Coding task

This coding task aims to train and evaluate a neural network model (architecture of your choice, PyTorch or TensorFlow) that classifies images of oranges into four different classes including three diseases (citrus canker, black spot and greening citrus) and fresh oranges.

The dataset is provided here: https://www.kaggle.com/datasets/jonathansilva2020/orange-diseases-dataset. There is one working solution on the kaggle webpage but we expect only original submissions.

Questions to answer in your solution.

- **Training description.**
  - Please motivate the choice of model architecture, training hyperparameters and metrics.
  - Use plots and text to support your claims.
- **Model evaluation.**
  - Please evaluate your trained model on the separate test data set, which is provided by kaggle and motivate the choice of the evaluation metric(s) you use.
- **Real-world application.**
  - In a real-world application, model size matters. Please provide a compressed version of your already trained model without significant loss in model performance (e.g., pruning, quantization, ….)

Additional questions:

Note that this additional task is not necessary!! However, if you would like to do an extended analysis of your model, you can proceed with one or both of the following tasks:

- **Model explanation.**
  - You can explain the difference between the model's decision-making strategies between "disease" classes and the healthy ones. (Hint: This could be done by analysing the results of local XAI methods (e.g., using captum, innvestigate).
- **Object detection.**
  - You can investigate ways to automatically detect the objects such as oranges or leaves (if possible) using any preferred Python library for automatic object detection.

Further instructions.

Please submit your solution (written in Python) as a zipped folder which includes a notebook of your coded solution (please specify the package requirements for how to run it). You may write auxiliary source code files but it is not mandatory. You are free to use the ML packages and frameworks such as TensorFlow and PyTorch of your own choice. Please use the Markdown cells to state your analysis clearly. Please name your submission according to your name i.e., "FIRSTNAME_LASTNAME.zip" and state the "ATB Coding Task 2023" as the email Subject.

We expect this task to finish in 2-4 hours. Please indicate the time you spent on the task (at the top of the notebook), including if you spent more time.

We value original, creative solutions!

Please send your solution as a response to this email at the latest 9th of March at 18.00 CET.

Lastly, please note that the coding task should be self-explaining, where we cannot give any further guidance or information (as this would create an unfair advantage).

Thank you for your time and have fun!