Plant Disease

February 2023

1 Define the problem: What is the input and What is the output?

Identification of the plant diseases in order to prevent the losses within the yield. In this project, we have described the technique for the detection of plant diseases with the help of their leaf's pictures.

input — leaves pictures
output — extract the image properties or useful information from the image to choose the high probability class from 38 class.

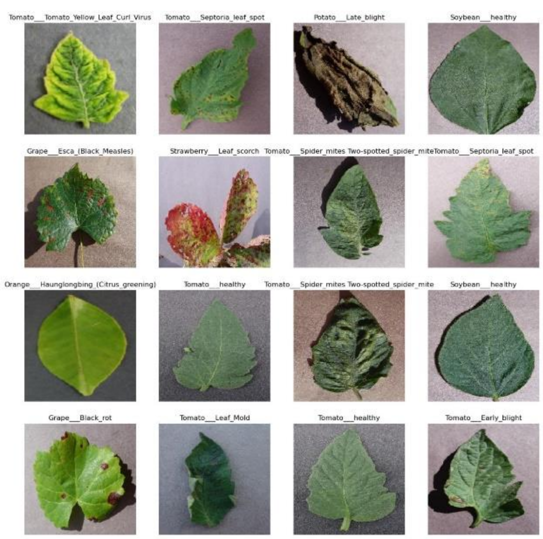

Figure 1- Random Image from train data

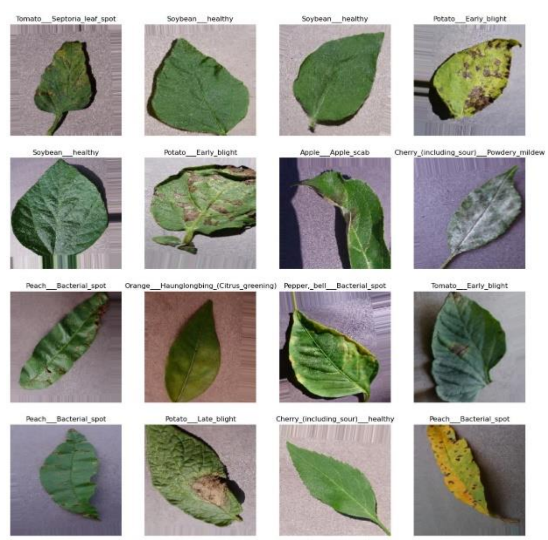

Figure 2- Random Image from test data

2 Why do this? What makes it difficult?

The motivation behind developing a project that detects diseases in plants is primarily to improve crop yield and ensure food security. Plant diseases can cause significant economic losses in agriculture, and detecting them early can help farmers take appropriate measures to prevent the spread of the disease andminimize crop damage.

By using machine learning and computer vision techniques to develop aplant disease detection system, farmers can quickly and accurately identify plant diseases, even in their early stages. This enables them to take action, such as using targeted treatments or adjusting irrigation and fertilization practices, to prevent the disease from spreading and potentially causing crop failure.

Moreover, plant disease detection systems can also reduce the amount of pesticides and fungicides used in agriculture, which can have negative impacts on the environment and human health. With the help of AI and ML-based disease detection systems, farmers can use fewer chemicals and only when necessary, leading to more sustainable and eco-friendly agriculture practices.

Overall, the development of a plant disease detection system using machine learning and computer vision techniques can have a significant impact on improving crop yield, ensuring food security, and promoting sustainable agricul- ture practices.

However, developing a reliable and accurate plant disease detection system is a challenging task due to several factors. Here are some of the reasons why this is a difficult problem to solve:

1- **Large Variability:** Plants can show a wide range of symptoms when they are affected by a disease, and these symptoms can vary depending on the type of disease, the stage of the disease, and environmental factors. This makes it challenging to develop a system that can accurately detect plant diseases under different conditions.

2- **Large Datasets:** Collecting and labeling large datasets of images of healthy and diseased plants can be a time-consuming and costly process. Moreover, the quality of these datasets can affect the accuracy and generalizability of themodels.

3- **Class Imbalance:** Plant diseases are often rare, and in some cases, onlya small percentage of plants in a field may be affected by a disease. This can lead to a class imbalance problem, where the model may be biased towards the majority class (healthy plants) and may not perform well on the minority class (diseased plants).

4- **Environmental Factors:** The performance of a plant disease detection sys-tem can be affected by various environmental factors such as lighting, camera quality, and the presence of other objects in the background. These factors can make it challenging to develop a system that can work reliably under different conditions.

5- **Deployment:** Deploying a plant disease detection system in the field canbe a challenging task, as it requires the system to work in real-world conditions and handle various types of interference.

3 Related work: What has been done? What are the problems?

Firstly, Links for related work papers:

- https://www.frontiersin.org/articles/10.3389/fpls.2016.01419/full
- https://ieeexplore.ieee.org/document/9399342
- https://arxiv.org/ftp/arxiv/papers/2106/2106.10698.pdf

Modern technologies have given human society the ability to produce enough food to meet the demand of more than 7 billion people. However, food secu- rity remains threatened by a number of factors including climate change, the decline in pollinators (Report of the Plenary of the Intergovernmental Science-PolicyPlatform on Biodiversity Ecosystem and Services on the work of its fourth session, 2016), plant diseases, and others. Plant diseases are not only a threatto food security at the global scale, but can also have disastrous consequences for smallholder farmers whose livelihoods depend on healthy crops. The com- bined factors of widespread smartphone penetration, HD cameras, and high-performance processors in mobile devices lead to a situation where disease di- agnosis based on automated image recognition, if technically feasible, can be made available at an unprecedented scale. Here, we demonstrate the technical feasibility using a deep learning approach utilizing 54,306 images of 14 crop species with 26 diseases (or healthy) made openly available through the project.

Neural networks provide a mapping between an input—such as an image of a diseased plant—to an output—such as a crop disease pair.

The occurrence of plant disease has a negative impact on agricultural early detection is the basis for control of plant diseases disease infected plant usually show obvious marks on leaves, stems in most cases, agricultural experts is used to identify on sit or farmers this method is not subjective use of image processing techniques has become a hot reach topic in recent years deep

learning technology in the study of plant disease made more progress we have introduced the basic

knowledge of deep learning recognition using deep learning provided suffi- cient data for training In most of the researches, the PlantVillage dataset was used to evaluate the performance of the DL models. Although this dataset hasa lot of images of several plant species with their diseases, it was taken in the lab. Therefore, it is expected to establish a large dataset of plant diseases inreal conditions.

4 Algorithm: What did you try? What are thealternatives? Why you choose to try this?

In our project we have been used EfficientNet B0, is a deep neural network architecture that has been shown to achieve state-of-the-art performance on image classification tasks while being computationally efficient. The architecture uses a combination of convolutional layers and squeeze-and-excitation blocks to efficiently learn features from input images. There are several alternatives to EfficientNet B0, such as ResNet, Inception, and VGG architectures. These architectures have also shown to achieve high accuracy on image classification tasks, but they may require more computationalresources than EfficientNet B0. Therefore, the choice of architecture depends on the specific requirements of the task, such as the available computational resources and the desired accuracy.

In the case of plant disease detection, EfficientNet B0 may be a suitable choice as it can efficiently learn features from input images and achieve high accuracy on image classification tasks. Moreover, EfficientNet B0 has been shown to generalize well to different datasets and can be fine-tuned for specific tasks with a small amount of data. Therefore, EfficientNet B0 can be a good choice for developing a plant disease detection system that can be deployed in the field with limited computational resources. However, the choice of algorithm also depends on other factors such as the size and quality of the dataset, the distribution of plant diseases in the dataset, and the availability of labeled data. Therefore, it is essential to carefully considerthese factors when choosing an algorithm for plant disease detection.

Also, the alternative solution we have been used DNN (Deep neural network)Deep neural networks without transfer learning involve training a deep learning model from scratch on a specific dataset for a particular task. This means that the model learns to extract features directly from the input images during the training process.

One advantage of training a deep neural network without transfer learningis that the resulting model can be more specific to the task at hand. Since the model is trained from scratch on the specific dataset, it can learn features that are specific to the characteristics of the input images and the task. This can result in higher accuracy than using a pre-trained model in some cases.

However, training a deep neural network without transfer learning can be challenging, especially when dealing with limited datasets. The model may require a large amount of labeled data to learn meaningful features from the input images. Additionally, the training process can be computationally expensive and time-consuming, especially for deeper models.

Therefore, the choice of using deep neural networks with or without transfer learning depends on the specific requirements of the task, such as the size and quality of the dataset, the available computational resources, and the desired accuracy. In the case of plant disease detection, if a large labeled dataset is available and computational resources are sufficient, training a deep neural network from scratch may be a viable option to achieve high accuracy. However,if the dataset is limited, transfer learning may be a better option to achieve reasonable accuracy with a smaller amount of labeled data.

5 Some visualization of our result: screen capture, plot, example outputs

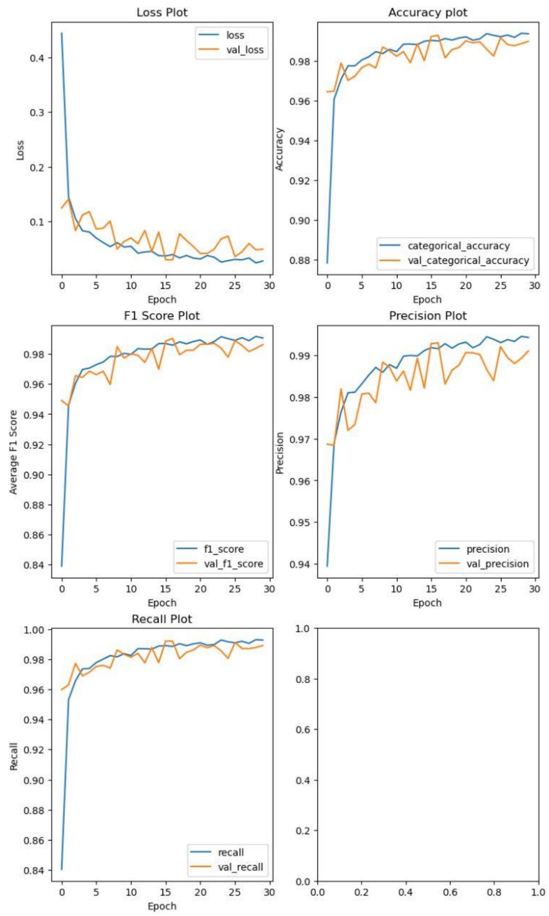

Figure 3- Model evaluation

6 Evaluation: Does it work? How well it works?

Based on the visualization we provided above, our model seems to be performing very well on the validation dataset. The low loss and high accuracy, F-1 score, precision, and recall all indicate that the model is able to accurately classify images of plant diseases. However, it is important to note that the performance on the validation dataset may not necessarily generalize to new, unseen data. Additionally, it is important to consider the specific requirements of the task when evaluating the performance of the model. For example, in the case of plant disease detection, false negatives (misclassifying a diseased plant as healthy) could be more damaging than false positives (misclassifying a healthy plant as diseased). Therefore, precision and recall may be more important metrics to consider than accuracy alone.

Overall, based on the information provided, our model seems to be perform-ing well on the validation dataset, but further evaluation on a test dataset and consideration of the specific requirements of the task may be necessary to fully assess its performance.

7 Analysis: Why it works or doesn't work? Canwe make it better?

Our model worked very well in validation set because we do everything that makes our model good due to some reasons (we have used transfer learning , fine tuning , and data ....) but the problem for our project only it cannot detect multiple diseases in the same image. This limitation can pose a challenge for real-world applications where multiple diseases may affect the same plant orcrops.

For instance, if a single plant or crop is affected by multiple diseases, your model may only detect one of them and miss the others, leading to incorrect diagnoses and ineffective treatments. This can potentially result in decreased crop yield and economic losses. To overcome this limitation, we could explore approaches such as

multi-label classification or object detection, which can detect multiple diseases or objects in the same image. These approaches can enable our model to detect and identifyall the diseases affecting a single plant or crop, providing a more accurate and comprehensive diagnosis.

However, implementing these approaches may require significant modifications to our current model architecture and training pipeline, as well as potentially larger and more diverse datasets. Nonetheless, overcoming this limitation could significantly improve the practical value and impact of our plant disease detection project.

8 Contribution division: Who did what?

However, implementing these approaches may require significant modifications to our current model architecture and training pipeline, as well as potentially larger and more diverse datasets. Nonetheless, overcoming this limitation could significantly improve the practical value and impact of our plant disease detection project.

Data preprocessing and Report:

1. Kholoud Walid Ebrahim Elsherbeny
2. Sohaila Mohammed
3. Ebrahim Mohammed Shabaan
4. Gasr Mohammed

Build model:

1. Amgad Mohammed Mohammed Shalaby
2. Abdelrahman Ahmed Fathi Mohammed Elhanafi
3. Ahmed Mohammed Abdelkder