RICARDO B. SOUSA (UP201503004@FE.UP.PT | RICARDO.B.SOUSA@INESCTEC.PT)

Title: Herbarium 2021 – Half-Earth Challenge – FGVC8

URL: https://www.kaggle.com/c/herbarium-2021-fgvc8/overview

Student: Ricardo Barbosa Sousa (up201503004@fe.up.pt | ricardo.b.sousa@inesctec.pt)

Dataset: https://www.kaggle.com/c/herbarium-2021-fgvc8/data (training and test set containing more than 2.5M images from approximately 65000 species of vascular plants – standardised plant list LCVP v1.0.2)

Abstract:

There are approximately 3000 herbaria world-wide that represent massive repositories of plant diversity data. All specimens included in these repositories not only maintain their morphological features but also include collection dates and locations, the name of the person who collected the specimen, and their reproductive state. This information provides the framework for understanding plant diversity on a massive scale and learning how it has changed over time.

So, the competition Herbarium 2021 (FGVC8) represent the first step to speed the discovery of new species. Given that there is an overwhelming amount of unnamed and new plants specimens, the main goal is to classify unidentified plant specimens in the dataset provided by the New York Botanical Grarden, Bishop Museum, Naturalis Biodiversity Center, Queensland Herbarium, and Auckland War Memorial Museum.

Expected Software to develop

- Pre-process and analyse the dataset
- Create machine learning models based on deep convolutional neural networks
- Training the implemented models
- Testing the implemented models
- Compare the performance of different models

Reference Papers:

Rawat, W., Wang, Z.: Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review. In: Neural Computation **29**(9):2352–2449 (2017). doi: https://doi.org/10.1162/neco_a_00990

Freiberg, M., Winter, M., Gentile, A., Zizka, A., Muellner-Riehl, A.N., Weigelt, A., Wirth, C.: LCVP, The Leipzig catalogue of vascular plants, a new taxonomic reference list for all known vascular plants. In: Scientific Data **7**(416) (2020). doi: https://doi.org/10.1038/s41597-020-00702-z

Barré, P., Stöver, B.C., Müller, K.F., Steinhage, V.: LeafNet: A computer vision system for automatic plant species identification. In: Ecological Informatics **40**:50–56 (2017). doi: https://doi.org/10.1016/j.ecoinf.2017.05.005