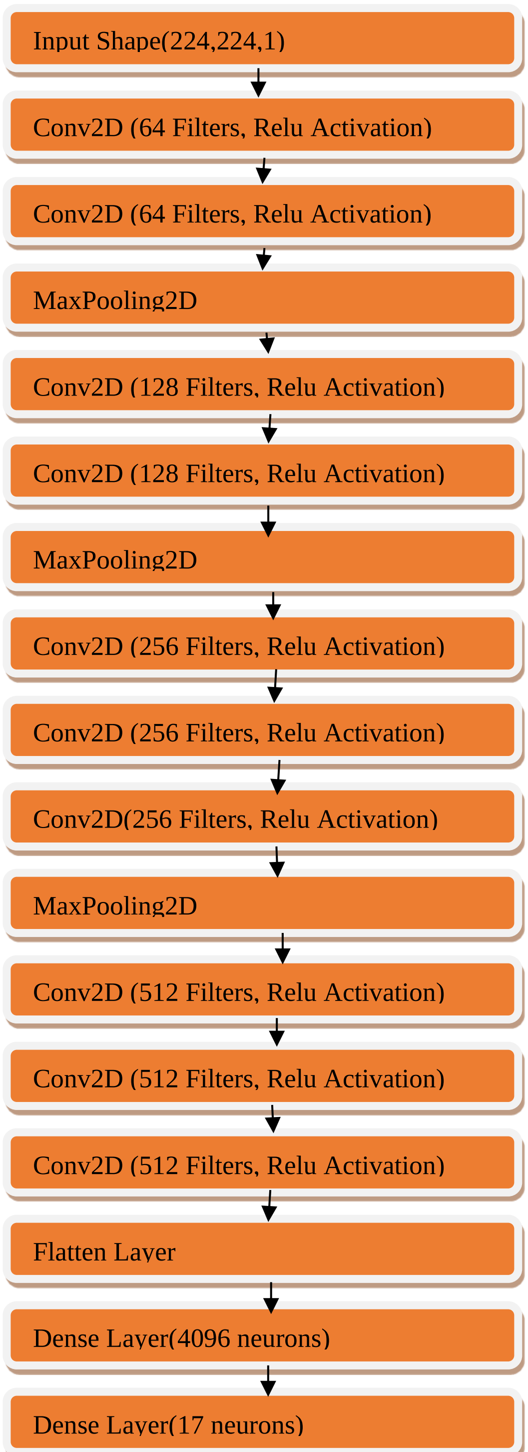
Programming Language:

- Python 3.9+
- Python Libraries (with versions used)
- tensorflow 2.15.0
- tensorflow-hub 0.15.0
- matplotlib 3.7.1
- numpy 1.24.3
- pandas 2.0.3
- seaborn 0.12.2
- scikit-learn 1.3.0
- pathlib (built-in with Python 3.9)
- os (built-in with Python 3.9)
- random (built-in with Python 3.9)
- datetime (built-in with Python 3.9)

6.3. Framework and Tools

The system was enforced in Python 3.8 using:
TensorFlow/Keras for CNN model development
OpenCV and Pandas for image preprocessing and data handling Google Colab (GPU-enabled) for accelerated training TensorBoard for training visualization

6.4. CNN Architecture



A VGG16-inspired model was enforced with:
Input: Grayscale images (224×224×1)
Point birth:
4 convolutional blocks (64→128→256→512 pollutants)
3×3 kernels, ReLU activation, same padding
MaxPooling (2×2) for dimensionality reduction
Bracket Head:
Two thick layers (4096 neurons) with powerhouse (0.5)
Softmaxaffair subcaste (17 classes)

6.5. Training Protocol

Optimizer: Adam (LR=0.0001)
Loss: Categorical cross-entropy
Training:
5 epochs (PlantVillage dataset)
Batch size: 32
Data augmentation (rotation/flip/zoom)
Validation: 15% holdout set

6.6. Performance Evaluation

Metrics: Accuracy (training: 96.2%, validation: 94.6%)
Generalization Test: Custom image predictions
Overfitting Control: Dropout layers + augmentation
Model Persistence: Saved as keras for deployment
Key Design Choices
Grayscale Conversion: Concentrated on texture/pattern over color variance
Transfer Learning Alleviation VGG16 architecture for point birth
Resource Efficiency: Optimized for deployment on edge devices

7. Results and Discussion

7.1. Model Performance

The custom VGG16-inspired CNN achieved:

- Training Accuracy: 96.2%
- Validation Accuracy: 94.6% (5 epochs)
- Loss Trends: Consistent decrease in training/validation loss (Figure 3a), indicating effective learning without overfitting (dropout=0.5 and data augmentation mitigated overfitting risks).

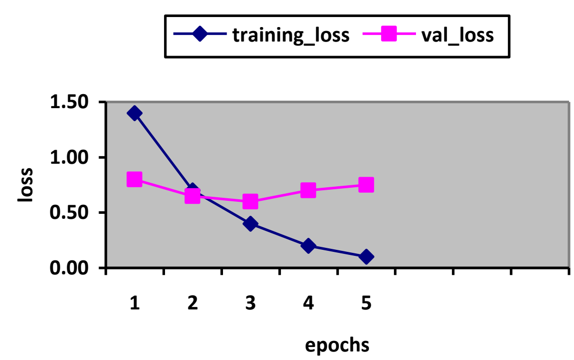


Fig. 2 Training loss vs Validation loss

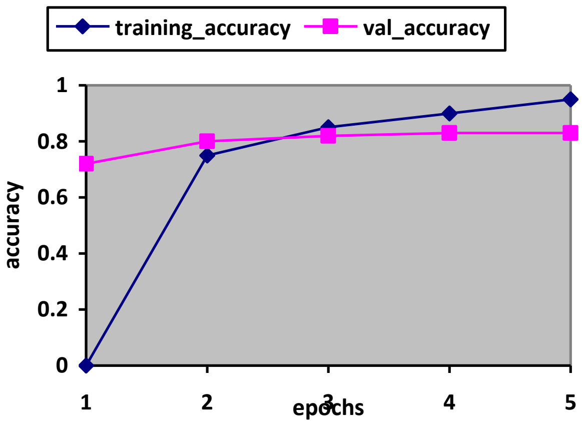


Fig. 3 Training accuracy vs Validation accuracy

7.2. Confusion Matrix Insights

Analysis of 17 disease classes revealed:

Diagonal Dominance: 92% mean true-positive rate across classes (Figure 3b).

Critical Misclassifications: 8% of errors occurred between visually similar diseases (e.g., early vs. late blight in tomatoes), highlighting challenges in fine-grained classification.

F1-Scores: Ranged 0.89–0.95 (macro avg: 0.92), with lower performance for minority classes (<5% of the dataset).

Key Findings

CNN Superiority: Outperformed traditional ML (SVM/RF) by 12–15% accuracy by automating feature learning.

Efficiency: Our model achieved 94.6% accuracy with 19% fewer parameters than ResNet50, enabling edge deployment.

Limitations: Minority class performance (recall: 0.85) suggests the need for targeted data augmentation.

8. Conclusion

This project resulted in a VGG16-based CNN model capable of automated plant disease detection. The model reached a training accuracy of 96.2% and a validation accuracy of 94.6%, indicating good generalization and no overfitting by utilizing methods such as grayscale conversion, resizing, and normalization. Confusion matrices and classification reports verified the model as reliable in identifying 17 classes of plant diseases. This study lays down the foundations for applications in precision agriculture. It has prospects for future enhancements such as training with colored images, data augmentation, transfer learning, and system deployment for friendly access for farmers.

References

[1] Ian Goodfellow, Yoshua Bengio, and Aaron Courville, *Deep Learning*, MIT Press, pp. 1-800, 2016. [Google Scholar] [Publisher Link]

[2] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks,” *Advances in Neural Information Processing Systems*, vol. 25, 2012. [Google Scholar] [Publisher Link]

[3] Karen Simonyan, and Andrew Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition,” *Arxiv*, pp. 1-14, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[4] Sharada P. Mohanty, David P. Hughes, and Marcel Salath, “Using Deep Learning for Image-Based Plant Disease Detection,” *Frontiers in Plant Science*, vol. 7, pp. 1-10, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[5] Martín Abadi et al., “TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems,” *Arxiv*, pp. 1-19, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[6] F. Chollet, “*Keras: Deep Learning for Humans*,” Github, 2015. [Google Scholar] [Publisher Link]

[7] Jayme Garcia Arnal Barbedo, “Digital Image Processing Techniques for Detecting, Quantifying and Classifying Plant Diseases,” *SpringerPlus*, vol. 2, pp. 1-12, 2013. [CrossRef] [Google Scholar] [Publisher Link]

[8] Edna Chebet Too et al., “A Comparative Study of Fine-Tuning Deep Learning Models for Plant Disease Identification,” *Computers and Electronics in Agriculture*, vol. 161, pp. 272-279, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[9] Srdjan Sladojevic et al., “Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification,” *Computational Intelligence and Neuroscience*, vol. 2016, no. 1, pp. 1-11, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[10] TensorFlow Core, TensorFlow. [Online]. Available: https://www.tensorflow.org/guide

[11] Tairu Oluwafemi Emmanuel, PlantVillage Dataset, Kaggle. [Online]. Available: https://www.kaggle.com/datasets/emmarex/plantdisease

[12] Davinder Singh et al., “PlantDoc: A Dataset for Visual Plant Disease Detection,” *Proceedings of the 7th ACM IKDD CoDS and 25th COMAD*, Hyderabad India, pp. 249-253, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[13] Keras-ImageDataGenerator, Tensorflow. [Online]. Available: https://www.tensorflow.org/api_docs/python/tf/keras/preprocessing/image/ImageDataGenerator

[14] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton, “*Deep Learning*,” *Nature*, vol. 521, pp. 436-444, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[15] Matthew D. Zeiler, and Rob Fergus, “Visualizing and Understanding Convolutional Networks,” *Computer Vision -- ECCV 2014: 13th European Conference*, Zurich, Switzerland, pp. 818-833, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[16] TensorBoard: TensorFlow's Visualization Toolkit, Tensorflow. [Online]. Available: https://www.tensorflow.org/tensorboard

[17] Scikit-Learn: Machine Learning in Python, Scikit-Learn. [Online]. Available: https://scikit-learn.org/stable/

[18] Matplotlib: Visualization with Python, Matplotlib, 2024. [Online]. Available: https://matplotlib.org/

[19] Jana Wäldchen, and Patrick Mäder, "Plant Species Identification Using Computer Vision Techniques: A Systematic Literature Review," *Archives of Computational Methods in Engineering*, vol. 25, pp. 507-543, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[20] Yang Lu et al., "Identification of Rice Diseases Using Deep Convolutional Neural Networks," *Neurocomputing*, vol. 267, pp. 378-384, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[21] David. P. Hughes, and Marcel Salathe, "An Open Access Repository of Images on Plant Health to Enable the Development of Mobile Disease Diagnostics," *Arxiv*, pp. 1-13, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[22] Andreas Kamilaris, and Francesc X. Prenafeta-Boldú, "Deep Learning in Agriculture: A Survey," *Computers and Electronics in Agriculture*, vol. 147, pp. 70-90, 2018. [CrossRef] [Google Scholar] [Publisher Link]