Plant Disease Detection

Nikhil Garuda
Dept of CSE
PES University,EC Campus
Bengaluru,Karnataka
nikhilgaruda6@gmail.com

Abstract—Plant disease detection is the key to help in preventing agricultural losses. Every year 14.1 percent of agricultural produce is affected due to plant diseases. Monitoring the health of the plant is very critical for sustainable agriculture.This paper discusses the various methods involved in plant disease detection using their leaves images.

Index Terms—component, formatting, style, styling, insert

I. Introduction

As farmers and agriculture fields are an important part of our lives, farmers are the root-level building blocks in the economy of any country. They work really hard for a whole season to grow a specific crop for the survival of their family. Sometimes the crops to which they dedicated, the whole 3-6 months to nurture these crops get diseases as a result of which they can't sell their crops at the price they except.

Knowing the plant disease beforehand, by the use of specific pesticides and fertilizers, these diseases can be prevented from aggravating. Using deep learning techniques we can help farmers know about a specific disease so that they can be ready before harvesting their crops. This can be done using Resnet50.

II. Literature Survey

One notable work includes ”Plant Disease Detection and Classification by Deep Learning ”[1] .This work provides a comprehensive overview of the development and application of deep learning (DL) methods for plant disease detection and classification, focusing on advancements made from 1943 to the present day, particularly after 2012. The paper examines various classification accuracy (CA), precision, recall, F1 score, and other performance metrics, as well as visualization strategies and different deep learning (DL) models like AlexNet, GoogLeNet, and ResNet for identifying and classifying different plant diseases. Using the PlantVillage dataset, the study shows that having sophisticated deep learning (DL) models that can adjust to different environmental situations is really important. However, it also points out that some datasets use simple backgrounds which might not show the real conditions needed for plants to be identified correctly.

The paper also talks about areas where more research could be done in the future. One of these is to have more realistic experiments and another is to develop even better Deep Learning models that are made just for specific environments.

Another significant work conducted in this field is ”Plant Disease Detection and Classification Using Deep Neural Networks” [2].This paper provides a summarized discussion of processing deep learning approaches in particular YoloV3 and ResNet18 models for the contribution of disease detection and categorization of plants. The study does both by highlighting the important role of the early diagnosis and different control measures in crop loss prevention as to the damaging effect plant diseases could make on yields. These experiments used versions of deep neural networks such as ResNet-50, GoogleNet, and VGGNet-16 for classifying different plant diseases with performance figures differing among other architectures. Science has found out application of CNN, Caffe framework as well as back propagation for training plant disease classifiers keeping in mind model selection and training techniques more profoundly. This experimentation has proved assignment's feasibility with special care given to the creation of a pre-trained model via transfer learning with ResNet18 to fine-tune specific layers for the features of the specific task. It systematizes the insights gained from using the PlantVillage dataset, 29 classes with over 36,000 images, as an illustration of how DL approaches can be leveraged in the working environments. Through YOLOv3 object detection to extract leaves and consecutive ResNet18 models to conduct standard analysis of diseases in various plants, data preprocessing and augmentation strategies put together a strong program to categorize diseases in various species.

III. Proposed Methodology

A. Dataset Preparation

Dataset Preparation: The dataset is divided into an 80/20 ratio of training and validation sets. There are 38 classes of plant disease images , approximately 1700-1800 images in the training set and 17572 images in the validation set. Each size of image being (256, 256, 3).

B. Dataset Preprocessing

Basic image augmentation techniques such as rotation, scaling, and flipping have been implemented, ensure that these

augmentations are applied during training to further enhance the diversity and robustness of the dataset. Integration of Gaussian blur as an additional preprocessing step helps reduce noise and emphasize important features associated with plant diseases.

C. Transfer Learning with ResNet-50

Using transfer learning with the pre-trained ResNet-50 model. The original output layers of ResNet-50 are replaced with a Global Average Pooling layer which is followed by a Dense layer consisting of 128 units and ReLU activation, and lastly , a Dense layer with 38 units for classification.

D. Model Training

The model is trained on the modified ResNet-50 model using the prepared training dataset. Implementation of callbacks such as EarlyStopping, ModelCheckpoint, and ReduceLROnPlateau to prevent overfitting, saving the best-performing model, and adjusting the learning rate during training, respectively.

E. Model Evaluation

Evaluation of the performance of the trained model is done on the validation set using metrics such as accuracy, precision, recall, and F1-score.

F. Fine-tuning and Optimization

Fine-tune hyperparameters such as learning rate, batch size, and optimizer choice based on the model's performance on the validation set. Monitoring for any signs of overfitting and adjusting regularization techniques as necessary.

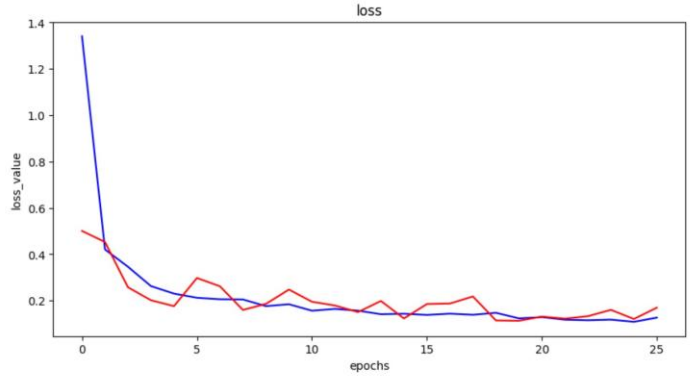
G. Testing and Deployment

Testing the final model on a separate test dataset to assess its generalization ability. Deploy the model for real-world plant disease detection applications, ensuring ethical considerations such as data privacy and model bias are addressed.

IV. Results

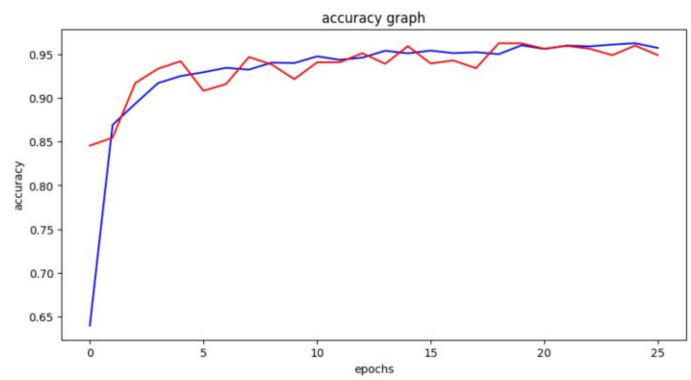
A. Loss Function

The loss function depicts a line graph, which shows the relationship between a machine learning model's training process and its performance. X-axis (horizontal): This axis represents the number of epochs. An epoch signifies one complete pass through the entire training dataset. So, the higher the value on the X-axis, the more times the model has been trained on the complete data. Y-axis (vertical): This axis represents the loss value. The loss value is a numerical metric indicating how well the model's predictions align with the actual data. In simpler terms, a lower loss value signifies better model performance. By analyzing the lines on the graph: A flat line, suggests the model's performance has reached its optimal performance for the given data. The model has been trained on 30 epochs, for which early stopping has occurred at 26 epochs. The loos function values after the final epoch is 0.1265.



B. Accuracy

The accuracy function depicts a line graph showing the relation of a machine learning model over a course of training epochs. X-axis (horizontal): An epoch signifies one complete pass through the entire training dataset. So, the higher the value on the X-axis, the more times the model has been trained on the complete data. Y-axis (vertical): Model accuracy is a performance metric that reflects how often the model makes correct predictions. By analyzing the line on the graph the accuracy value has reached 0.96233



V. Conclusion

A plant disease identification system in which we implemented augment the dataset further and improve plant disease detection accuracy, we introduced Gaussian blur and achieved an accuracy of 96 percent was developed ,and Loss is 0.2.The implementation of more accurate deep learning systems could help in better diagnosis of diseases in crops as these analyze down to the smallest unit of an image, a pixel. This level of detail cannot be analyzed using the naked eye. blur

VI. Future scope

Considering the number of epochs longer than the current 26 would give the model much more freedom to seek deeper patterns in the given data and further refine its representations. And thus, having a wider plant species embedded in the dataset will help the model to be well diversified and learn to generalize its capabilities across different plants. This broader dataset could be the basis for the model to understand different disease patterns more accurately, which will eventually lead to the deployment of the project in the real world.

Although we recognize the possible benefits of avoiding additional trials that can be done, the scope of our work is limited by computing constraints. This drawback impedes us a wide variety of epochs or spending hours on experiments with an aim to improve the algorithm.

In terms of handling these issues, the future research could be all about optimizing algorithm efficiency using compression techniques and the use of distributed training. Besides, investigation of this transfer learning approach, or creation of more efficient neural networks that will tackle this disease prediction in a more effective way could become, actually, the most efficient for an accurate outcome.

This could further be enhanced by introducing object detection technique on the images of the plants.

References

[1] Saleem MH, Potgieter J, Mahmood Arif K. Plant Disease Detection and Classification by Deep Learning. Plants (Basel). 2019 Oct 31;8(11):468. doi: 10.3390/plants8110468. PMID: 31683734; PMCID: PMC6918394.

[2] Venkataramanan, Aravindhan Agarwal, Pooja. (2019). Plant Disease Detection and Classification Using Deep Neural Networks.