Image-based Disease Detection in Bell Pepper Plants using CNN

This repository/folder contains code for training a Convolutional Neural Network (CNN) to detect and classify diseases in potato plants based on leaf images.

Dataset Details

The dataset used in this project is the "PlantVillage" dataset, which consists of RGB images of healthy and diseased potato plant leaves of the size 256*256 pixel. The dataset contains 2, 1 for Pepper_bell_healthy and 1 for Pepper_bell_Bacterial_spot diseases leaves images.

Dependencies

The code is implemented in Python using the TensorFlow library. Make sure you have the following dependencies installed:

- TensorFlow
- Matplotlib
- Numpy
- Scikit-learn
- Seaborn

Usage

Install the required dependencies:
“ pip install (dependencies) ”
Type this code to your terminal or jupyter notebook or any other IDE or in Anaconda Navigator.

Run the code:

You can run the code by opening the file named “*Pepper_Leaf_Disease_Classification-Final.ipynb*” in any software like Jupyter Notebook, Google Colab, VS Code or any other Python IDE after installing the required dependencies to your system.

- The code will preprocess the dataset, split it into training, validation, and testing sets, and train the CNN model.
- After training, the code will evaluate the model's performance on the testing set and display the test loss and accuracy.
- Additionally, it will plot the training and validation accuracy and loss curves for analysis.
- The code will also generate a set of sample predictions on the testing set, displaying the actual and predicted classes along with the confidence level.
- Finally, a confusion matrix and classification report will be generated to assess the model's performance in classifying different potato leaf diseases.

Model Saving

The trained model will be saved in the project directory with the name *"pepper.h5"* with some extra files related to this model in folder named *“model_1”*. It can be loaded and used for making predictions on new data using FastAPI.

Result

For this model we have tried different combination of train-test split like 70:30, 80:20 and few more. We save the output like accuracy and loss graph, confusion matrix and statistical values of different combination with folder named *“60train,40test Result”, “70train,30test Result”, “75train,25test Result” and “80train,20test Result”.*

Conclusion

This code provides a pipeline for training a CNN model to detect and classify diseases in potato plants based on leaf images. By using this code as a starting point, you can further enhance the model's performance by adjusting hyperparameters like increasing or decreasing the batch size, epochs or adding some regularization like L1 or L2, exploring different CNN architectures, or applying techniques like data augmentation.

Contributors

Utpal Kant Singh (utpalkant.4855@gmail.com)