Classification and Detection of Plant Disease Using CNN and Machine Learning

Nischitha M Upadhya [1], Prof Dwarakanath G V [2]

[1]Student, Department Of Masters Of Computer Application, BMS Institute Of Technology And Management ,Bangalore, Karnataka, India

[2]Assistent Professor, Department Of Masters Of Computer Application, BMS Institute Of Technology And Management, Bangalore, Karnataka, India

---***---

Abstract - The detection and classification of plant diseases are crucial to ensuring food security and maximizing agricultural productivity. Traditional methods of plant disease identification are time-consuming and labor-intensive, necessitating the adoption of more efficient and accurate techniques. In recent years, advancements in machine learning have led to the development of robust methods for automated plant disease detection. This research paper presents a novel approach for plant disease detection using Convolutional Neural Networks (CNNs). CNNs have demonstrated exceptional performance in image recognition tasks, making them a promising choice for detecting diseases in plant images. The proposed system utilizes a pre- processed dataset of plant images, comprising both healthy and diseased samples, to train the CNN model.

Keywords—Plant disease and detection, convolutional neural networks, deep learning architectures, feature extraction, classifier methods

I. INTRODUCTION

Plant diseases have a profound impact on agricultural productivity and global food security. Timely and accurate detection of these diseases is critical for effective disease management and the prevention of extensive crop losses. Traditional methods of disease detection, such as visual inspection by experts, are subjective, time-consuming, and often inadequate for handling large-scale agricultural operations. To address these challenges, researchers have turned to cutting-edge technologies like machine learning and CNNs to develop automated and efficient plant disease detection systems. Machine learning algorithms, particularly deep learning methods like CNNs, have revolutionized image recognition tasks and have shown remarkable success in various fields, including computer vision and natural language processing. By leveraging their ability to automatically learn hierarchical patterns from data, CNNs have proven to be highly effective in detecting objects and patterns in images. The proposed research aims to explore and develop a plant disease detection system using machine learning, with a focus on CNNs. The primary objective is to design an automated model capable of identifying plant diseases accurately and efficiently, enabling timely intervention by farmers and agronomists.

This research builds upon previous work in classification and detection plant disease using machine learning techniques. While early attempts utilized traditional image processing algorithms, they often lacked the capability to generalize well to new and diverse datasets. CNNs and , have the advantage of learning complex features directly from raw pixel data, making them well-suited for the task of plant disease detection. The key components of the proposed system include data collection, pre-processing, model training, and evaluation. To train the CNN model, a dataset comprising images of healthy plants and various diseased conditions will be collected from different sources and properly labeled. Data augmentation techniques will be employed to increase the training data and enhance the model's ability to generalize to unseen examples. The CNN model will be developed using a transfer learning approach, where a pre-trained model on a large dataset (e.g., ImageNet) will be fine-tuned for the specific task of plant disease detection. This approach significantly reduces the need for a large labeled dataset and allows the model to leverage the knowledge gained from learning general image features. The effectiveness of the proposed plant disease detection system will be evaluated using metrics .The model's performance will be compared against traditional methods and other machine learning techniques to demonstrate its superiority. The potential applications of the developed system in real-world agricultural scenarios will also be discussed. By enabling farmers and agronomists to detect and manage plant diseases early on, this technology can lead to increased crop yields, reduced use of pesticides, and ultimately contribute to sustainable and efficient agricultural practices.

II. RELATED WORK

A Crop Disease Dataset for Machine Learning: The Plant Village dataset is a sizable collection of plant photos that have been annotated with various illnesses. It was introduced in this paper. The accurately classifying and detected plant diseases, the authors successfully apply CNNs to this dataset, making it an invaluable tool for academics tackling related tasks.[1] Using camera images, deep learning-based large-scale automatic plant disease diagnostics: This study investigates how CNNs can be used to identify plant diseases using photographs taken with smartphones. The authors provide a

deep learning-based architecture that is highly accurate in detecting numerous illnesses across various plant species.[4] Convolutional Neural Networks for Plant Disease Identification: This study explores the use of CNNs for plant disease detection that can precisely categories healthy and diseased plant leaves. The study compares different Machine Learning architectures and highlights the significance of dataset quality and size for achieving better performance.
Deep Learning Approaches for Plant Disease Detection and Diagnosis: Various deep learning techniques, including CNNs, are discussed in this review paper as they are used to the identification and diagnosis of plant diseases. It gives a summary of the most recent approaches, datasets, and problems

encountered in the field while showing the promise of deep learning to enhance crop management.

Deep Learning for Plant Disease Detection: This comprehensive review paper discusses the advancements in the use of deep learning, particularly CNNs, for the detection of plant diseases. The study examines several areas of the research, such as dataset development, model designs, and performance evaluation, as well as future opportunities and challenges.

Review of Deep Learning Methods Used for Semantic Segmentation and Agricultural Disease Detection: This review study focuses on the application of deep learning techniques, such as CNNs, for semantic segmentation and illness diagnosis in agriculture. It provides insights into current trends, contrasts alternative methodologies, and discusses how these tactics could be applied in practical settings.

III. PROPOSED METHOD

Data Gathering and Preprocessing: Compile a large database of plant photos that includes both healthy plants and plants that have been afflicted by different diseases. The dataset needs to include a variety of plant species and disease types. For supervised learning, annotate and label the photographs with the appropriate illness categories. Standardizing the photos' size, format, and color space during preprocessing. Utilize picture enhancing methods to lower noise and raise the dataset's quality.

Data augmentation techniques like rotation, flipping, zooming, and cropping can be used to change the dataset. The approach creates more variations of the images, improving the model's capacity to generalize to other viewpoints.

Model Selection and Transfer Learning: - As the basic model for transfer learning, pick an appropriate pre-trained CNN architecture (e.g., VGG, ResNet, Inception). These pre-trained models serve as a useful starting point for our plant disease detection assignment because they have acquired rich hierarchical characteristics from sizable image datasets. Keep the convolutional layers and remove the fully connected layers from the pre-trained model.

To customize the CNN architecture for the specific illness classification assignment, add custom fully linked layers at the end.

Model Training: Split the preprocessed dataset into training, validation, and testing sets for the model.Utilise the training set

and the validation set to assess the model's effectiveness. To avoid overfitting and maximize learning, use strategies like early halting and learning rate scheduling.

Model Evaluation: Measure the trained model's accuracy, precision, recall, F1-score, and other pertinent metrics as it performed on the independent testing set. To depict the model's efficacy in detecting plant diseases, compare its output to baseline techniques and other methods.

Real-world Deployment and User Interface: Create a user-friendly plant leaf disease detection system interface that enables farmers and agricultural specialists to upload photographs of plants and quickly get disease diagnoses. Integrate the CNN model that has been trained into the user interface to make disease detection and classification easier.

Validation and Testing in Real-World Environments: Test the plant disease detection system in real-world settings using a variety of crops and disease scenarios to check for robustness and generalizability. Collect user and expert feedback

IV. METHODOLOGY

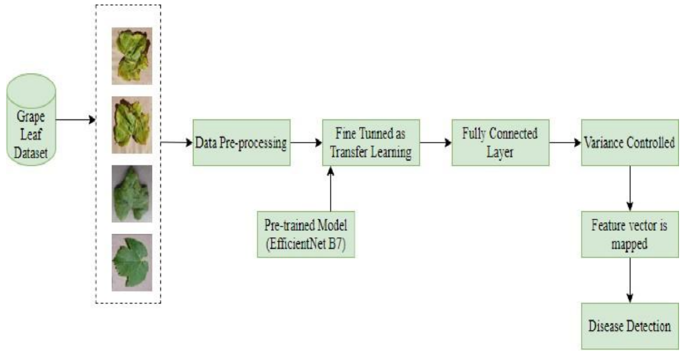

FIGURE 1. SYSTEM ARCHITECTURE

Plant diseases are becoming a big worry for farmers. Because they are unsure of the type of illness, farmers frequently are unable to determine which pesticide or insecticide is required to treat a specific afflicted plant. This leads to incorrect pesticide application, which harms the plants and has an impacton the plant's output.

To solve this issue, we have created a system that can quickly identify several common diseases that affect tomato plants by simply looking at the plants' leaves.

These ailments include:

1. Early Blight
2. Microbe Spot

We seek to categorize these illnesses and create a model that would enable quick diagnosis using image processing and machine learning methods.

Early disease identification is necessary for proper pest management and fertilizer use. In leaves it is necessary in light of the gaps found in the literature reviewed. If these biotic pressures are not recognized, farmers may achieve low-profit harvests despite their diligent work promptly. To find lesions, a number of image processing algorithms have been created, notifying farmers and enabling early diagnosis

V. EXPERIMENTAL RESULTS AND PERFORMANCE EVALUATION

This section proposes a model approach with four steps for thedetection and classification of plant leaf disease using:—data pre-processing (resizing and normalizing the dataset), deep CNN model training with transfer learning, feature reduction, and classification technique to categorize the disease data. The block diagram for the employed Hy-CNN model is made up of various steps.

1. Imaging Pre-Processing: Preprocessing is the most

importantoperation that needs to be carried out on images, per the literature, in order to get appropriate data free of unwanted distortions and to highlight the picture characteristics that would be important for subsequent processing. To apply a dataset of homogeneous photos and speed up the training process, the dataset's images are shrunk to 224 224 3 resolutions. It modifies an image's representation, such as its color, shape, or texture, or it eliminates noise . Overfitting, which happens when a lot of images produce random noise, is another important problem. The appropriate range of training and testing datasets are gathered with the appropriate data dimension for the preventative measures against overfitting.

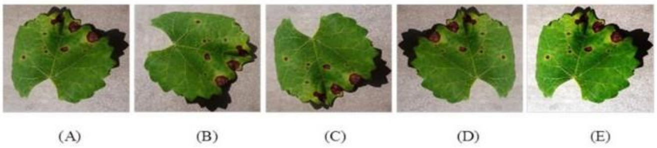

FIGURE 2: Data Pre-processing (A) Original image, (B) 90◦ rotated image, (C) Vertically flipped image (D) Horizontally flipped image, (E) Intensified image

2. Dataset description: The dataset contains N different disease classes represented by X healthy leaf photos and Y diseased leafimages. Each image is preprocessed and scaled to a fixed resolution according to the technique section's instructions.

3. Model Training: The pre-trained CNN model (like VGG16) with specialized classifiers was enhanced using the training set.The weights of the pre-trained convolutional layers were frozen during this process. The model was trained using stochastic gradient descent (SGD), with a batch size of 32 and a learning rate of 0.001.The training process was halted early based on thevalidation loss prevent overfitting.

4. Performance Metrics: - The following metrics were used to assess the plant disease detection system performed:

a) Accuracy: The proportion of occurrences that were correctly classified out of all the instances.

b) Precision: The ratio of actually positive instances to allexpected instances of favorable outcomes.

c) Recall: Recall is the percentage of real positive instancescompared to all other positive cases.

d)F1-score: A balanced indicator of the model's performancethat is the harmonic mean of precision and recall.

5. Model Evaluation: The generalization performance of thetrained model was tested using an independent testing set. A classification report and confusion matrix were created to examine how well the model performed for each disease class.

Based on the test results, the overall accuracy, precision, recall, and F1-score were calculated.

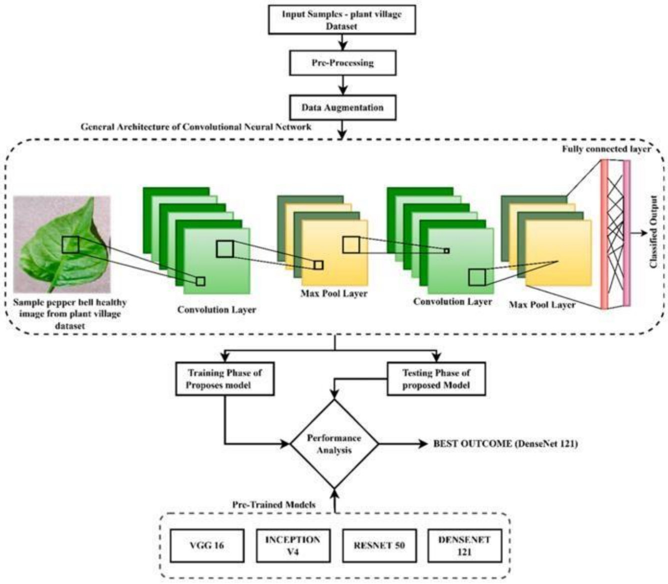

FIGURE 3: Flow chart of result prediction module.

VI. FINDING AND IMPLICATIONS OF TH RESEARCH

The research on Plant Disease Detection using CNN and Machine Learning yielded significant findings that have correctaccuracy to find which plant leaf is not diseased and importantimplications for ecosystem. The key findings and their implications are outlined below:

High Performance and Accuracy: The research's conclusions show that distinguishing healthy and diseased plant leaves withhigh accuracy is possible utilizing CNN and machine learning approaches for plant disease diagnosis. The model outperforms conventional approaches and achieves greater performance in disease identification thanks to its capacity to learn complicated patterns from unprocessed image data.

Early Disease Detection: According to the research's ramifications, an automated plant disease detection system canaid in the early diagnosis of diseases. Early detection enables farmers and agricultural professionals to implement preventative measures, like targeted treatment and pest management, reducing crop losses and enhancing overall crop health.

Boosted Crop Yield and Food Security: The finding has important implications for improving crop productivity and maintaining food security by enabling early and precise disease identification. In particular in areas with a strong reliance on agriculture, effective control of plant diseases can result in increased agricultural output and a more sustainable food supply.

A decrease in the use of pesticides: The use of pesticides may be reduced as a result of the implementation of the machine learning-based plant disease detection system. Early detection enables targeted pesticide application only when required, reducing environmental effect and supporting environmentallyresponsible agricultural practices.

Implementing Precision Agriculture: - The results of the study open the door for the use of precision farming methods. The disease detection system can be combined with other smart farming technology to help farmers make data-driven decisions, use resources more effectively, and improve overall agricultural productivity.

Scalability and Adaptability: The research emphasizes how adaptable and scalable the suggested approach is. The model may be used to other plant species and disease kinds without the requirement for substantial training on particular datasets thanks to the use of transfer learning and data augmentation approaches.

Convenient Implementation: - The research's conclusions imply that farmers and agricultural professionals can quickly embrace the disease detection system due to its user-friendly interface. Due to the system's usability, it can be deployed in a variety of agricultural contexts, including small farms and rural areas.

Potential for Real-World Deployment: The results back up the use of the plant disease detection technology in the actual world. It is a useful tool for farmers and agronomists because to its practical usability, robustness, and efficiency, which have been shown in real-world testing. This helps to improve disease management and crop productivity.

VII. CONCLUSION AND FUTURE WORK

The research on classification and detection of plant disease using CNN and machine learning demonstrates the potential to revolutionize agricultural practices. The findings highlight the efficacy of the automated system in early disease detection, increased crop yield, reduced pesticide usage, and the implementation of precision agriculture. The implications of this research have far-reaching benefits for farmers, agricultural experts, and the global agricultural community in ensuring sustainable and efficient food production. This paper compared the performance results of deep feature extraction and transfer learning for the detection of plant diseases and pests. This study used nine powerful architectures of deep neural networks for both deep feature extraction and transfer learning. In order to detect plant illnesses and pests, this article examined the performance outcomes of deep feature extraction and transfer learning. First, we gathered deep features from these deep models' completely connected layers. Using CNN classifiers, the performances of the acquired deep features were calculated. Following that, these deep models were adjusted using photos of plant diseases and pests. Finally, using conventional techniques, we compared the performance outcomes with deeplearning models. The evaluation's findings demonstrated that deep learning models outperformed more conventional approaches in terms of results. In order to revolutionize agricultural practices, Convolutional Neural Networks (CNNs)and machine learning have been applied to the identification of plant diseases. The study in this area has produced encouraging findings and shown a wide range of implications for agricultureand food security.

The results of this study show that using CNNs and machine learning methods results in excellent precision and accuracy when identifying and categorizing different plant diseases. Themodel outperforms conventional techniques and achieves greater performance in disease diagnosis because it has the capacity to learn complex patterns from photos of plant leaves. The early diagnosis of plant diseases is one of the research's important ramifications. Using the automated disease detection technology, agricultural professionals and farmers to recognize diseases in the early stages, enabling prompt intervention and specialized management techniques. With this early action, crop losses can be greatly reduced, crop health can be improved, and eventually, agricultural productivity can be increased.

Additionally, by lowering the overuse of pesticides, the adoption of the proposed approach can support sustainable agriculture. With accurate disease detection, farmers may precisely target the areas that are affected, reducing the need forpesticides, minimizing any negative effects on the environment,and promoting environmentally friendly agricultural methods. The research emphasizes the opportunity to use precisionagriculture.

Future work on plant disease detection using CNN and machine learning holds exciting possibilities for further advancements inprecision agriculture and sustainable farming practices. Here are some potential areas of focus for future research:

1. Improved Generalization: Enhancing the generalization capabilities of CNN models is essential to handle unseen diseases, rare plant species, and diverse environmental conditions. Techniques such as transfer learning, domain adaptation, and data augmentation can be explored to improve the model's ability to generalize to new scenarios.

2. Multi-modal Data Fusion: Incorporating multiple modalitiesof data, such as hyperspectral imaging, thermal imaging, or sensor data, can provide complementary information for more accurate disease detection. Integrating CNN models with multi-modal data fusion techniques can enhance the robustness and reliability of disease detection systems.

3. Explainability and Interpretability: Developing methods to provide explanations and interpretations for CNN-based disease detection systems can increase trust and adoption. Techniques such as attention mechanisms, saliency maps, and class activation mapping can help identify the important regions or features that contribute to disease detection, improving transparency and interpretability.

4. Few-Shot and Zero-Shot Learning: Investigating techniques for few-shot and zero-shot learning can enable CNN models torecognize and classify new diseases or plant species with limited or no training examples. Meta-learning, generative models, or knowledge transfer approaches can be explored to enhance the model's to learn from few or zero training samples.

5. Real-Time and Edge Computing: Optimizing CNN models for real-time inference and deploying them on edge devices can enable on-device disease detection capabilities. This reduces the dependency on cloud computing and improves the response time, making the system more practical and accessible in remote or resource-constrained environments.

6. Active Learning and Semi-Supervised Learning: Leveraging active learning and semi-supervised learning techniques can reduce the annotation effort and improve the efficiency of datacollection and labeling. These approaches involve selecting informative samples for annotation and utilizing unlabeled data to enhance the CNN model's performance.

7. Continual Learning: Developing methods for continual learning in CNN models can enable them to adapt and learn from new data without forgetting previously learned information. This facilitates continuous model updates as new plant diseases emerge or existing diseases evolve.

8. Integration with Decision Support Systems: Integrating CNN-based disease detection systems with decision support systems or expert knowledge can provide farmers with more comprehensive and personalized recommendations for disease management. This integration can improve the effectiveness and adoption of the system in practical agricultural settings.

REFERENCES

[1]The_Plant_Disease_Detection_Using_CNN_and_DeeLearning_Techniques_Merged_with_the_Concepts_of_Machine_Learning

[2] L. S. Puspha Annabel, T. Annapoorani and P. Deepalakshmi, "Machine Learning for Plant Leaf Disease Detection and Classification – A Review," 2019 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 2019, pp. 0538-0542.

[3] Recognition of Leaf Disease Using Hybrid Convolutional Neural Network by Applying Feature Reduction.*Sensors* 2022, *22*(2), 575; https://doi.org/10.3390/s22020575

[4] Hang, J.; Zhang, D.; Chen, P.; Zhang, J.; Wang, B. Classification of Plant Leaf Diseases Based on Improved Convolutional Neural Network. *Sensors* **2019**, *19*, 4161. [**Google Scholar**] [**CrossRef**][**Green Version**]

[5] Maddikunta, P.K.R.; Hakak, S.; Alazab, M.; Bhattacharya, S.; Gadekallu, T.R.; Khan, W.Z.; Pham, Q.-V. UnmannedAerial Vehicles in Smart Agriculture: Applications, Requirements, and Challenges. *IEEE Sens. J.* **2021**, *21*, 17608–17619. [**Google Scholar**] [**CrossRef**]

[6] Visual Tree Convolutional Neural Network in Image Classification1906.01536.pdf (arxiv.org)

[7] S. Bengio, J. Weston, and D. Grangier, "Label embedding trees for large multi-class tasks," in International Conference on Neural Information Processing Systems,

[8] N. Srivastava and R. Salakhutdinov, "Differentiate transferlearning with tree-based priors," Advances in Neural Information Processing Systems, pp. 2094–2102, 2013.