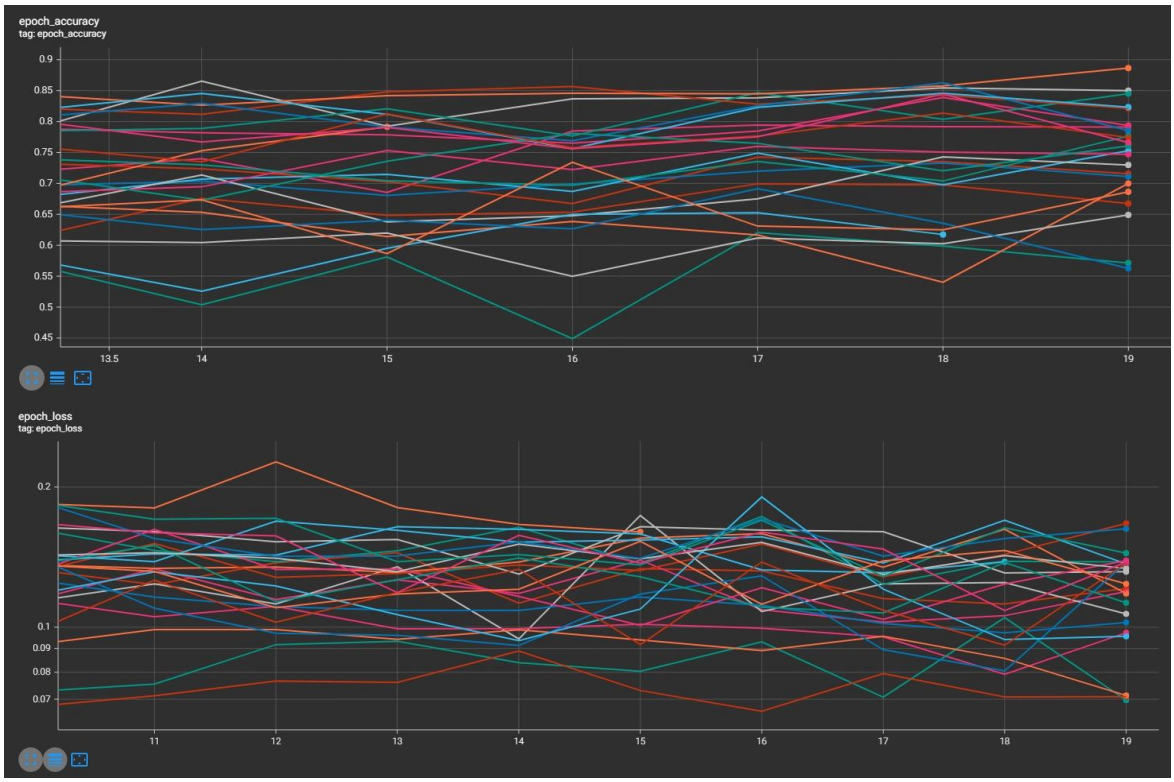
REPORT ANNDL2021_Homework1

TEAM: Cooper
- Davide Canali
- Filippo Caliò
- Natalia Bagnoli

We started our work trying the easiest solution (3 convolutional layer and 2 dense layer) to verify if everything worked properly.

To avoid overfitting, we decided to use the early stopping technique.

Through for loops, we tried 25 different configurations of the network to better understand the optimal quantity of convolutional layers, dense layers, and neurons/filters.
Thanks to the use of Tensor Board, we analyzed results taken on the validation test, inferring that the configurations that brought better results were the ones with more convolutional layers and 0 or 1 dense layer.
Therefore, one established this solution, instead of the Flattening technique we decided to apply the Global Average Pooling operation but, however, we do not obtain results significantly different from the previous ones.

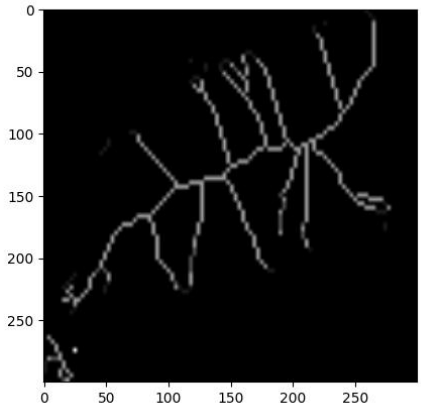


Based on the graphs the 3 best overall results were those with:
- 3 convolutional with 128 filters and no dense layer.
- 3 convolutional with 128 filters each and an Average Global Pooling at the end.
- 3 convolutional with 32 filters each and 1 dense layer with 128 neurons.

As a result, we applied the Transfer Learning technique using firstly VGG16 and RESNET architectures and then Xception which proved to be one of the best convolutional neural network models for this task.
We tested the model more times training only some layers (the last 10, 20, 30) but we found that the best results are obtained by retaining the whole model.

Noting that the dataset was unbalanced (5693 files for tomato images and 264 for raspberry images), we normalized it decreasing the size and keeping 500 images for each class but even this solution brough almost the same results of the techniques used previously.

Other solutions that did not bring notable changes were the addition of noise to data and the use of different preprocessing: we no longer used only rotation, inversion, and brightness changes but we tried to light more significant features of the leaves. Though several studies and researches we found that an operation widely used for this type of images is highlighting only the veins of the leaves replacing the red channel of the image with the skeleton.



One of the latest ideas was to operate with the Edge Enhancement filter to increase the importance of the shape of the leaves. We also tried to crop and center images by finding the first and last column and row that were not black and used that index as the coordinate for the bounding box of the cropped image.

But even those last ideas didn’t bring any improvements and so our best model was a fully retrained Xception on the entire dataset.