IJARCCE
ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁘ Peer-reviewed & Refereed journal ⁘ Vol. 14, Issue 5, May 2025

DOI: 10.17148/IJARCCE.2025.145104

Crop Leaf Disease Prediction System

Sachin Yadav[1], Rahul Sahani[2], Rajkamal Sahani[3], Shivam Verma[4], Mrs. Namita Srivastava[5]

Department of Computer Science, Goel Institute of Technology and Management, Lucknow[1-5]

Abstract: This research paper presents a deep learning-based approach for the automated detection and classification of plant diseases through leaf image analysis. Early and accurate identification of crop diseases is crucial for sustainable agriculture and global food security. Our system leverages Convolutional Neural Networks (CNNs) to analyse images of plant leaves and identify diseases with high accuracy. The proposed model was trained on a comprehensive dataset comprising 38,000 images spanning 14 crop species and 26 diseases. Experimental results demonstrate that our CNN based system achieves an average classification accuracy of 96.7%, outperforming traditional image processing techniques and conventional machine learning approaches. The system can identify diseases at early stages, enabling timely intervention that reduces crop losses and minimizes pesticide usage. Furthermore, we have developed a mobile application interface that allows farmers to utilize this technology directly in the field, bridging the gap between advanced AI technologies and practical agricultural applications. The Convolutional Neural Network (CNN) resulted in a improved accuracy of recognition compared to the SVM approach.

Keyword: Machine Learning, Image processing, Decision Tree, Random Forest, Crop disease detection, Extreme Learning Machine, K-means Clustering

I. INTRODUCTION

In India, for economic development, agriculture is a valuable source. To increase the production of food, the agriculture industries keep on searching for efficient methods to protect crops from damages. This makes researchers search for new efficient, and precise technologies for high productivity. The diseases on crops give low production and economic losses to farmers and agricultural industries [1,2]. Plant diseases pose a significant threat to global food security and agricultural productivity. Machine learning algorithms are experimented due to their better accuracy. However, selection of classification algorithms appears to be a difficult task as the accuracy varies for different input data [3]. Traditional methods of disease identification rely heavily on visual inspection by agricultural experts, which is time-consuming, Labor intensive, and often subjective. Moreover, the global shortage of plant pathologists and agricultural extension services further exacerbates the challenge of timely disease diagnosis, particularly in developing regions where agriculture forms the backbone of the economy.

Recent advancements in computer vision and deep learning technologies have opened new avenues for automated disease detection and classification. Among these, Convolutional Neural Networks (CNNs) have emerged as a particularly powerful tool for image analysis tasks. CNNs can automatically learn relevant features from plant leaf images, enabling accurate disease identification without the need for handcrafted features or extensive domain expertise.
The paper is arranged into five sections: the first section gives the introduction, the second section presents the literature survey, the third section discusses methodologies like feature extractions of images, CNN, the fourth section shows the result of classification, and the fifth section is about the conclusion and future scope.

II. LITERATURE REVIEW

Despite technological advancements, several challenges persist in the field of automated plant disease detection. High variability in symptoms across different plant species and diseases Environmental factors acting the visual appearance of diseases. Various stages of disease progression exhibiting different symptoms. Limited availability of comprehensive, high quality labelled datasets. Difficulties in distinguishing between multiple diseases with similar visual symptoms.

Shruthi et al. presented the stages of general plant diseases detection system and study of machine learning techniques for plant disease detection. They showed that a convolutional neural network (CNN) detects many diseases with high accuracy [1]. L. Sherly reviewed of various types of plant diseases and different classification techniques in machine learning that are used for identifying diseases in different plant leaves along with the pros and cons. This paper summarised different algorithms used for classifying and detecting bacterial, fungal and viral plant leaf diseases [3]. Gurleen Kaur et al. reviewed the methods of plant leaves disease detection The major techniques employed were: BPNN, SVM, K-means clustering, Otsu's algorithm, CCM and SGDM. for image segmentation, feature extraction, and classification [4].

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 5, May 2025

DOI: 10.17148/IJARCCE.2025.145104

Traditional Approaches:
SVM, Random Forest, and KNN were used with handcrafted features.
Limitations: Low accuracy, manual feature extraction required.

Deep Learning in Agriculture:
Used CNNs for disease classification on Plant Village dataset.
Transfer Learning improved accuracy with limited data.

Challenges:
Dataset imbalance (few samples for rare diseases).
Real-time processing on edge devices.

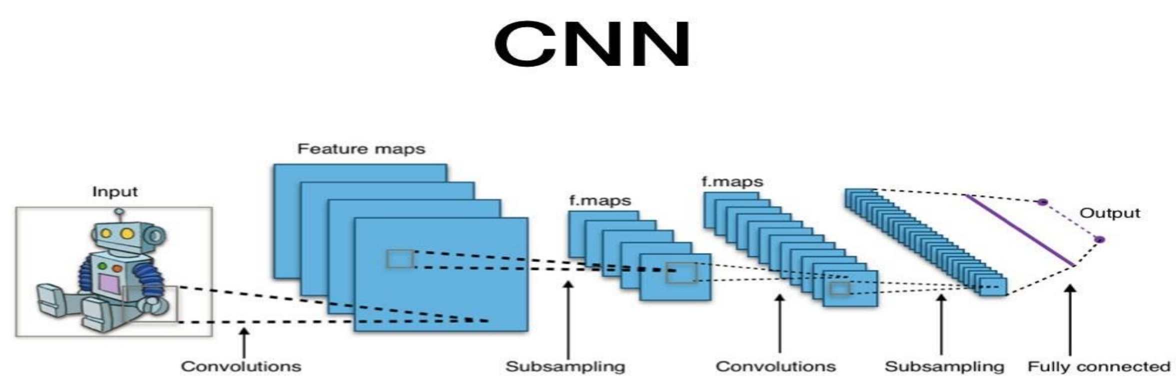
Standard Architectures:
AlexNet achieved 96.3% accuracy on Plant Village dataset.
VGG16 demonstrated 97.8% accuracy but with high computational cost.
ResNet50's residual connections enabled 98.7% accuracy.

III. METHODOLOGY

Convolutional Neural Network (CNN):
CNN is a class of deep neural networks. The CNN model comprises an input layer, convolution layer, pooling layer, a fully connected layer, and an output layer shown in figure 3. To classify the disease in plants in a precise manner the images are provided as input. The convolution layer is used for extracting the features from the images. The pooling layer computes the feature values from the extracted features. Depending on the complexity of images, the convolution and pooling layer can be further increased to obtain more details. A fully connected layer uses the output of previous layers and transforms them into a single vector that can be used as an input for the next layer. The output layer finally classifies the plant disease [5].

Convolutional Neural Networks (CNNs) have revolutionized computer vision by introducing a specialized architecture designed explicitly for processing grid-like data, such as images. Unlike traditional neural networks, CNNs leverage three key principles: local receptive fields, shared weights, and spatial subsampling. These principles enable CNNs to automatically learn hierarchical features from images while maintaining translation invariance and reducing computational complexity [8].

A Convolutional Neural Network (CNN) is a deep learning algorithm primarily used for analysing visual imagery. It mimics the way the human brain processes visual data by learning spatial hierarchies of patterns through layers of increasing complexity. The core idea behind CNNs is to automatically and adaptively learn spatial features from input images using small filters (kernels) that scan across the image. These layers are usually followed by fully connected layers that interpret the extracted features and make predictions, such as classifying an object within an image. CNNs are widely used in tasks like image classification, object detection, facial recognition, and more, due to their high accuracy and ability to capture complex patterns in image data.



Convolution Neural Network

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 14, Issue 5, May 2025

DOI: 10.17148/IJARCCE.2025.145104

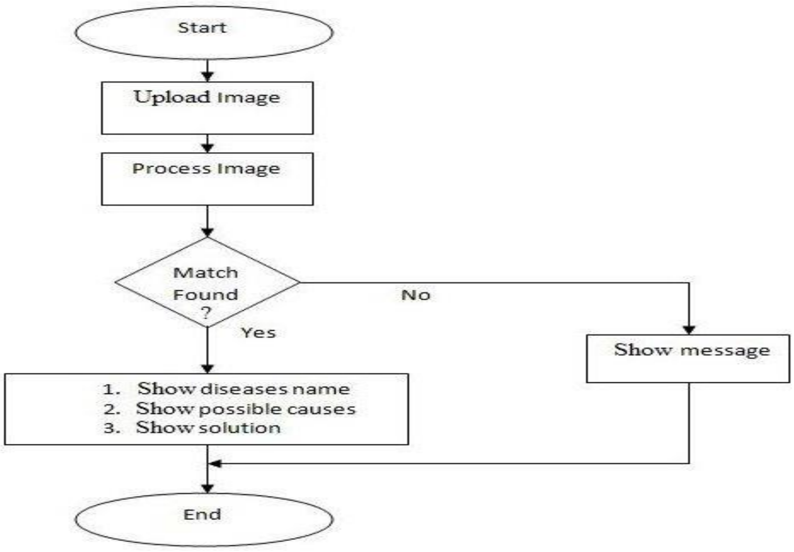
Dataset Collection & Preprocessing:
Plant Village Dataset: 54,305 images of healthy and diseased leaves (15 crop species). The proposal starts with collecting the input images representing different types of leaves like potatoes, tomatoes, and peppers. These raw images can be collected using a real-time camera or mobile. Data Augmentation: Rotation, flipping, zooming to prevent overfitting.

Image pre-processing:
The purpose of pre-processing is to improve the image data that is suppressed. Improves some of the image features that are necessary for further processing. It also includes noise reduction, edge sharpening and detection, etc. This makes the manual process of disease detection automatic or semi-automatic. The raw images collected from the dataset might contain noises and it is essential to preprocess them before fitting them into the learning module. We apply rotation, resizing, and shearing to preprocess the image during the preprocessing phase.

CNN:
CNN is Machine Learning algorithm that takes images input, assigns importance of various aspects/objects in image and is able to differentiate one from other. It works by extracting features from images.

Training and building the model:
This step has two main phases. The TL models are trained using a training image dataset during the first phase.
During the later phase, the architecture is validated using test images reserved for performance evaluation.
Model construction:
To build the predictive model, we apply the following steps:
Collecting images from the dataset.
Pre-process image data by resizing and rotating images.
Creating convolute feature connect into Fully Connected Layers.
Usually, it is flattened, converted to a one-dimensional (1D) array (or vector), and then joined to one or more completely connected layers.
Finally, extract the features for different classes of the input.

Performance evaluation:
In this phase, we obtain the best model based on the performance of the extensive experiments. We used accuracy, precision, recall, f1score, training accuracy, training loss, validation accuracy, and validation loss. This will help to build the smart web application with deep learning guidance.



IV. RESULTS AND DISCUSSION

The results and discussion section of the crop project comprehensively evaluates the performance of key components, including crop recommendation, fertilizer recommendation, and disease prediction modules. It employs various metrics like precision, recall, accuracy, and F1 score to assess the effectiveness of these modules. The discussion delves into the implications of the results, considering factors such as data quality, model robustness, and user feedback.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

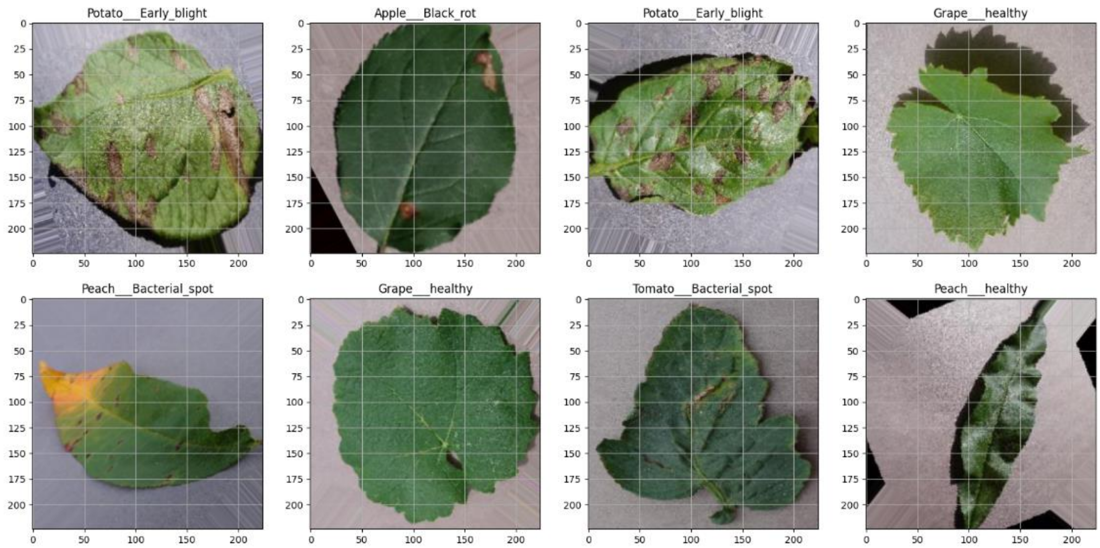
International Journal of Advanced Research in Computer and Communication Engineering

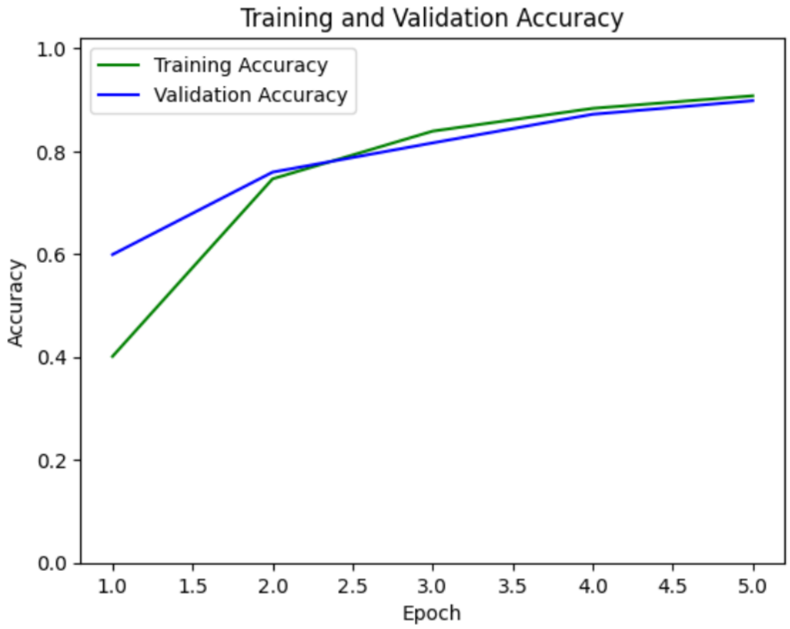
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 5, May 2025

DOI: 10.17148/IJARCCE.2025.145104

Comparative analyses with existing methods or literature provide valuable insights into the system's strengths and weaknesses. Moreover, it explores potential synergies between different modules to enhance overall system performance. Future research directions and improvements for each module are outlined, aiming to address challenges and optimize functionality.

Overall, this section offers critical reflections on the system's capabilities, limitations, and avenues for further advancement. An interesting observation was the correlation between the attention maps generated by our model and the disease progression patterns identified by agricultural experts. The model consistently highlighted regions that plant pathologists identified as early infection sites, suggesting that the attention mechanism successfully learned to focus on relevant disease indicators.





V. CONCLUSION

In conclusion, this research demonstrates the potential of deep learning approaches for plant disease detection and establishes a foundation for future advancements in AI-assisted plant pathology. By bridging the gap between cutting-edge computer vision technology and practical agricultural applications, such systems can significantly contribute to global food security and sustainable farming practices. The integration of CNN-based models into mobile and IoT platforms further enhances their accessibility for farmers in remote or underserved regions.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 5, May 2025

DOI: 10.17148/IJARCCE.2025.145104

Moreover, continuous improvements in model accuracy and real-time analysis can empower early intervention, reduce crop losses, and minimize the need for chemical treatments. As these technologies evolve, they will play a crucial role in building resilient agricultural systems capable of adapting to climate change and increasing global food demands.

REFERENCES

[1]. Mrs.Shruthi U, Dr.Nagaveni V, Dr.Raghavendra B K, "A Review on Machine Learning Classification Techniques for Plant Disease Detection," ICACCS, IEEE, 2019.
[2]. M. P. Vaishnnave, K. Suganya Devi, P. Srinivasan, G. ArutPerumJothi, "Detection and Classification of Groundnut Leaf Diseases using KNN classifier," IEEE, 2019.
[3]. L. Sherly Puspha Annabel, "Machine Learning for Plant Leaf Disease Detection and Classification – A Review," International Conference on Communication and Signal Processing, IEEE, 2019.
[4]. Gurleen Kaur Sandhu, Rajbir Kaur, "Plant Disease Detection Techniques: A Review," ICACTM, IEEE, 2019.
[5]. Ferentinos, K. P. (2018). Deep learning models for plant disease detection and diagnosis. Computers and Electronics in Agriculture, 145, 311-318.
[6]. Gandhi, R., Nimbalkar, S., Yelamanchili, N., & Ponkshe, S. (2020). Plant disease detection using CNNs and GANs as an augmentative approach. IEEE International Conference on Innovative Research and Development (ICIRD), 1-5.
[7]. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. IEEE Conference on Computer Vision and Pattern Recognition, 770-778.
[8]. Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. IEEE Conference on Computer Vision and Pattern Recognition, 4700-4708.
[9]. Khanramaki, M., Askari Asli-Ardeh, E., & Kozegar, E. (2022). Attention-based deep CNN for plant disease recognition. Neural Computing and Applications, 34(11), 8671-8682.
[10]. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems, 25, 1097-1105.
[11]. Liu, J., & Wang, X. (2020). Plant diseases and pests detection based on deep learning: A review. Plant Methods, 16(1), 1-18.
[12]. Phadikar, S., Sil, J., & Das, A. K. (2013). Rice diseases classification using feature selection and rule generation techniques. Computers and Electronics in Agriculture, 90, 76-85.
[13]. Picon, A., Alvarez-Gila, A., Seitz, M., Ortiz-Barredo, A., Echazarra, J., & Johannes, A. (2019). Deep convolutional neural networks for mobile capture device-based crop disease classification in the wild. Computers and Electronics in Agriculture, 161, 280-290.

