MSc. in Computing
Practicum Approval Form

Project Title:	Automatic Classification of Nutritional Deficiencies in Coffee Leaves Images Using Transfer learning
Student ID:	23267908, 23266257
Student name:	Sarath Mukundan Adavakkat, Suhara Fathima Shahul Hameed
Student email	sarath.mukundanadavakkat2@mail.dcu.ie, suhara.shahulhameed2@mail.dcu.ie
Chosen major:	Data Analytics
Supervisor	Claudia Mazo
Date of Submission	24-11-2023

What is the topic of your proposed practicum?

Peruvian coffee has become well-known for its high-quality, specialty coffee. Peru is now the second largest exporter of organic coffee beans [6]. Coffee production in Peru significantly strengthens rural employment, providing vital opportunities for farmers and agricultural workers. Its robust export volume enhances the country's trade balance, while the cultivation of coffee acts as a crucial driver in elevating rural incomes [7].

Several important nutrients are required for the healthy growth of coffee plants, and deficits in these minerals can manifest apparent symptoms. These deficiencies manifest in various forms, including leaf discolourations, stunted growth, limited fruiting, and twisted leaves, resulting from insufficient quantities of critical nutrients such as nitrogen, phosphorus, potassium, magnesium, iron, calcium, sulfur, zinc, and boron [8].

Coffee crop productivity can deteriorate as a result of delayed disease identification in crops. Detecting nutritional deficits in coffee plants requires paying close attention to leaf symptoms such as yellowing, shape deformation, discoloration, as well as abnormal growth patterns which is quite a challenging and laborious task for farmers on a daily basis as its time consuming process to analyse and take corrective actions is required. Early identification gives farmers the foresight that they need to make informed decisions, boosting effective resource management and increasing overall output in coffee cultivation.

Our proposed practicum project focused on the automatic detection of nutritional deficiencies such as potassium, boron, calcium, and iron in coffee plants using transfer learning techniques and data augmentation strategies.

Please provide details of the papers you have read on this topic

1. Tuesta-Monteza, V.A., Mejia-Cabrera, H.I. and Arcila-Diaz, J. (2023). CoLeaf-DB: Peruvian coffee leaf images dataset for coffee leaf nutritional deficiencies detection and classification. Data in Brief, [online] 48, p.109226. doi:https://doi.org/10.1016/j.dib.2023.109226. [link]

2. Monsalve, D., Trujillo, M. and Chaves, D. (2015). Automatic Classification of Nutritional Deficiencies in Coffee Plants. doi:https://doi.org/10.1049/ic.2015.0317. [link]

3. Vassallo-Barco, M., Vives-Garnique, L., Tuesta-Monteza, V., Mejía-Cabrera, H., Raciel, Y. and Toledo (2017). Automatic Detection of Nutritional Deficiencies In Coffee Tree Leaves Through Shape And Texture Descriptors. Journal of Digital Information Management, [online] 15. Available at: https://www.dline.info/fpaper/jdim/v15i1/jdimv15i1_2.pdf [Accessed 15 Nov. 2023]. [link]

4. Kumar, M., Gupta, P., Madhav, P. and Sachin (2020). Disease Detection in Coffee Plants Using Convolutional Neural Network. 2020 5th International Conference on Communication and Electronics Systems (ICCES). doi:https://doi.org/10.1109/icces48766.2020.9138000. [link]

5. Dutta, L. and Rana, A.K. (2021). Disease Detection Using Transfer Learning In Coffee Plants. doi: https://doi.org/10.1109/gcat52182.2021.9587602. [link]

6. UNDP. (n.d.). Coffee Sector in Peru Infographic | United Nations Development Programme. [online] Available at: https://www.undp.org/facs/publications/coffee-sector-peru-infographic. [link]

7. ElevaFinca (2023). Peruvian Coffee: Economic Importance, Sustainable Practices, and Market Demand. [online] ElevaFinca. Available at: https://www.elevafinca.com/post/peruvian-coffee-economic-importance-sustainable-practices-and-market-demand. [link]

8.Vikaspedia.in. (2022). vikaspedia Domains. [online] Available at: https://vikaspedia.in/agriculture/crop-production/integrated-pest-managment/ipm-for-commercial-crops/ipm-strategies-for-coffee/coffee-nutritional-deficiencies-disorders. [link]

How does your proposal relate to existing work on this topic described in these papers?

In [1] an automatic diagnosis of nutritional deficiencies in coffee leaves was proposed using a naive Bayes classifier another classifier based on neural networks on the shape and colour of coffee leaves. In [2] a Random Forest approach to automatically diagnose nutritional deficits in coffee plants using global and local features derived from leaves is conducted. In [3] an automatic detection of nutritional deficiencies in coffee tree leaves through shape and texture descriptors in KNN, Naive Bayes, and Neural Network classifiers is analysed. In [4] disease detection in coffee plants was suggested using convolutional neural network by Transfer learning (using InceptionV3 model). In [5] disease detection using transfer learning

in coffee plants was put forward by a combination of deep learning and MobileNetV2 which is a powerful transfer learning model.

Our research proposal with reference to previous work on detecting and classifying nutritional deficiencies in coffee plants compares different transfer learning models using MobileNetV2, InceptionV3, VGG19, and EfficientNetV2 architectures with the CoLeaf dataset.

- What are the research questions that you will attempt to answer?

In this research project focusing on nutritional deficiencies in coffee plants, we are attempting to answer the follow research question:

- Is it possible to classify early nutritional deficiencies (e.g. Potassium, Boron, Calcium and Iron) in Peruvian coffee plants using transfer learning techniques?

- What software and programming environment will you use?

- Google Colab, short for Google Colaboratory, is a Google cloud-based platform that provides a free, collaborative, and Jupyter Notebook-based environment for running Python programs. NumPy, Pandas, Scikit, Seaborn, and Matplotlib are also used.

- What coding/development will you do?

- Our preprocessing steps include data augmentation such as rotation, scaling, and shearing. We will utilize the pre-trained Transfer learning architectures for training in Python.

- What data will be used for your investigations?

- The CoLeaf dataset contains 1006 leaf images grouped according to their nutritional deficiencies which were collected from Peru.
- The dataset consists of the following classes with No. of images:

  Healthy leaf -6, Nitrogen (N) -64, Phosphorus (P) -246, Potassium (K) -96, Magnesium (Mg) -79, Boron (B) -101, Manganese (Mn) -83, Calcium (Ca) -162, Iron (Fe) -65 and more than one deficiency -104 [link].

- Is this data currently available? If not, where will it come from?

- We have access to a public dataset in the Mendeley data dataset.

- What experiments do you expect to run?

- As the dataset is limited, we plan to conduct the training on 70% of the dataset, and the remaining 30% will be used to test the models.

- What output do you expect to gather?

- Based on the analysis performed, we plan to compare different transfer learning models for detection of nutritional deficiencies such as potassium, boron, calcium, and iron in coffee plants at an earlier stage.

- How will the results be evaluated?

The result is the model's accuracy, precision, F1-score, and recall for prediction which can be analysed by comparing it with the existing works.