Crop Disease Detection using VGG16

line 1: Sameh Raouf Helmy   
line 2: *Student*  
line 3: FCAI at Cairo universityline 4: Cairo, Egypt  
line 5: samehraouf2000@gmail.com

line 1: -  
line 2: *-*  
line 3: *-*line 4: -, -  
line 5:-line 1: 2nd Osama Ibrahim Marzok  
line 2: *Student*  
line 3: FCAI at Cairo universityline 4: Cairo, Egypt  
line 5: osamaabrahim72@gmail.com

line 1: -  
line 2: *-*  
line 3: *-*line 4: -, -  
line 5: -line 1:-  
line 2: *-*  
line 3: *-*line 4: -, -  
line 5: -

line 1: -  
line 2: *-*  
line 3: *-*line 4: -, -  
line 5: -

*Abstract*—This document presents a comprehensive analysis and implementation of a plant disease classification model using a hybrid architecture combining elements of VGG16 and Inception. The dataset consists of images of healthy and diseased plants, specifically classified into three categories: Early Blight, Healthy, and Late Blight. The proposed model is trained on this dataset and evaluated using various performance metrics. Keywords : VGG16, Disease

# Introduction

Plant disease identification is crucial for effective agricultural management. In this work, we employ a transfer learning approach, combining the strengths of VGG16 and Inception architectures, to create a robust classification model for detecting plant diseases. The dataset includes images of healthy and diseased plants, with a focus on Early Blight, Healthy, and Late Blight categories.

# Literature Review

## Traditional methods

Problem Definition:

The problem of crop disease detection refers to the identification and diagnosis of diseases and disorders that affect agricultural crops. It is a critical issue in agriculture as the health of crops directly impacts food production, crop yield, and the livelihood of farmers. The primary goal of crop disease detection is to enable early and accurate identification of diseases, which can help in the timely application of treatments or preventive measures to minimize crop damage and ensure a healthy harvest. Here's a more detailed problem definition:

1. Identification of Crop Diseases: The first and foremost objective is to identify the presence of diseases or abnormalities in crops. This includes identifying diseases caused by pathogens (e.g., fungi, bacteria, viruses), pests, environmental stressors (e.g., drought, nutrient deficiencies), and other factors that can impact crop health.

2. Disease Classification: Once a disease or disorder is detected, the next step is to classify it accurately. Different diseases may exhibit similar symptoms, making it essential to distinguish between them to determine the appropriate treatment or control measures.

3. Early Detection: Early detection is crucial because it allows for a prompt response to contain the spread of the disease and minimize crop damage. Timely intervention can help save crops and reduce economic losses.

4. Disease Severity Assessment: It's important to assess the severity of the disease. Some diseases may only affect a small portion of the crop, while others can devastate an entire field. Assessing severity aids in decision-making regarding whether and to what extent crop management practices are required.

5. Monitoring and Management: Continuous monitoring of crop health is necessary to track disease progression and the effectiveness of interventions. Crop disease detection should facilitate ongoing management strategies, including the application of pesticides, irrigation adjustments, and other treatments.

6. Precision Agriculture: The use of technology and data-driven approaches, such as remote sensing, drones, and sensor networks, can enhance the precision and efficiency of disease detection and management.

7. Data Collection and Analysis: The collection of relevant data, such as images of crops, environmental conditions, and historical disease records, is essential. Data analysis and machine learning techniques can be employed to process and interpret this data for accurate disease detection.

8. Decision Support: Crop disease detection should provide actionable insights to farmers and agricultural experts, aiding them in making informed decisions regarding disease control and mitigation.

9. Cost-Effective Solutions: Developing cost-effective and scalable methods for crop disease detection is crucial, as this enables broader adoption of the technology among farmers, especially in resource-constrained regions.

10. Education and Awareness: Promoting awareness and education among farmers and agricultural stakeholders about crop diseases, their symptoms, and prevention measures is also an important aspect of addressing this problem.

In summary, crop disease detection involves the timely and accurate identification, classification, and management of diseases and disorders that affect crops, with the ultimate goal of ensuring food security and sustaining agricultural productivity. Technology and data-driven approaches play a significant role in addressing this critical issue in agriculture.

## Dataset description:

Plant disease is a deviation from the normal state of a plant that disrupts or alters its vital functions. Plant diseases can lead to significant yield losses, with estimated global potential losses of up to 16%. As a result, studying plant diseases and developing methods to diagnose and control them is an essential area of research in plant pathology.dataset contain 405 samples for testing and 409 for validation and 3100 for training and the classes is Early Blight and Healthy and Late Blight , and then we used Data Augmentation algorithm for increase number of samples in training samples then data train is now 5150 samples .

The proper identification of plant diseases is crucial for effective control measures, as without them, control efforts can be ineffective and a waste of resources. Image processing algorithms have been developed to detect plant diseases by analyzing the color features of the infected leaves. One such algorithm involves using the K-means method for color segmentation and the Gray-Level Co-Occurrence Matrix (GLCM) for disease classification. This method of vision-based plant disease detection has shown promising results and has the potential to be an efficient and effective tool for disease diagnosis.

To understand the relationship between plant diseases and yield loss, it is necessary to consider the epidemiology of the disease, the physiology of the crop, the yield development, the damage mechanisms involved, and the effect of management practices. By integrating this information, it is possible to improve our understanding of the relationship between plant diseases and crop loss. However, it is important to note that yield loss studies are resource-intensive and can be difficult to interpret, as crops are rarely affected by only one pest or pathogen at a time.

In conclusion, the detection of plant diseases is an important aspect of agriculture, as it is essential for effective disease control and management. Image processing algorithms have shown promising results in the detecdetectingseases, and the integration of various aspects of plant physiology, disease epidemiology, and management practices can help increase our understanding of the relationship between plant diseases and crop loss. The goal of plant pathology research is to reduce yield losses and develop integrated pest management strategies based on economic thresholds, which can be achieved through a better understanding of the relationship between plant diseases and crop loss.

# Methodology

## Model Architecture

The proposed hybrid model comprises a VGG16 backbone followed by an Inception module. The VGG16 component consists of convolutional layers with max-pooling, while the Inception module introduces parallel convolutional paths with different filter sizes. The model is compiled using the Adam optimizer and categorical cross-entropy loss function.

# Data Preprocessing

## Data Acquisition

The dataset is obtained by uploading a compressed file (archive.zip) containing plant images using the Google Colab interface. The uploaded file is then extracted into a designated folder (your\_extracted\_folder) for subsequent processing.

## Data Loading and Labeling

## Images from the training, testing, and validation sets are loaded using the Python Imaging Library (PIL). The dataset is organized into folders based on disease categories, and labels are assigned accordingly. The images are converted to NumPy arrays, forming the basis for the model's input data.

## Data Shuffling

To ensure a balanced distribution during training, the order of samples in the training, testing, and validation sets is randomized using the random.shuffle function.

## One-Hot Encoding:

Categorical labels are transformed into one-hot encoded vectors to facilitate the training of the neural network. The OneHotEncoder from scikit-learn is utilized for this purpose.

# results

## Model Training

The model is trained using the training set, and the training process is monitored using the validation set. The number of epochs and batch size are adjustable parameters to fine-tune the model's performance. The training progress, model summary, and evaluation metrics are displayed throughout this phase. Then we used CNN and VGG16 and Mobilenet and mix with Inception Module and Vgg16 for training the model and used Decision Tree algorithm for train model because is the best algorithm for this problem and take low time in training and Use Ant Colony Algorithm Optimization

For this problem

## Evaluation

The model's performance is evaluated on both the training and testing sets using accuracy as the primary metric. Additionally, a classification report and confusion matrix are generated to provide a detailed analysis of the model's predictive capabilities. The results are visualized using heatmaps for better interpretation.

CNN with Neural Network has accuracy 40% after one epoch and take long time , then use Vgg16 for feature extraction with Decision Tree algorithms have accuracy 73% and using Ant Colony algorithm the increase accuracy 74% , then use MobileNet Archtecture with Decision Tree and with Ant Colony give accuracy 75.5 %

# Discussion

## Limitations and Future Directions

Despite the success of our CNN model, it is necessary to acknowledge its limitations. The model may face challenges with very complex scenes, ambiguous color situations, or cases where ground truth data is scarce. Future research could explore solutions to these challenges, which may include incorporating additional contextual information or advanced attention mechanisms.

## Real-World Applications

Successful Crop disease detection using CNNs opens the door to countless real-world applications. potential impact is wide-ranging. As we continue to refine and improve our model, its application in practical scenarios becomes increasingly promising.

## Impact of Architectural Choices

Ablation studies highlight the contributions of specific architectural choices in our model. The inclusion of skip connections, choice of best model for classification ,structures, and attentional mechanisms has been pivotal in enhancing the model's ability to capture global context and fine detail. Understanding the influence of these components contributes to developing knowledge of the optimal architectural design of crop diseases detection

## Conclusion

The hybrid Inception-VGG model exhibits strong potential for plant disease classification. Its ability to leverage features from both architectures contributes to robust and accurate predictions. Further optimization and exploration of hyperparameters may enhance its performance, making it a valuable tool for real-world agricultural applications.

##### References

1. Review on Techniques for Plant Leaf Classification and Recognition *(references) (*[*https://www.researchgate.net/publication/336707258\_Review\_on\_Techniques\_for\_Plant\_Leaf\_Classification\_and\_Recognition*](https://www.researchgate.net/publication/336707258_Review_on_Techniques_for_Plant_Leaf_Classification_and_Recognition) *(*
2. Rice Crop Disease Detection Using Machine Learning Algorithms *(references) (*<https://www.researchgate.net/publication/374236434_Rice_Crop_Disease_Detection_Using_Machine_Learning_Algorithms> *)*
3. Image-based crop disease detection with federated learning. *(references) (*<https://www.researchgate.net/publication/375417843_Image-based_crop_disease_detection_with_federated_learning> *)*
4. Revolutionizing Agriculture: Machine Learning- Driven Crop Recommendations and Disease Detection in Fertilizer Management. *(references) (*<https://www.researchgate.net/publication/374784563_Revolutionizing_Agriculture_Machine_Learning-_Driven_Crop_Recommendations_and_Disease_Detection_in_Fertilizer_Management> *)*
5. Crop Leaf Disease Detection in Soybean Crop Using Deep Learning Technique. *(references) (*<https://www.researchgate.net/publication/374141414_Crop_Leaf_Disease_Detection_in_Soybean_Crop_Using_Deep_Learning_Technique> *)*
6. Crop Plant Disease Detection Using Deep Learning. *(references) (*<https://www.researchgate.net/publication/373199213_Crop_Plant_Disease_Detection_Using_Deep_Learning> *)*
7. Recent Advances in Crop Disease Detection Using UAV and Deep Learning Techniques . (references) (<https://www.researchgate.net/publication/370620371_Recent_Advances_in_Crop_Disease_Detection_Using_UAV_and_Deep_Learning_Techniques> )
8. A Novel Hybrid Deep Learning Model for Crop Disease Detection Using BEGAN

(<https://www.researchgate.net/publication/371780489_A_Novel_Hybrid_Deep_Learning_Model_for_Crop_Disease_Detection_Using_BEGAN> )

1. Plant Disease Expert . (data) (<https://www.kaggle.com/datasets/sadmansakibmahi/plant-disease-expert> )