Apple Leaf Disease Detection and Classification System Using Deep Learning

Dr. P.D. Khandait[1], Prathmesh G. Ghangare[2], Pranay G. Zingre[3], and Ashish T. Mahalle[4]
[1]*Professor, Department of Electronics and Telecommunication Engineering, KDK College of Engineering, Nagpur, Maharashtra, India*
[2,3,4]*Student, Department of Electronics and Telecommunication Engineering, KDK College of Engineering, Nagpur, Maharashtra, India*

***Abstract*—The agricultural sector plays a vital role in global food security and economic stability but faces challenges such as crop diseases that impact productivity and quality. Manual disease detection is labor-intensive and error-prone, especially in large fields, necessitating automated solutions. This study introduces a deep learning framework using Convolutional Neural Networks (CNNs) integrated with a Streamlit-based web interface for detecting cotton and apple leaf diseases. The methodology involves data collection, preprocessing through normalization and augmentation, model selection using DenseNet121, ResNet50V2, and Xception, and deployment via Streamlit. Transfer learning with pre-trained models enhances accuracy while optimizing training efficiency.**

***Index Terms*—Convolutional Neural Networks (CNNs), DenseNet121, Image Processing, InceptionV3, ResNet50V2, VGG16, VGG19.**

I. INTRODUCTION

I am Global food security and economic stability depend on agriculture, yet it faces many obstacles, especially from plant diseases that jeopardize crop quality and quantity. Apple and cotton crops, which are vital to the food and textile industries, are particularly vulnerable to a number of leaf diseases that, if not identified in time, can result in significant financial losses. Conventional disease detection techniques depend on expert manual inspection, which is labor-intensive, time-consuming, and prone to human error. With the rise of artificial intelligence (AI) and deep learning, automated solutions offer a promising alternative for efficient and accurate disease detection.

Convolutional Neural Networks (CNNs), widely used in image classification tasks, provide a robust approach to diagnosing plant diseases by learning patterns from thousands of labeled images. This capability makes deep learning an ideal technology for developing scalable and accurate plant disease detection frameworks.

The proposed framework leverages CNN architectures such as ResNet50 and DenseNet121, combined with Streamlit, an open-source Python library, to create an interactive and user-friendly web application for detecting cotton and apple leaf diseases. The methodology involves data preprocessing steps, including image resizing, normalization, and data augmentation, to enhance model robustness.

Transfer learning is employed to improve classification accuracy by utilizing pre-trained models. Once trained, the model is evaluated based on performance metrics such as accuracy, precision, recall, and F1-score to ensure reliable predictions. The trained model is then deployed in a Streamlit application, allowing users to upload images of leaves and receive instant disease classification results along with confidence scores. The interactive interface also provides educational resources and disease management recommendations, making the system accessible to farmers and agricultural professionals. This deep learning-based disease detection framework offers several advantages, including real-time disease identification, improved accuracy, and reduced dependency on manual inspections.

The use of a web-based interface makes the system easily accessible to users with minimal technical expertise, while its scalability allows for future expansion to additional crops and diseases. However, challenges such as data collection, environmental variability, and computational requirements must be addressed to ensure the model's effectiveness in real-world conditions. Future enhancements could include integrating mobile applications, real-time image capture using drones, and edge computing to improve accessibility and efficiency. By combining deep learning with user-friendly deployment, this framework has the potential to revolutionize agricultural disease management,contributing to increased productivity and sustainable farming practice.

II. RELATED WORK

Recent advancements in deep learning have significantly enhanced the detection and classification of apple leaf diseases. Notably, researchers have developed models like the Coordination Attention EfficientNet (CA-ENet), which integrates coordination attention mechanisms to improve feature representation, achieving high accuracy in identifying various apple diseases. [11] Similarly, the AppleLeafNet framework offers a lightweight solution, employing a 37-layer deep learning model to distinguish between healthy and diseased apple leaves, further subclassifying diseases such as rust, scab, and frogeye leaf spot. To address challenges in real-time detection, modified versions of existing architectures have been proposed. For instance, an improved YOLOv5-based method, termed A-Net, has been developed to efficiently detect apple leaf disease spots, enhancing both detection speed and accuracy.

Furthermore, a lightweight YOLOv8 model has been developed specifically for the detection of apple leaf disease in natural settings. This model can be easily deployed on mobile and embedded devices, offering useful solutions for disease monitoring on-site. In this field, transfer learning strategies have also been successfully applied. By fine-tuning pre-trained models like EfficientNetV2S on apple leaf disease datasets, researchers have achieved impressive classification accuracies, demonstrating the potential of leveraging existing models for specific agricultural applications.

Several studies have explored the use of deep learning models for the detection and classification of apple leaf diseases, leveraging convolutional neural networks (CNNs) and transfer learning approaches. Researchers have employed architectures such as VGG16, ResNet, Inception, and Xception, demonstrating their ability to extract high-level features from leaf images for accurate disease identification. Studies have shown that pre-trained models fine-tuned on agricultural datasets outperform traditional machine learning techniques, as they effectively learn complex patterns in diseased leaves. Additionally, some works have integrated hyperspectral imaging and attention mechanisms to enhance classification accuracy further. However, challenges such as dataset imbalance, model generalization, and real-time deployment continue to be areas of ongoing research in plant disease detection systems.

III. METHODOLOGY

The methodology for implementing a Convolutional Neural Network (CNN) for image classification follows a structured approach to ensure an efficient model capable of performing well on unseen data. The first step in the process is data preprocessing, which prepares the raw images for model training. Images are rescaled to normalize pixel values, converting the pixel range to a scale of 0 to 1 by dividing each pixel value by 255. This normalization ensures that the model does not assign excessive importance to certain pixels, leading to faster and more stable training. Data augmentation is then applied to introduce variability into the dataset through transformations such as rotation, shifting, shearing, zooming, and horizontal flipping.

By ensuring that the model generalizes patterns rather than memorizes specific details, these techniques help the model learn robust features and avoid overfitting. To give an objective assessment of the model's performance, the test data is only rescaled without being enhanced, maintaining its original structure.

Once data preprocessing is completed, the CNN architecture is designed using a Sequential model, which allows layers to be added in a linear fashion. The architecture consists of multiple convolutional layers, followed by max-pooling layers, a flattening layer, fully connected layers, and a SoftMax output layer. The first convolutional layer has 32 filters to extract basic image features such as edges and textures, while each subsequent layer increases the filter count—64 in the second and 128 in the third—enabling the model to learn more complex and abstract features. Max-pooling layers follow convolutional layers to reduce spatial dimensions by selecting the highest pixel value from small patches of the feature map, improving computational efficiency while retaining essential information.

Prior to being processed by the fully connected layers, which eventually result in the SoftMax output layer, the output from the final convolutional layer is flattened into a one-dimensional vector. The chance that a picture belongs to a certain class is represented by the SoftMax activation function, which makes sure that output values add up to

1. The Adam optimizer, which effectively modifies learning rates to maximize convergence speed and accuracy, is used to compile the model. Categorical cross-entropy, a loss function that quantifies the discrepancy between the actual labels and the predicted probability distribution, is employed. This function is particularly effective for multi-class classification problems.

The evaluation metric chosen is accuracy, which represents the percentage of correctly classified instances. While accuracy is a straightforward metric, additional performance measures such as precision, recall, and F1-score are used during evaluation to account for class imbalances. The CNN model is then trained for 20 epochs, meaning the entire dataset passes through the network 20 times. Each training cycle involves backpropagation, where the model adjusts weights based on the error calculated using the loss function. A validation set is also used to monitor performance, helping identify potential overfitting when training accuracy improves but validation accuracy stagnates or declines.

After training, the model's capacity for generalization is assessed using a test dataset. Test accuracy, loss, confusion matrix, and classification report are important evaluation indicators. The confusion matrix, which displays the number of true positives, false positives, true negatives, and false negatives for each class, offers comprehensive insights into the model's classification performance. The classification report further includes precision (proportion of correctly classified positives), recall (correct identification of all relevant instances), and F1-score (a harmonic mean of precision and recall).

These metrics are essential in assessing the model’s real-world applicability. After evaluation, the model is tested on individual images from outside the dataset to ensure its effectiveness in practical scenarios. Once confirmed, the trained model is saved for future deployment, eliminating the need for retraining, thus saving computational resources and time. The methodology for implementing a DenseNet121 model follows a systematic process that leverages transfer learning to improve image classification performance. Like CNN, the first step is data preprocessing, which includes rescaling images to normalize pixel values between 0 and 1. Data augmentation is applied to the training dataset to artificially increase diversity, ensuring the model learns generalizable features. Augmentation techniques, such as random rotations, shifting, shearing, zooming, and horizontal flipping, enhance the model’s ability to handle variations in real-world data. The test dataset undergoes only rescaling, ensuring that evaluation results reflect the model's generalization ability without additional transformations affecting the results.

The DenseNet121 model, pre-trained on ImageNet, is used as the base model for transfer learning. DenseNet’s distinctive feature is its dense connectivity, where each layer receives inputs from all previous layers. This structure enables feature reuse, reducing the number of trainable parameters and improving efficiency. Since DenseNet121 was originally trained for a different classification task, its top layers are removed, and new custom layers are added for the specific image classification task. These include a global average pooling layer, which reduces the spatial dimensions of feature maps, followed by a dense layer with 128 units and ReLU activation for learning high-level representations.

By randomly setting a portion of the input neurons to zero, a dropout layer is incorporated to prevent overfitting and force the model to generalize more effectively. The SoftMax activation function is used in the final output layer to produce class probabilities that add up to one, where the predicted class is indicated by the highest probability. Like CNN, DenseNet121 is compiled using the Adam optimizer and categorical cross-entropy loss function after the model architecture is defined. To identify overfitting, the model's performance is tracked on a validation set after 20 epochs of training. During training, the confusion matrix, classification report, and accuracy metrics are analyzed to assess performance across different classes.

Since DenseNet reuses features effectively through its dense connections, it often converges faster than traditional CNN models, making it a powerful choice for image classification tasks. After training, the model undergoes final evaluation using test accuracy, loss, and detailed classification metrics to measure its effectiveness on unseen data.

A prediction function is developed to classify and display random images from the test set by comparing predicted and actual labels, and once the model is verified to be accurate, it is saved for subsequent use. In conclusion, both CNN and DenseNet121 follow a structured methodology that includes data preprocessing, model design, training, evaluation, and deployment. CNN constructs an architecture from scratch, but DenseNet uses transfer learning to increase efficiency and accuracy, making it a preferred choice for complex classification tasks requiring high generalization capability.

Each epoch represents a full pass through the training dataset, with performance metrics like accuracy and loss plotted over time. The validation dataset makes sure that the model learns generalizable patterns rather than just memorizing the training data, and if needed, adjustments like early stopping, learning rate scheduling, and regularization techniques are applied. During the training phase, all three models—ResNet50V2, InceptionV3, and Xception—are assessed using a separate validation dataset to track learning progress and mitigate overfitting.
Training typically takes 20 epochs, though this can change based on the processing capabilities and complexity of the dataset. An independent test dataset that was not used for training or validation is then used to assess the performance of the trained model. This offers an objective evaluation of the model's generalizability to actual images.

IV. MODEL ARCHITECTURE

During training, a number of deep learning architectures are used to identify characteristics suggestive of diabetic retinopathy, such as DenseNet121, InceptionV3, ResNet50, ResNet101, VGG16, VGG19, and Inception-ResNetV3. The accuracy and loss of the model are used to evaluate its performance.
graphs, a confusion matrix, and classification results. Once training is complete, the model is saved as model.h5 and tested on unseen data to generate predictions. The final step involves evaluating key performance metrics such as precision, recall and F1-score to access the modals effectiveness in detecting diseases in plant leaf .

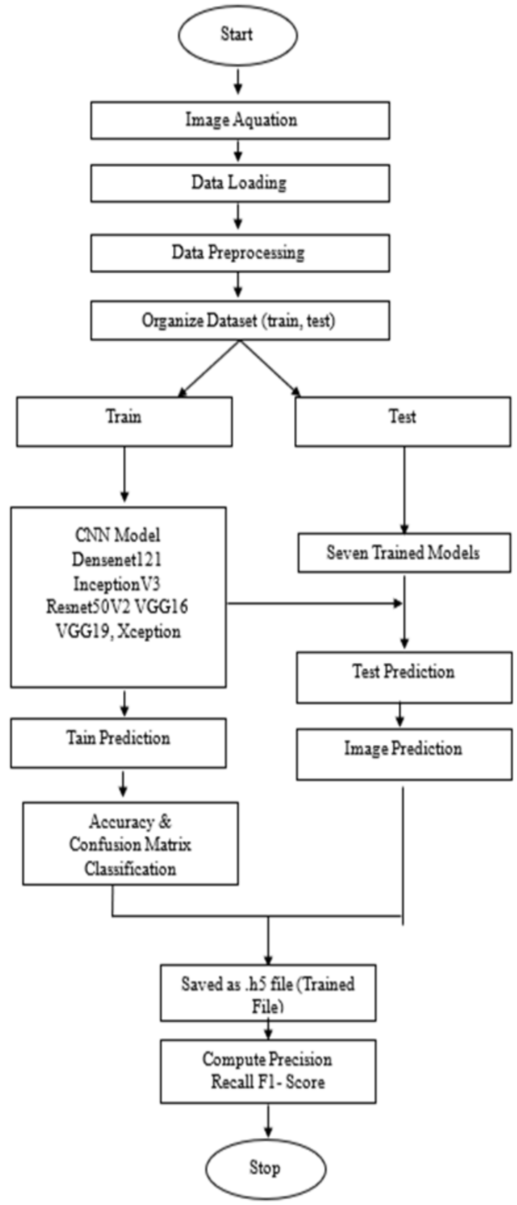

Figure1. Apple Leaf Disease Model Training And Evaluation Workflow.

Result
The classification report for apple leaf disease detection presents a comparative analysis of various deep learning models, including CNN, ResNet50, VGG19, Xception, and Inception V3, evaluated using key performance metrics such as F1-score, precision, recall, and accuracy. Among these models, VGG19

and Xception demonstrate superior performance, both achieving an F1-score and precision of 0.95, with VGG19 slightly outperforming Xception in recall at 0.96 compared to 0.95, while maintaining the highest accuracy of 0.95. With an F1-score and precision of 0.93, a recall of 0.93, and an accuracy of 0.95, Inception V3 also demonstrates strong classification capabilities, making it a competitive alternative to VGG19 and Xception. CNN and ResNet50, on the other hand, perform noticeably worse, with CNN achieving an F1-score, precision, recall, and accuracy of just 0.25 and ResNet50 recording the lowest F1-score of 0.19 and an accuracy of 0.22. These findings show that deeper and more sophisticated architectures like VGG19, Xception, and Inception V3 use their sophisticated feature extraction capabilities to achieve high classification accuracy and robustness, whereas basic CNN and ResNet50 architectures have difficulty classifying apple leaf diseases. In the end, Xception and VGG19 prove to be the most successful models, closely followed by Inception V3, proving their applicability for precise apple leaf disease detection in practical settings. Table below summarizes the comparison of model performance:

Model	F1-Score	Precision	Recall	Accuracy
CNN	0.25	0.25	0.25	0.25
RESNET50	0.19	0.22	0.22	0.22
VGG16	1.00	1.00	1.00	1.00
VGG19	0.95	0.95	0.96	0.97
XCEPTION	0.95	0.95	0.95	0.95
INCEPTION V3	0.93	0.93	0.93	0.93

Table 1. Analysis of Comparative Table

V. CONCLUSION

Deep learning-based models, especially convolutional neural networks (CNNs) and advanced architectures like Xception, Inception V3, VGG19, and ResNet50, have greatly improved the accuracy and speed of disease identification. Traditional methods of disease detection, which rely on manual inspection and expert knowledge, are frequently time-consuming, error-prone, and impractical for large-scale farming. Deep learning-based models have shown remarkable advancements in agricultural technology, offering a robust and efficient approach to plant disease management.

Through the use of deep learning, farmers and researchers can deploy automated and real-time disease detection systems, decreasing reliance on human expertise while improving agricultural productivity and crop health. These models process high-dimensional image data, extracting complex features that allow precise classification of various leaf diseases.

In the classification results, models like Xception and VGG19 achieved accuracy rates above 93%, demonstrating their reliability in differentiating between healthy and diseased apple leaves; the Inception V3 model also performed well, while ResNet50 and simple CNN architectures showed lower classification accuracy, indicating that deeper models with optimized feature extraction layers are more effective. The comparative analysis of various deep learning models has shown that advanced architectures like Xception and VGG19 outperform traditional CNN models in terms of precision, recall, and accuracy.

This implies that in order to achieve high-performance disease classification, choosing the appropriate model architecture is essential. Furthermore, by avoiding overfitting and guaranteeing strong generalization, data augmentation methods, hyperparameter adjustments, and transfer learning approaches have improved model performance even further.

REFERENCES

[1] T. Setiyono, T. Gentimis, F. Rontani, D. Duron, G. Bortolon, R. Adhikari, B. Acharya, K.J. Han, W.D. Pitman, Application of TensorFlow model for identification of herbaceous mimosa (Mimosa strigillosa) from digital images, Smart Agricultural Technology, Volume 7, 2024, 100400, ISSN 2772-3755, https://doi.org/10.1016/j.atech.2024.100400

[2] Jie Yuan, Performance Analysis of Deep Learning Algorithms Implemented Using PyTorch in Image Recognition, Procedia Computer Science, Volume 247, 2024, Pages 61-69, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2024.10.008

[3] Julien Demagny, Camille Roussel, Maïlys Le Guyader, Eric Guiheneuf, Véronique Harrivel, Thomas Boyer, Momar Diouf, Michaël Dussiot, Yohann Demont, Loïc Garçon, Combining imaging flow cytometry and machine learning for high-throughput schistocyte quantification: A SVM classifier development and external validation cohort, eBioMedicine, Volume 83, 2022, 104209, ISSN2352-3964, https://doi.org/10.1016/j.ebiom.2022.104209

[4] Xudong Liu, the educational resource management based on image data visualization and deep learning, Heliyon, Volume 10, Issue 13, 2024, e32972, ISSN 2405-8440, https://doi.org/10.1016/j.heliyon.2024.e32972

[5] Rajiv Kumar Srivastava, Sanju Purohit, Edris Alam, Md Kamrul Islam, Advancements in soil management: Optimizing crop production through interdisciplinary approaches, Journal of Agriculture and Food Research, Volume 18, 2024, 101528, ISSN 2666-1543, https://doi.org/10.1016/j.jafr.2024.101528

[6] H.M. Forhad, Md. Ripaj Uddin, R.S. Chakravorty, A.M. Ruhul, H.M. Faruk, Sarker Kamruzzaman, Nahid Sharmin, AHM Shofiul Islam Molla Jamal, Md. Mezba-Ul Haque, AKM M Morshed, IoT based real-time water quality monitoring system in water treatment plants (WTPs), Heliyon, Volume 10, Issue 23, 2024, e40746, ISSN 2405-8440, https://doi.org/10.1016/j.heliyon.2024.e40746

[7] Samsul Huda, Yasuyuki Nogami, Maya Rahayu, Takuma Akada, Md. Biplob Hossain, Muhammad Bisri Musthafa, Yang Jie, Le Hoang Anh, IoT-Enabled Plant Monitoring System with Power Optimization and Secure Authentication, Computers, Materials and Continua, Volume 81, Issue 2, 2024, Pages 3165-3187, ISSN 1546-2218, https://doi.org/10.32604/cmc.2024.058144

[8] Fabian Sittaro, Michael Vohland, IPS Monitor – A habitat suitability monitoring tool for invasive alien plant species in Germany, International Journal of Applied Earth Observation and Geoinformation, Volume 135, 2024, 104252, ISSN 1569-8432, https://doi.org/10.1016/j.jag.2024.104252

[9] Arie Pieter Paulus Ravensbergen, Martin K. van Ittersum, Renske Hijbeek, Corné Kempenaar, Pytrik Reidsma, Field monitoring reveals scope to reduce environmental impact of ware potato cultivation in the Netherlands without compromising yield, Agricultural Systems, Volume 220, 2024, 104091, ISSN 0308-521X, https://doi.org/10.1016/j.agsy.2024.104091

[10] Nelson Kevin Sinisterra-Solís, Neus Sanjuán, Javier Ribal, Vicent Estruch, Gabriela Clemente, Dataset to monitor regionalised environmental impacts of the main agricultural products in Spain, Data in Brief, Volume 46, 2023, 108883, ISSN 2352-3409, https://doi.org/10.1016/j.dib.2023.108883

[11] S.M. Nuruzzaman Nobel, Maharin Afroj, Md Mohsin Kabir, M.F. Mridha, Development of a cutting-edge ensemble pipeline for rapid and accurate diagnosis of plant leaf diseases, Artificial Intelligence in Agriculture, Volume 14, 2024, Pages 56-72, ISSN 2589-7217, https://doi.org/10.1016/j.aiia.2024.10.005

[12] Yahya Faqir, Abdul Qayoom, Elizabeth Erasmus, Marietjie Schutte-Smith, Hendrik G. Visser, A review on the application of advanced soil and plant sensors in the agriculture sector, Computers and Electronics in Agriculture, Volume 226, 2024, 109385, ISSN 0168-1699, https://doi.org/10.1016/j.compag.2024.109385

[13] Sandesh Bhagat, Manesh Kokare, Vineet Haswani, Praful Hambarde, Trupti Taori, P.H. Ghante, D.K. Patil, Advancing real-time plant disease detection: A lightweight deep learning approach and novel dataset for pigeon pea crop, Smart Agricultural Technology, Volume 7,2024,100408, ISSN 2772-3755, https://doi.org/10.1016/j.atech.2024.100408

[14] Javed Rashid, Imran Khan, Ghulam Ali, Shafiq ur Rehman, Fahad Alturise, Tamim Alkhalifah, Real-Time Multiple Guava Leaf Disease Detection from a Single Leaf Using Hybrid Deep Learning Technique, Computers, Materials and Continua, Volume 74, Issue 1, 2022, Pages 1235-1257, ISSN 1546-2218, https://doi.org/10.32604/cmc.2023.032005

[15] Weiyue Xu, Ruxue Yang, Raghupathy Karthikeyan, Yinhao Shi, Qiong Su,GBiDC-PEST: A novel lightweight model for real-time multiclass tiny pest detection and mobile platform deployment, Journal of Integrative Agriculture, 2024, ISSN 2095-3119, https://doi.org/10.1016/j.jia.2024.12.017

[16] Nueraili Aierken, Bo Yang, Yongke Li, Pingan Jiang, Gang Pan, Shijian Li,A review of unmanned aerial vehicle based remote sensing and machine learning for cotton crop growth monitoring, Computers and Electronics in Agriculture, Volume 227, Part 2, 2024,109601, ISSN 0168-1699, https://doi.org/10.1016/j.compag.2024.109601
[17] L.J. Velazquez-Chavez, A. Daccache, A.Z. Mohamed, M. Centritto,Plant-based and remote sensing for water status monitoring of orchard crops: Systematic review and meta-analysis, Agricultural Water Management, Volume 303,2024,109051, ISSN 0378-3774, https://doi.org/10.1016/j.agwat.2024.109051
[18] Kushagra Sharma, Shiv Kumar Shivandu, Integrating artificial intelligence and Internet of Things (IoT) for enhanced crop monitoring and management in precision agriculture, Sensors International, Volume 5, 2024,100292, ISSN 2666-3511, https://doi.org/10.1016/j.sintl.2024.100292
[19] Shubhika Shubhika, Pradeep Patel, Rickwinder Singh, Ashish Tripathi, Sandeep Prajapati, Manish Singh Rajput, Gaurav Verma, Ravish Singh Rajput, Nidhi Pareek, Ganesh Dattatraya Saratale, Aakash Chawade, Kamlesh Choure, Vivekanand Vivekanand,Application of artificial intelligence techniques to addressing and mitigating biotic stress in paddy crop: A review,Plant Stress,Volume 14,2024,100592, ISSN 2667-064X, https://doi.org/10.1016/j.stress.2024.100592
[20] Mireille Gloria Founmilayo Odounfa, Charlemagne D.S.J. Gbemavo, Souand Peace Gloria Tahi, Romain L. GlÃ¨lÃ¨ KakaÃ¯, Deep learning methods for enhanced stress and pest management in market garden crops: A comprehensive analysis, Smart Agricultural Technology, Volume 9, 2024,100521, ISSN 2772-3755, https://doi.org/10.1016/j.atech.2024.100521