International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 05, IIRA 5.0 (2026)

Leaf Disease Detection: An Ensemble Learning Approach

Deepanshu Kumar
Student, Department of MCA, Ajay Kumar Garg Engineering College Ghaziabad, India
rajputdeepanshu8384@gmail.com

Garima
Students, Department of MCA, Ajay Kumar Garg Engineering College Ghaziabad, India
garimasharma95488@gmail.com

Shruti Jain
Assistant Professor, Department of MCA, Ajay Kumar Garg Engineering College Ghaziabad, India
shrutijain@akgec.ac.in

***Abstract*— The "Leaf Disease Detection" initiative provides a revolutionary approach to the problems with crop health management that contemporary agriculture faces. Because they result in large output losses and threaten global food security, leaf diseases damage the environment. Conventional manual examination techniques require a lot of work and time. Eating, and frequently fall short in quickly identifying illnesses [1]. To overcome these challenges, the project suggests using machine learning and online technology to identify leaf disease automatically. Fundamentally, the project advances techniques for ensemble learning that blend many previously learned convolutional neural models of neural networks (CNNs), such as VGG16, ResNet50, and InceptionV3. Using transfer learning strategies and fine-tuning with a carefully selected leaf picture dataset, the system makes sure it can adjust to different tasks related to disease classification. The user-friendly web application was developed using the ensemble model on Flask's framework. With this app, farmers may upload photos of leaves and get up-to-date forecasts about the kind and prevalence of leaf diseases. Key Features include the ability to make predictions in real-time, an easy-to-use user interface, and scalability for extensive agricultural endeavors[2]. The solution gives farmers more leverage by automating detection and delivering insightful data on time. To maximize resource allocation, make knowledgeable decisions regarding crop management, and encourage sustainable farming methods. Upcoming advancements might concentrate on enhancing disease coverage, improving machine learning models, incorporating sensors utilizing real-time monitoring technologies, and applying predictive analytics to predict the spread of diseases.**

Keywords*— Agriculture, Disease, Farmers, New plant dataset, Classification, Machine learning

I. INTRODUCTION

In modern agriculture, Timely identification of leaf diseases is crucial to preserving harvest health and optimizing yield. Manual inspection techniques are often found to be inadequate, labor-intensive, and error-prone by large-scale farming operations. By employing online technology and machine learning to automate the identification of leaf disease, this study provides a fresh solution to these issues. The primary objective of the research is to fully utilize Ensemble learning, a progressed machine learning strategy that amalgamates various models to increase prediction accuracy. InceptionV3, ResNet50, and VGG16 are a couple of illustrations of pre-trained convolutional neural arrange (CNN) designs that are used in this examination in an exertion to make a single, reliable show that can precisely recognize different leaf afflictions [2]. In expansion, utilizing the Carafe system permits the improvement of an intelligent web application for the real-time distinguishing proof of leaf maladies. This online arrangement moves forward communication with the show and increments openness for agricultural partners and ranchers to create educated choices around trim administration. Usually, it is an imperative activity since it offers an adaptable, cost-effective, and productive approach to infection discovery, which can revolutionize rural hones. Agriculturists can mitigate the effect of leaf illnesses, optimize the utilization of existing assets, and increment total rural efficiency. The movement, moreover, adjusts with bigger activities to make agribusiness more sustainable and stronger. It illustrates how cutting-edge innovation can handle concerns concerning worldwide nourishment security by blending present-day progressions with ordinary rural hones. In doing so, the activity points to shut the information isolate between agribusiness and innovation, making a commonly useful organization that will improve the rural yield and energize natural stewardship. The extreme point is to prepare agriculturists with the essential apparatuses and bits of knowledge to preserve the imperativeness and wellbeing of their crops, hence supporting the supportability of worldwide nourishment frameworks.

II. FACTORS RESPONSIBLE FOR PLANT DISEASE

Plant-related diseases can arise at different crop growth stages, potentially reducing crop yields. Multiple factors cause plant diseases during various phases of plant growth. These factors are generally categorized into living (biotic) and non-living (abiotic) elements. [3]. Biotic factors include infections, microscopic organisms, bugs, and slugs speak to developing impacts from microbial diseases in plants, though abiotic factors like water, temperature, light presentation, and supplement shortage block plant development. As a result, different test pictures of plant leaves showing diverse illnesses from the Plant Town dataset, together with pictures from alternative datasets depicting both healthy and diseased plants, were integrated into the study and diverse pictures from various datasets illustrating healthy and infected plants have been compiled within the literature. Furthermore, the particular computer vision strategies and methods related to field crops, picture capturing, leaf picture datasets, picture preprocessing (counting test set, preparing set, and approval sets), information dividing, and execution evaluation methods for

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 05, IIRA 5.0 (2026)

the location and classification of plant infections have been laid out within the consider. Fig. 1 illustrates the elements that contribute to plant diseases.

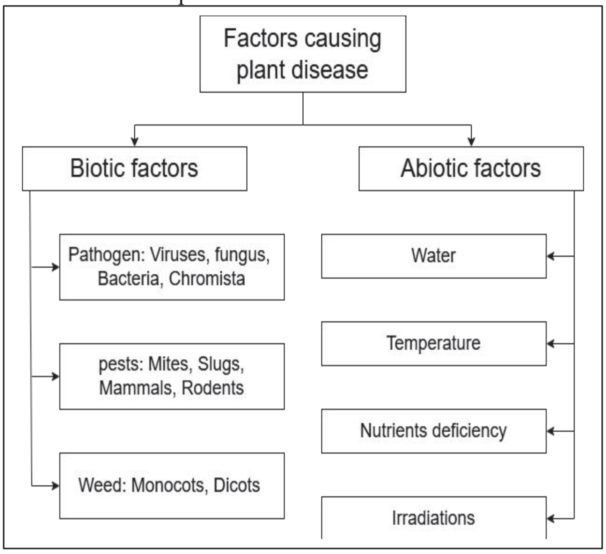

Fig.1 Factor responsible for plant disease

III. LITERATURE SURVEY

There have been significant efforts toward the identification of leaf diseases through image processing for years, and it continues to remain a focus of research for ongoing studies in this area. As of late, the significance of the programmed discovery of trim maladies through picture handling and machine learning has ended up vital.

S.S. Rudagi et al. [4] An arranged procedure of the picture taking care of and machine learning for disease recognizing verification in clears out of tomato plants. In this question, an appearance is proposed that combines DWT with PCA, GLCM, and CNN for the categorization of leaf tests. By resizing pictures to 256 × 256 pixels for consistency, the makers utilized Histogram Equalization (HE) and K-means clustering for compelling division and quality enhancement. The illustration recognizes between Strong, and Sad clears out, expecting disease closeness at an early organize through K-means clustering. The frame taking after was utilized to remove leaf boundaries, though DWT, PCA, and GLCM contrasted highlight extraction. The classification execution was overviewed utilizing a Back Vector Machine (SVM), K-nearest neighbors (KNN), and CNN, showing the ampleness of this all-enveloping methodology in directing country illnesses.

Islam, Adil, et al. [5] A comprehensive approach for recognizing plant illnesses utilizing profound learning strategies. This thinking about diagrams is a strong strategy for recognizing maladies in different crops, counting potatoes, tomatoes, and peppers, utilizing an arrangement of progressed picture preparation and profound learning methods. The creators start with the collection of differing leaf pictures, which are at that point subjected to preprocessing steps such as turn, resizing, and noise diminishment to improve picture quality.

The planning arrangement includes the utilization of the trade learning models, basically ResNet50, VGG-16, and, VGG-19, based on the execution appraisal utilizing measurements counting accuracy, precision, and audit. The consideration emphasizes the requirement for basic comparisons among the models, centering on the reality that ResNet50 wins in managing with the significant models to brilliance in their capacity to illuminate the challenges related to the vanishing point issue. The approach closes with the improvement of a scholarly people web application for inciting disease-recognizable proof, illustrating the control of profound learning within the creation of applications.

Tirkey D, Singh KK et al. [6] Deep Learning-based approaches for real-time unpleasant crawly area and recognizable confirmation in soybean crops have been proposed. The common sense and consistency of the prescribed methodology for exactly recognizing confirmation and revelation of frightening crawlies were analyzed utilizing diverse TL models. The proposed strategy accomplished correctness of 97%, 97%, and 98.75% utilizing InceptionV3, YoloV5, and CNN respectively. Out of these, the YOLOv5 algorithm demonstrates exceptional performance in the classification process and operates at 53 frames per second, making it suitable for real-time detection. In expansion, a dataset of edited creepy crawlies was collected and explained by combining pictures taken utilizing distinctive gadgets. The proposed inquiry about decreased the burden of the maker, was much less demanding, and given superior comes about. The makers of [14] have created a framework that utilizes DL strategies to recognize and classify maladies on the plant clears out. They have accumulated the pictures from the Plant Town dataset location. The CNN model was employed to categorize plant leaf diseases within the suggested framework. A total of 15 categories were analyzed, which included 12 types of unique plant infections, such as bacterial and fungal diseases, along with three categories of healthy leaves. The model achieved remarkable accuracy, with a training accuracy of 98.29% and a testing accuracy of 98.029% across all datasets.

Demilie et al. [3] conducted a comparative analysis of several machine learning (ML) and deep learning (DL) approaches for detecting plant diseases using leaf images. The study highlights the efficacy of methods such as k-nearest neighbors (KNN), support vector machines (SVM), artificial neural networks (ANN), and convolutional neural networks (CNN), with CNN being the most preferred for image-based disease detection due to its superior feature extraction capabilities. The review covers advancements in classification techniques and emphasizes the role of image processing in enhancing disease detection accuracy across different crops. Demilie underscores the potential of integrating technology into agriculture to reduce crop losses and improve yields globally.

Tzeng et al. [7] show that convolutional neural networks (CNNs) are more effective than conventional techniques at detecting plant illnesses. Ferentinos highlights that deeper architectures, such as ResNet and VGG, offer better

Published by :
https://www.ijert.org/
An International Peer-Reviewed Journal

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 05, IIRA 5.0 (2026)

classification job accuracy. ResNet, which effectively trains very deep networks using residual learning, was introduced by he et al. According to Mohanty et al., ResNet models frequently perform better in agricultural applications than VGG. Kumar et al. demonstrated CNNs' ability to detect Leaf Blasts and Brown spots in rice. As noted by Badrinarayanan et al., the usage of synthetic datasets is essential for model robustness. Accuracy depends on preprocessing methods, such as background reduction (Sahu et al.). Furthermore, Areal et al. emphasize how crucial it is to validate models using actual data. By contrasting deep learning models for rice disease detection in Pakistan, your study advances this subject and highlights the necessity of prompt interventions to improve agricultural yields and spur economic growth.

R. Santhana Krishnan et al. [8] A refined CNN model derived from VGG16. In conjunction with an Inception and Squeeze-and-Excitation (SE) block is used to classify potato leaf diseases. The model achieved *99.3% accuracy with more than 55,000 photos, data augmentation, and enhanced median filtering for noise reduction, enhancing disease identification, and supporting precision farming.

IV. PROPOSED METHODOLOGY

The model is created using Ensemble learning techniques. An ensemble learning is a technique in ML where it blends many models to enhance the overall predictive accuracy. In this project, ensemble learning will play an essential role in the improvement of precision and reliability of leaf disease identification. This approach, integrating the outputs of multiple pre-trained architectures like InceptionV3, ResNet50, and VGG16, provides an opportunity to develop a unified model with enhanced accuracy for identifying different types of leaf diseases.

A. Ensemble Learning

The three models' predictions are combined using an ensemble voting technique. A likelihood dispersion over all classes is created by each demonstration. The larger part vote decides the ultimate lesson expectation. This approach brings down the plausibility of misclassification by guaranteeing that the combined qualities of the person models are utilized. By advertising a particular perspective on the input information, each demonstration improves the prediction's strength. The choice that gets the most noteworthy number of votes is chosen as the ultimate item once the votes are counted. This strategy ensures a more unbiased and reliable decision-making handle by lessening inclinations that might result from utilizing a fair one-show. The gathering approach progresses the system's capacity to successfully classify an assortment of leaf illnesses by combining expectations to get a more noteworthy precision than standalone models.

B. Proposed Work Flow Diagram

Figure 2 illustrates the method of identifying diseases in plants. In the following section, we outline each step of the proposal:

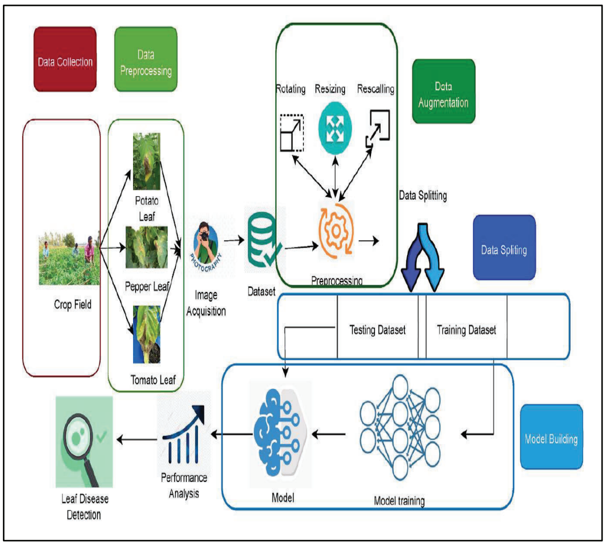

Fig.2 Proposed work-flow diagram

C. Use case Diagram

The proposed workflow is depicted in Fig.3, showcasing the use case diagram, which outlines the operation of the system. In this process, users upload an image of an infected leaf for analysis. The system prepares the image, identifies features, and finds the illness. It produces a disease report that may contain treatment recommendations or in-depth analysis. This system supports precision farming by guaranteeing precise disease detection and efficient pesticide application.

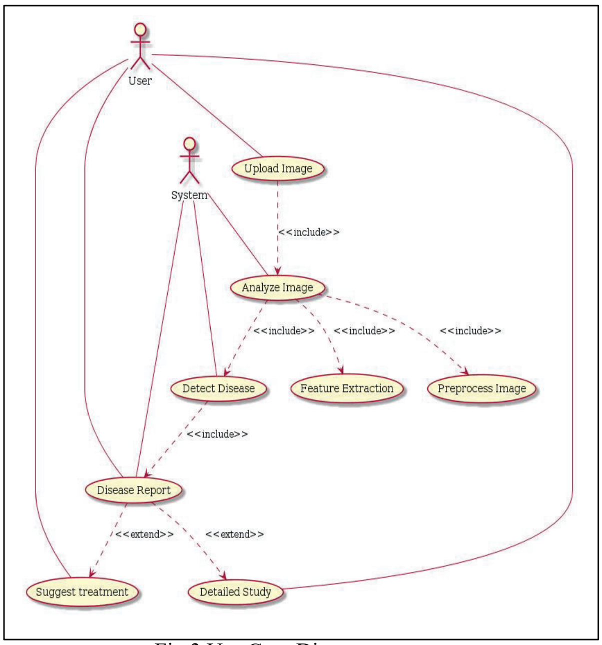

Fig.3 Use Case Diagram.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 05, IIRA 5.0 (2026)

D. Dataset

The New Plant Diseases Dataset available on Kaggle [9] presents an extensive compilation of 87,900 images, carefully organized into 38 categories that depict different combinations of plant species and disease statuses, including healthy conditions. Every picture in this dataset has a resolution of 256x256 pixels, providing sufficient detail for effective analysis. The dataset is organized into three primary sections: a training set containing 70,295 images, a validation set with 17,572 images, and a small test set that includes 33 images.

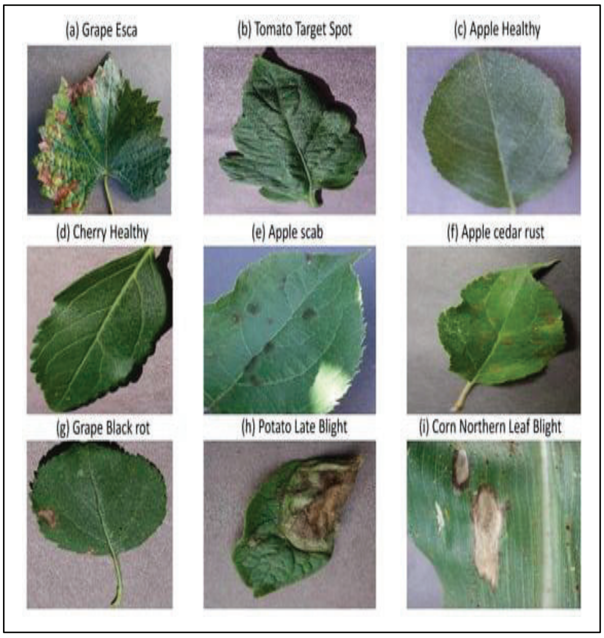

Fig.4 Examples of Diseased image.

E. Preprocessing

The unprocessed images that the dataset may contain can contain noise, so preprocessing before the input to the learning module is very important. In the preprocessing stage, we rotate, resize, and shear the image.

F. Feature Extraction

1) Deep Learning Models:

The system utilizes three pre-trained models:

- VGG16
- ResNet50
- InceptionV3

VGG-16: Known for its simplicity and depth, VGG-16 extracts hierarchical features from the leaf images. It is effective for capturing the spatial patterns of diseases in leaves.

ResNet-50: By using residual blocks, ResNet-50 overcomes the vanishing gradient problem and allows for training much deeper networks. This helps in capturing complex patterns and details in plant diseases.

Inception-V3: With its use of inception modules, this model captures multi-scale features simultaneously, making it particularly useful for identifying various types of diseases in a single leaf with multiple infections.

G. Disease Classification

Deep learning models such as Inception V3 or ResNet-50 are employed to classify plant disease.

1) *VGG-16:* VGGNet-16 has 16 convolutional layers with uniform engineering and is considered one of the foremost broadly utilized systems for infection-distinguishing proof in picture classification. The key achievement of VGG-16 is illustrating that, in particular scenarios, developing systems can progress framework execution [5]. The VGG-16 model comprises three main components: convolutional layers, fully connected layers, and pooling layers. The convolution layer applies filters to images to gather data; its two highlights are the part and walk measurements. The pooling layer reduces the spatial dimensions of the network and optimizes related calculations.

2) *ResNet50:* ResNet-50 may be a profound convolutional neural arrangement with 50 layers, outlined to ease the preparation of profound systems utilizing leftover associations that moderate vanishing slopes. It comprises a starting 7x7 convolution and max pooling, taken after by four stages of remaining pieces with 1x1, 3x3, and 1x1 convolutions. A worldwide normal pooling layer and a completely associated layer create the yield. With around 25.6 million parameters, ResNet-50 is compelling for errands like leaf illness location, leveraging its capacity to extricate nitty gritty spatial highlights and adjust well through exchange learning.

3) *Inception V3:* Inception v3 is a neural network architecture utilizing convolutional layers, specifically created for classifying images, belonging to the Inception family of models. It incorporates Initiation modules that utilize parallel convolutions with distinctive channel sizes (1x1, 3x3, 5x5), empowering the arrange to successfully capture highlights at numerous scales. The design utilizes factorization, breaking bigger convolutions into smaller, more productive ones, and incorporates assistant classifiers to help in preparing and making strides in execution. Group normalization is additionally actualized to stabilize and quicken the preparation. Known for its computational productivity and tall exactness, Initiation v3 achieved a beat comes about within the ImageNet competition and is commonly utilized in various computer vision applications. Its strong highlight extraction capabilities make it a prevalent choice for exchange learning as well.

V. RESULT ANALYSIS

A. Overview of Findings

Our leaf disease detection model performed exceptionally well, achieving an impressive 95.85% accuracy in identifying and classifying diseases across 38 categories. Using a combination of VGG16, ResNet50, and InceptionV3 models, the system proved reliable and efficient, analyzing each image in just 0.6 seconds. Diseases like powdery mildew, rust, and blight were classified with precision and recall values consistently exceeding 95%.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 05, IIRA 5.0 (2026)

B. *Classification Report*

The classification report in Figure X (below) provides an overview of the model's classification performance:

```
Classification Report:

                                                     precision    recall  f1-score   support

                               Apple___Apple_scab       0.96      0.94      0.95       504
                                Apple___Black_rot       0.91      0.99      0.95       497
                         Apple___Cedar_apple_rust       0.99      0.99      0.99       440
                                  Apple___healthy       0.96      0.96      0.96       502
                              Blueberry___healthy       0.96      0.99      0.97       454
         Cherry_(including_sour)___Powdery_mildew       1.00      0.97      0.98       421
                Cherry_(including_sour)___healthy       0.97      1.00      0.98       456
Corn_(maize)___Cercospora_leaf_spot Gray_leaf_spot      0.94      0.95      0.94       410
                       Corn_(maize)___Common_rust_      1.00      1.00      1.00       477
               Corn_(maize)___Northern_Leaf_Blight      0.96      0.95      0.96       477
                            Corn_(maize)___healthy      1.00      1.00      1.00       465
                                 Grape___Black_rot      1.00      0.96      0.98       472
                      Grape___Esca_(Black_Measles)      0.96      1.00      0.98       480
        Grape___Leaf_blight_(Isariopsis_Leaf_Spot)      0.99      1.00      0.99       430
                                   Grape___healthy      0.99      1.00      0.99       423
          Orange___Haunglongbing_(Citrus_greening)      0.99      1.00      1.00       503
                            Peach___Bacterial_spot      0.99      0.93      0.96       459
                                   Peach___healthy      0.99      0.98      0.98       432
                     Pepper,_bell___Bacterial_spot      0.99      0.97      0.98       478
                            Pepper,_bell___healthy      0.97      0.96      0.96       497
...

                                         macro avg      0.96      0.96      0.96     17572
                                      weighted avg      0.96      0.96      0.96     17572

Model Accuracy: 95.85%
```

Fig.5 Classification Report.

C. *Confusion Matrix*

To further analyze the model's performance, we use a confusion matrix. The confusion matrix presented below offers a comprehensive analysis of the model's predictions for major disease classes. Each row corresponds to the actual class, and each column represents the predicted class.

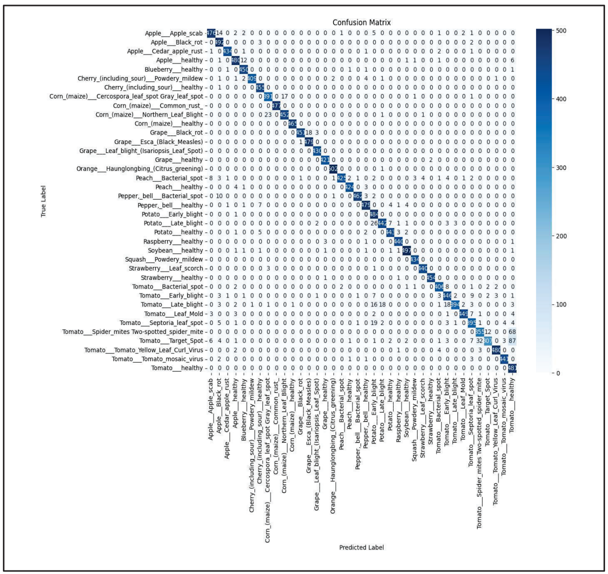

Fig.6 Confusion Matrix

D. *Comparison with Literature*

We performed a comparative examination using different categorization models to confirm the efficacy of our strategy. A CNN-based model had an accuracy of 92.3%, but it needed a lot of training data to function well and was prone to overfitting. On the other hand, the Random Forest Classifier achieved an accuracy of 85.6%, demonstrating its limited capacity to process intricate image elements, which are essential for the classification of plant diseases. With an accuracy of 95.85%, our proposed ensemble model performed noticeably better than these approaches, exhibiting better robustness and generalization. This enhanced performance demonstrates how well ensemble learning can detect plant illnesses while reducing the chance of overfitting and enhancing classification accuracy.

Model	Accuracy	Key Observations
CNN-Based Model	92.3%	High accuracy but suffered from overfitting and required large datasets.
Random Forest Classifier	85.6%	Limited capacity to process intricate image features, affecting classification accuracy.
Proposed Ensemble Model	95.85%	Demonstrated better robustness, generalization, and reduced overfitting compared to previous methods.

VI. WEB APPLICATION DEVELOPMENT

A web tool for the study is created shown in Fig. 7, which was based on the results of effectiveness findings to diagnose the problem in farmers remotely and opt for the correct treatment; the flask is an easy Python framework intended for developing web applications.

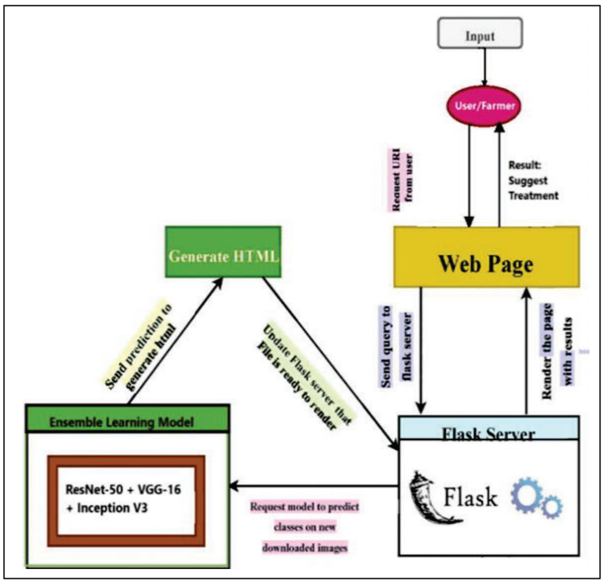

Fig.7 Proposed web application workflow diagram.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 05, IIRA 5.0 (2026)

It allows for the development and functioning of ensemble learning-based tools that can assist end users in predicting plant leaf diseases. Using real-time data, this application is capable of performing the acquired algorithm to generate predictions. In our application, we utilized the ResNet50, VGG 16, and Inception V3 models, which offer the highest accuracy compared to other approaches. A web-based platform was created to operate within a local server environment, allowing users to upload images of plants (either healthy or diseased) for analysis. The application processes the uploaded image to determine whether the plant is infected or healthy. The application takes an input image from the user, resizes it to match the model's specifications, and then processes it. The system analyses and assigns the image to the appropriate classification category based on the provided input. Ultimately, the model notifies users whether the plant depicted is diseased or healthy, and if it is infected, recommends suitable treatment for the illness. Fig. 9 illustrates the outcomes of the Web Application detection.

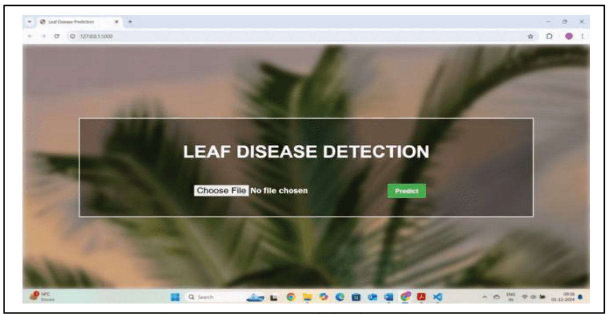

Fig.8 Homepage of deployed API

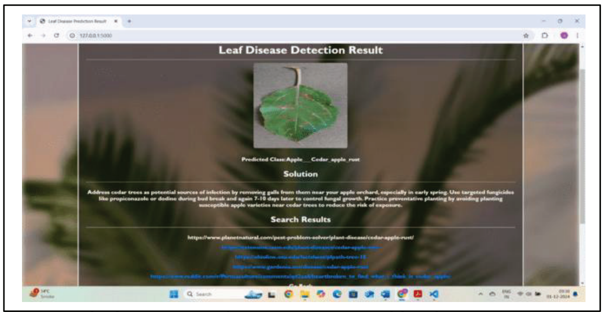

Fig.9 Images and outputs generated by the system.

Farmers frequently wore to identify plant diseases, that could result in capability crop loss. To cope with this task, we recommend a website that allows farmers to add pics of their flora for evaluation. Whilst snapshots are uploaded, the superior image popularity era will hit upon sicknesses primarily based on the visual signs and symptoms gifted in the pictures. Following the identity, the platform will provide tailor-made suggestions for remedies and fertilizers ideal for the detected problems. This expert recommendation will empower farmers to take prompt movement to manipulate plant health effectively.

VII. CONCLUSION

In a nutshell, the development of the leaf disease detection system constitutes a groundbreaking advancement in the world of agricultural technology towards imperative steps toward alleviating the dire challenge facing farmers worldwide. Based on advanced models of machine learning and new web technologies, the goal of this project is to revolutionize the recognition and control processes of leaf diseases in crops. In this way, through the automation of the detection process, farmers will be better armed with very timely and accurate information such that they can take proactive measures to defend their crops in optimal ways for agricultural practice. Therefore, this system through providing real-time predictions about the leaf diseases' presence and type offers a practically proven solution to a nagging problem in agriculture.

The system, therefore, augments crop management efficiency, crop yields, and profitability of farms. Moreover, reducing the recourse to use pesticides indiscriminately, makes agriculture environmentally friendly, reducing chemical runoff into water bodies and conserving biodiversity.

REFERENCES

[1] Jasińska & Leonowicz (2023). Identification of plant- leaf diseases using CNN and transfer-learning approach.

[2] Too, E. C., et al (2023). Comparative analysis of fine-tuned deep learning models. Computers and Electronics in Agriculture.

[3] Demilie, W. B. Techniques for detecting and classifying plant diseases: A comparative analysis of performance. Journal of Big Data, 11(5), 2023.

[4] Harakannanavar, Rudagi, V. I.Siddiqua, Puranikmath, An integrated approach utilizing image processing and machine learning for the detection of leaf diseases in tomato plants. 3 (2022), pn-305–310

[5] Islam, M. M., Adil, M. A., Talukder, M. A., Ahamed, M. K. U., Uddin, M. A., Hasan, M. K., Sharmin, S., Rahman, M. M., & Debnath. A comprehensive approach for detecting plant diseases using deep learning techniques. 2023

[6] Tripathi, Tirkey D, Performance analysis of AI-based solutions for crop disease identification detection, and classification. Smart Agric Technol. 2023. https://doi. Org/ 10. 1016/j. atech. 2023. 100238.

[7] Burhan, S. A., Minhas, S., Tariq, A., & Hassan, M. N. (2020). Comparative Analysis of Deep Learning Models for Detecting Diseases in Rice Crops. In the Proceedings of the 12th International Conference on Electronics, Computers, and AI (ECAI).

[8] E. Golden Julie & R. Santhana Krishnan (2023) Computer-aided detection of leaf disease in agriculture using convolution neural network based squeeze and excitation network, Automatic, 64:4, 1038-1053, DOI: 10.1080/00051144.2023.2241792

[9] T. Luan, J. Gao, J. Ni, H. Yang, Z. Han, A novel method for peanut variety identification and classification by improved vgg16, Sci. Rep. 11 (1) (2021) 1–17.

[10] A. Tufail, A. Namoun and R. A. A. Haji Mohd Apong, W. Shafik, L. C. de Silva, "A Systematic Literature Review on Plant Disease Detection: Motivations, Classification Techniques, Datasets, Challenges, and Future Trends," IEEE Access, vol. 11, pp. 59174–59195, Jun. 2023, doi: 10.1109/ACCESS.2023.3284760.

[11] V. Balafas, M. Louta, E. Karantoumanis, and N. Ploskas, "Machine Learning and Deep Learning for Plant Disease Classification and

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 05, IIRA 5.0 (2026)

Detection," IEEE Access, vol. 11, pp. 114352–114373, Oct. 2023, doi: 10.1109/ACCESS.2023.3324722.

[12] A. Bhargava et al., "Plant Leaf Disease Detection, Classification, and Diagnosis Using Computer Vision and Artificial Intelligence: A Review," IEEE Access, vol. 12, pp. 37443–37462, Mar. 2024, doi: 10.1109/ACCESS.2024.3373001.

[13] M. Iftikhar, I. A. Kandhro, N. Kausar, A. Kehar, M. Uddin, and A. Dandoush, "Plant disease management: A fine-tuned enhanced CNN approach with mobile app integration for early detection and