Coffee Plant Nutritional Deficiencies Classification Using Transfer Learning

MSc Research Project
Data Analytics

Reyna Vargas Antonio
Student ID: x23127635

School of Computing
National College of Ireland

Supervisor: Paul Stynes

Coffee Plant Nutritional Deficiencies Classification Using Transfer Learning

Reyna Vargas Antonio X23127635
MSc in Data Analytics
National College of Ireland

Abstract

The production of coffee plants can be affected by different factors such as pests, diseases, and nutritional deficiencies for that reason, the early identification of any of these factors is important. These factors can be detected through leaves, stems, or berries, however, if the plant has any spot or affected area, it is usually related to a disease, leaving behind the lack of nutrients such as boron, calcium, iron, magnesium, manganese, nitrogen, phosphorus, potassium and others. Implementing deep learning methods that can classify leaves of coffee plants according to nutritional deficiencies is a challenge when it comes to with a small dataset. The proposed research is designed on using Convolutional Neural Networks-based model and different Transfer Learning Techniques, which allow to farmers provide the sufficient nutrients to coffee plants and prevent low quality of production. This research uses a combination of two datasets, CoLeaf dataset includes 1,290 images split into ten categories of boron (B), Calcium (C), Iron (Fe), Magnesium (Mg), Manganese (Mn), Nitrogen (N), Phosphorus (P), Potassium (P), Others and Healthy leaves. No disease file was added to this research from Coffee Plant disease dataset. Data Augmentation were implemented, followed by five transfer learning techniques for training models such as VGG-16, ResNet-50, DenseNet-121, MobileNet-V2 and Inception-V3. Results of the five models are presented in this paper based on accuracy and loss. At first instance this model shows best performance in VGG-16 and ResNet-50 in training dataset with an accuracy of 99.16% and 95.85% respectively, however , fine-tuned model displays better performance in ResNet-50 with an accuracy of 99.83% whilst VGG-16 obtains 99.69% of accuracy.

Keywords: Coffee Plant, Nutritional Lacks, Deep Learning, CNN, Transfer Learning Architecture.

1. Introduction

According to The World Bank, the Agriculture represent 4% of global gross domestic product (GDP) , its development means less risk of poverty and boost shared property due to it is estimated that in 2050, the population will be of 10 billion people[1]. However, there are multiple factors, which do not allow a growth in this sector, such as the impact of climate change, inappropriate treatment of farms, diseases or pests infection in the crops. In the last years, deep learning approaches have been developed in Agriculture for improving

[1] https://www.worldbank.org/en/topic/agriculture/overview

the productivity of crops and preventing risk (Singh Patel, 2021) by means of detecting and classifying coffee plant diseases. Usually, when a plant is infected by a disease or pest, their symptoms start to be visible marks mostly on their leaves, flowers or fruits, these symptoms allows to identify the type of disease or pest, based on the pattern or abnormality in the crops (S, 2022) . On the other hand, these symptoms could be sing of nutritional lack instead of a disease. Usually, coffee cultivation is produced by smallholder farmers, who have limited technology, resources and information, additionally, they use basic or general nutrients without any studied previous to land, resulting in only caring of presence of pest or disease.

The identification of nutritional deficiencies is the purpose of this research, which helps smallholder farmers prevent any disease and provide sufficient nutrients from the beginning of cultivation and if it is necessary during its growth, bringing with itself good quality of production. The implementation of a deep learning framework that uses image recognition and classification to identify the nutrients that are poor in coffee plants through their leaves is the primary contribution of this research. Having as resource for this research, the dataset CoLeaf that includes 1,290 Peruvian coffee leaf images divided into ten classes of Boron (B), Calcium (C), Iron (Fe), Magnesium (Mg), Manganese (Mn), Nitrogen (N), Phosphorus (P), Potassium (K), others and Healthy, this files show the nutritional deficiencies detection and classification. A model was chosen for the deep learning framework by comparing accuracy and loss. In order to provide support to smallholder farmers with new resources and technology for better production of their crops, it is essential to inform them of the use of these models through training, with the aim that they can feed the models by themselves and improve the accuracy.

This paper consider deep learning models applied for classification of different nutritional deficiencies and healthy leaves in coffee plants in Section 2 related work. The research methodology and specifications is explained in Section 3. The results obtained in this research are discussed and displayed in the Section 4. In Section 5 concludes the research.

2. Related Work

Coffee is among the most popular non-alcoholic drinks around the World and the second most-exported commodity, values for its fragrance and caffeine content. Coffee is grown in over 70 tropical countries, being Brazil, Vietnam and Colombia the principal leads in production[2]. However, the quality of the coffee depends of many factors that can differ in each countries, such as the type of disease that affect the plantation. Another factor that can influence the taste and quality of the coffee is the balance of nutrients or minerals of the land where are cultivated the coffee trees. Recent advancements in deep learning, particularly transfer learning and convolutional neural networks (CNN), have demonstrated their potential to address these challenges by detecting nutrient deficiencies in coffee plants by the use of image classification technique. Transfer Learning, a technique that adapts pre-trained models to new tasks (Brodzicki, 2020), has proven essential in Agriculture, particularly where labelled datasets are sparse. By leveraging models pre-trained on large datasets like ImageNet, researchers have reduced the load of data collection in agriculture, where gathering labelled data is costly and time-consuming. In nutrient deficiency detection, this technique has shown promise for a variety of crops, including maize, tomato, rice and coffee.

Several studies have demonstrated the effectiveness of **transfer learning** in detecting nutrient deficiencies in crops like maize and tomato. It has been utilized transfer learning models like DenseNet-201, VGG-16,

[2] https://worldpopulationreview.com/country-rankings/coffee-consumption-by-country

ResNet-50, GoogLeNet and MobileNet-V2 to detect phosphorus deficiency in maize plants (Ramos-Ospina, 2023), where DenseNet-201 achieved the highest accuracy of 96%, showcasing how transfer learning can adapt to models to agricultural tasks with minimal data. Following the use of transfer learning, an investigation related to classify nutrient deficiencies in rice plants using leaf images, but with the difference that the researchers worked with hydroponic experiments (Xu, 2020), which involves growing plants without soil by using nutrient-rich water solutions instead, collecting images that classify into ten types of nutrient deficiencies. This paper evaluated four state-of-the-art like Inception-V3, ResNet-50, NasNet-Large and DenseNet-121 architectures, achieving the last model the highest accuracy with 97.44%, the paper also identifies challenges such as misclassification among deficiencies that exhibit similar symptoms as iron and manganese deficiencies. Similarly, it was applied **transfer learning** for calcium and magnesium deficiency detection in tomato plants (Kusanur, 2021), using a combination of VGG-16 and Support Vector Machine (SVM) classifier to performance a 99.14% accuracy, while Inception-V3 with Random-Forest attained 98.71%. The pairing of CNN models with traditional machine learning classifiers, highlighting the utility of **hybrid models** in improving accuracy.

Another paper used three different approaches of deep learning models for identifying nutrient deficiencies of nitrogen, phosphorus and potassium in Rice (Kolhar, 2024). The proposed models were Xception, Vision Transformet and Multi-layer perceptron-bases (MLP) Mixer, where the best performance was by the transfer model with an accuracy of 95.14%, showing the effectiveness of convolutional model in this domain. However, the paper also explored newer architectures, which while effective, did not outperform CNN models like Xception. One of the first papers focused on **nutrient deficiency detection in coffee leaf plants** (Lewis, 2020), implemented a dataset of 1,000 images from various farms, the researchers of this paper classified deficiencies such as nitrogen, phosphorus and boron, achieving an overall accuracy of 91.94%. This study demonstrated the practicality of using CNN models for real-world agricultural challenges and emphasized the importance of developing models that can help farmers to detect and rectify nutrient deficiencies. In 2023, it was collected a dataset called CoLeaf-DB in Peru (Tuesta-Monteza V. A.-C.-D., 2023), the dataset consists of 1,006 high-resolution images of Peruvian coffee leaves, which are categorized by eight types of nutritional deficiencies, including nitrogen, phosphorus, potassium, magnesium, boron, manganese, calcium and iron. In this paper was applied a ResNet-50 architecture, the model achieved an accuracy of 87.75%. However, the aim of this research was to make suitable the dataset for various image analysis tasks like feature extraction, segmentation and classification. These papers emphasize the value of transfer learning in nutrient deficiency detection, highlighting its adaptability across different plant species and conditions. Despite the broader applications in agriculture, nutrient deficiency detection in coffee plants remains relatively underexplored compared to other crops. Nonetheless, a number of key studies have emerged that set a strong foundation for applying deep learning models, explicitly CNN-based transfer learning models, to this critical area of coffee production.

Further advancements have been made, it was introduced a method called PND- Net (Plant Nutrition Deficiency and Disease Network), integrating graph convolutional networks (GNC) with CNNs. The research based on this model used CoLeaf dataset (Bera, 2024), PND-Net obtained an accuracy of 90.54% in nutrient deficiency classification. The combination of GCN and CNN layers proved effective in extracting detailed features from coffee leaf images, allowing for better differentiation between nutrient deficiencies. **Ensemble learning** has also been explores in this context. Papers focused on rice plants (Sharma, 2022) showed that combining models such as InceptionResNet-V2 with DenseNet significantly boosted classification accuracy of 100% according to the quality of the dataset as it was proposed as individual transfer model, where

InceptionResNet-V2 obtained an accuracy of 90% and Xception Model had an accuracy of 95.83%. In coffee leaf disease classification, a research emphasizes the use of ensemble models to improve classification performance, combining various combinations of CNN architectures and Vision transfers, which are employed using **early** and **late fusion** strategies to enhance accuracy (Cong Pham, 2023). The results of these experiments achieved accuracy of 97.80% in the early fusion with the combination of MobileNet and EfficientNet, achieving the same accuracy in the late fusion of MibileNet and Vision Transformer. Ensemble methods in nutrient deficiency detection, while computationally expensive, offer robustness and generalizability, especially when applied to complex datasets like coffee leaves where deficiencies may present similar visual cues.

Another important technique that has been essential for overcoming the challenges of limited datasets in agriculture is data augmentation (Paulos, 2022). In the University of Rwandan (Hitimana, 2023), the researches implemented **data augmentation techniques** to their dataset of coffee leaves, which allowed for increased model robustness despite a relatively small dataset. The paper analysed leaf plants, which were affected by Rust, Miner and red spider mites disease, with the aim to training transfer learning models like Inception-V3, ResNet-50, Xception, VGG-16 and DenseNet, where the DenseNet Model achieved the highest accuracy of 99.57%, outperforming the other models. In the case that the dataset size does not represent any inconvenient, another techniques that can help to provide visual explanations for the classification of models decisions is combine **transfer learning** and **explainable AI (XAI)** (Mridha, 2023)**,** as it was implemented for classifying coffee leaf disease such as leaf rust, phoma, miner, cercospora and healthy leaves. The paper used techniques like Grad-CAM and Grad-CAM++, which helped visualize the regions of interest in coffee leaves, enhancing model interpretability and providing actionable insights for farmers. Must of the researches related to classification of coffee lead diseases have improved the application of transfers learning with another techniques, as it was mentioned previously.

In Kenyan Arabica coffee plants, the investigation was focused on comparing the performance of ResNet-50, DenseNet-121 and VGG-16 models, with the idea of improving the accuracy of the models throughout of **fine-tuning** (Binney, 2022), making them more effective in real- time, in this performance, the highest accuracy was of 95.44% by DenseNet-121 architecture, which after fine-tuning reached an accuracy of 99.36%. A diversity of techniques or approaches help to improve the application of deep learning models, the following paper applies the use of an ensemble method combining EfficientNet-B0, ResNet-152 and VGG-16, which as well is fine-tuning (Novtahaning, 2022). The dataset of this research was previously pre-processed with data augmentation due to the dataset contains 1,300 images divided into five classes. The proposed ensemble architecture outperformed an accuracy of 97.31%, highlighting the efficiency of transfer learning and fine-tuning, without to say that the combination of multiple CNN architectures improve the robustness and classification accuracy over single models. Additionally, it is possible to **develop a smartphone application** based on transfer learning approaches, as it was done for rice leaf images disease and nutrient deficiency detection (Nayak, 2023). The researchers implemented DenseNet-201, Xception, MobileNet-V2 and ResNet-59, integrating the models into an Android app called “Rice Disease Detector”, optimized for offline use in rural areas where real-time, on-site analysis is needed, The best performance model for smartphone applications with the MobileNet-V2 architecture, having a validation accuracy of 97.56%. The idea of integrating a model into a mobile app for farmers into a mobile app was also proposed for coffee leaf investigation, however, the idea was mentioned in the investigation of detecting and classifying coffee leaf disease and their severity (Rust disease categorized into four leaves) and healthy leaves (Chaubey, 2023). The methodology includes preprocessing with U2Net for background removal and DeepLab-V3 for segmenting the disease areas, followed by training CNN models such as VGG-16, Inception- V3 and MobileNet-V2 for disease severity classification, performing an accuracy of 97.99% the last approach.

3. Research Methodology

The research methodology designed in this paper is focus on the following steps namely Data Collection, Data Preparation, Modelling, and Evaluation as illustrate in Fig 1.

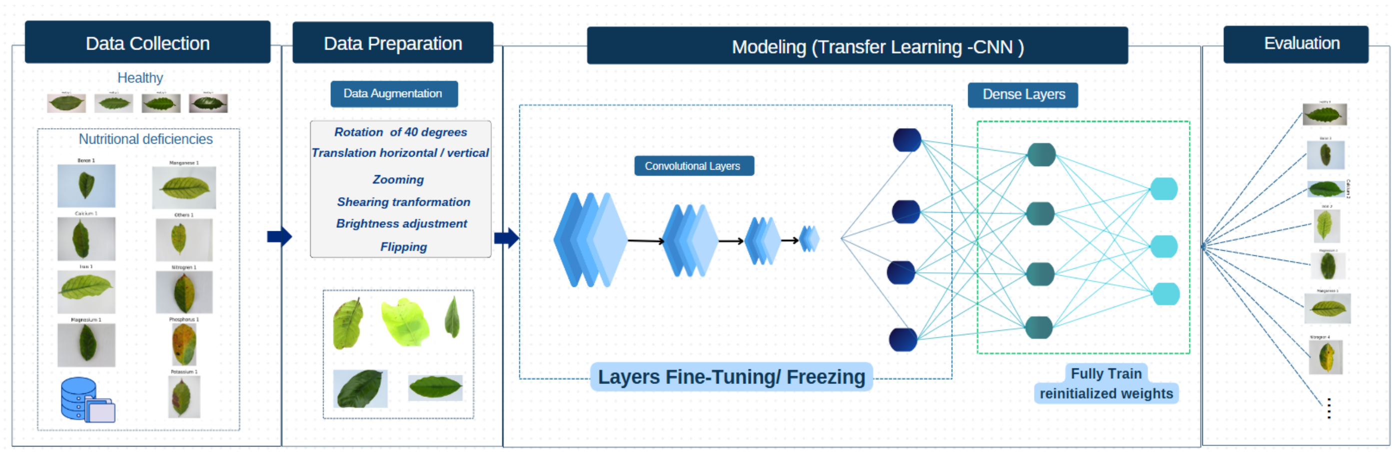

Fig. 1. Research Methodology flow diagram highlighting key steps.

3.1 Data Collection

3.1.1 Datasets

The dataset was collected from Mendeley Data public repository[3], this repository contains one dataset, known as CoLeaf dataset (Tuesta-Monteza V. M.-C.-D., 2023), which contains 1,006 images split into ten classes according to their nutritional deficiencies such as Boron (B), Iron (Fe), Potassium (K), Magnesium (Mg), Manganese (Mn), Nitrogen (N), Phosphorus (P), Other deficiencies and Healthy. The dataset belongs to Peruvian coffee leaf images that facilitate the use for training and validation of deep leaning algorithms in detection and classification for coffee leaf nutritional deficiencies through these images (see the distribution of the dataset in the table 1). In addition, it was used a dataset from Kaggle Website[4] called Coffee Lead Disease (Gaurav, 2023), from this dataset was used only healthy leaf files, owing to the other files corresponding to pest disease.

Classes	Number of images
Boron	101
Calcium	162
Iron	65
Magnesium	79
Manganese	83
Nitrogen	64

[3] https://data.mendeley.com/datasets/brfgw46wzb/2
[4] https://www.kaggle.com/datasets/gauravduttakiit/coffee-leaf-diseases

Phosphorus	246
Potassium	96
Other deficiencies	104
Healthy	6
Healthy[5]	284

Table 1. Distribution of CoLeaf Dataset, plus Nodisease file from Kaggle DB.

3.1.2 Visual Symptoms of nutritional deficiencies

Healthy coffee plants show oval-shaped leaves with a vibrant green color uniformly covering the entire leaf, including the elongated apex and veins. Following by characteristics specific according to nutritional deficiencies (Tuesta-Monteza V. A.-C.-D., 2023).

a) *Boron (B):* Usually, the deficiency appears in young leaves, causing them to be small, elongated, twisted, wrinkled, with irregular edges, deformed and leathery in texture. The leaves display a dull olive-green chlorosis that extends from the to the base. Older leaves develop yellowing at the tips, corky midribs, and secondary veins.
b) *Calcium (Ca):* The characteristic of this deficiency is by marginal chlorosis in young leaves, often accompanied by leaf deformation, causing a convex shape. Cork formation occurs in the veins on the underside of the leaves.
c) *Iron (Fe):* Plants affected by iron deficiency, in young leaves develop a greenish-yellow to pale-green close to white coloration, while the veins stay green.
d) *Magnesium (Mg):* This deficiency usually appears as interveinal chlorosis on older leaves, with yellowing beginning at the branch base and progressing towards the tip. Green stripes along the midrib form an inverted wedge shape cross the petiole, leading to rapid and severe leaf drop.
e) *Manganese (Mn):* It appears the lack of manganese in young leaves as a pale green color, with the main veins and a surrounding band remaining deep green, and the leaves become increasingly yellow.
f) *Nitrogen (N):* This begins as a uniform yellowing (chlorosis) that moves from the base to the tip of the leaf and from the central vein on the way to the edges. In more severe cases, the entire leaf becomes increasingly yellowed, eventually affecting the entire leaf blade.
g) *Phosphorus (P):* Initially, a yellow-brown band appears on the leaves, which later turns necrotic with a dark brown color. A yellow halo borders the necrotic area, affecting first the older leaves and eventually causing the edges and tips to curl upwards.
h) *Potassium (K):* This deficiency starts as a yellow-brown band on the leaves, later becoming necrotic with a dark brown color, bordered by a yellow halo. It first appears on older leaves, eventually causing the edges and tips to curl upwards.

[5] Nodisease File from Kaggle dataset.

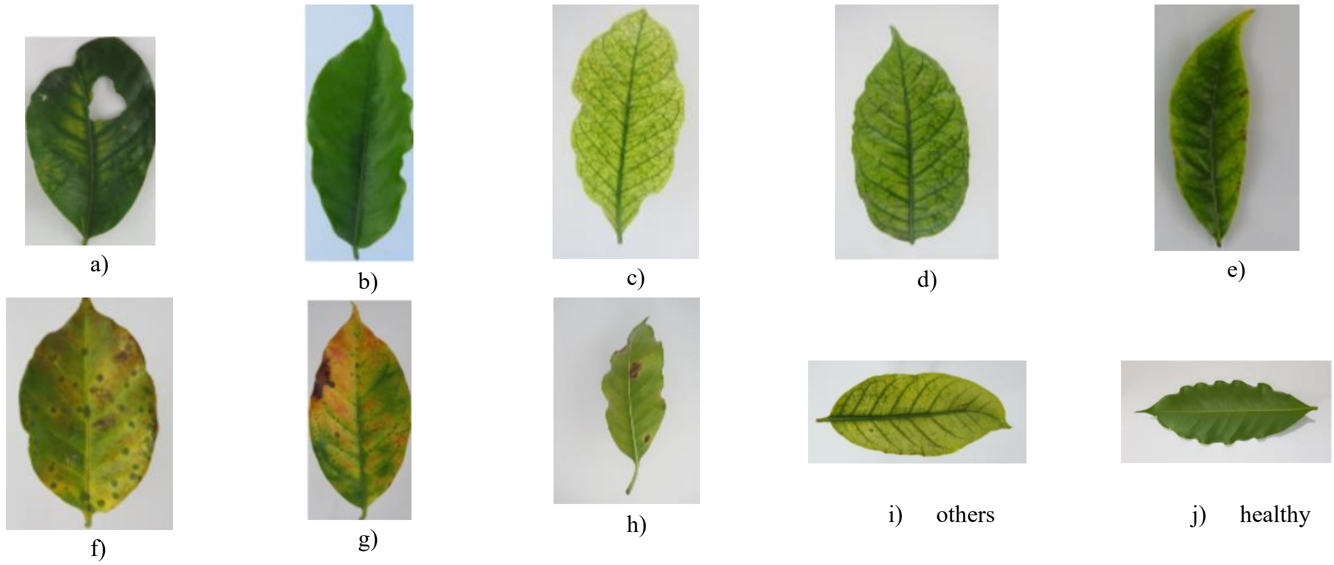

Fig. 2. Nutritional Deficiencies.

The possible bias presented in the collection of this dataset is the imbalance of samples for each class, in accordance with the health of coffee plantations has relied on how well the owners manage their cops. For the classification of coffee nutritional deficiencies was implemented a CNN ResNet-50, achieving an accuracy of 87.75% for testing test. The models was trained with a learning rate of 0.001 and 100 epochs.

3.2 Data Preparation

Data Augmentation is an important step in this research, as it was mentioned in the section 3.1.2 and the table 1 shows, the dataset presented imbalance, in consequence, the use of the next techniques will help to create additional training data without any modification of the corresponding labels (Haba, 2023).

i **Rotation:** Applying this technique to an image means turning it either clockwise or counterclockwise, measured in degrees. General rotation affects the X-Y plane, while tilting involves rotation in the Z plane. Skewing or shearing rotates the images across all there planes X, Y and Z.

ii **Shearing:** It refers to shifting parts of an image in opposite directions, offering a diverse perspective that can enhance model robustness. However, too much shearing can distort the image and obstruct the recognition. It is crucial to balance shearing to avoid excessive distortion and noise, making it a useful tool for improving model generalization without overdoing it .

iii **Zooming:** It is similar to enlarging or cropping an image while keeping the aspect ratio consistent.

iv **Flipping:** There are two types of flips, horizontal and vertical, the horizontal flips mirror the image and work well for most photos except directional ones, both of the flips can be used together, such as with aerial photos.

v **Brightness:** It is also an useful augmentation technique caused by objects are rarely seen under ideal lighting conditions. Training the model with images in various lighting scenarios is essential.

Fig. 3. Image of a leaf with Phosphorus deficiency before data augmentation

Fig. 4. Image of a leaf with Phosphorus deficiency after data augmentation

3.3 Modelling

3.3.1 Convolutional Neural Network-based Model

A convolutional neural network (CNN) or known as well as ConvNet belongs to deep neural networks, regularly, it is designed for identifying, recognizing and classifying images, by using convolution layers. The extraction of information through a convolution operation implies combining input data with filters to form a transformed feature map and finally the filters are modified. A simple CNN model is built by three main layers named as convolution layer, pooling layer and fully connected layer (Sewak, 2018).

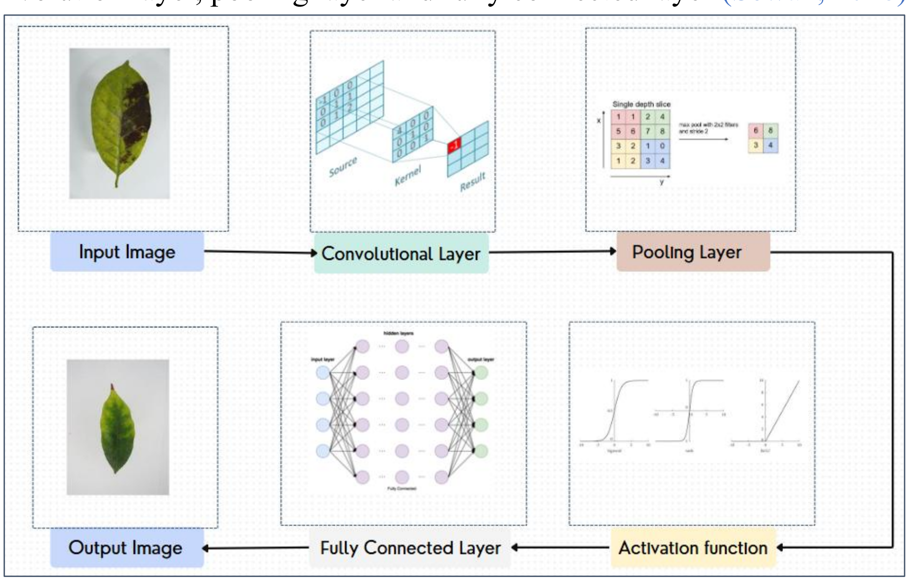

Fig. 5. Components in a typically Convolutional Neural Network simple model.

a. Convolutional Layer: this layer is on charge to extract features from the input image, in order to reduce into small set of matrix known as filters from the input matrix, with the aim to reduce the quantity of weights to be learned such as colors, padding, edges and orientation (Tharsanee, 2021), creating one feature map for each filter.

b. Pooling Layer: it is responsible for controlling overfitting by way of reducing the complexity of the following layer, it seems this is done in order to decrease the quality of each image, however, the filters are not affected (Albawi, 2017). The two approaches used for this step are max pooling and average pooling, in max pooling the size used is 2x2, frequently.

c. Fully Connected Layer: on this layer is taken features from the previous layer with the purpose of estimating class probabilities and classifying to the class that belongs to, these probabilities can be estimated by ReLU or Softmax function.

3.3.2 Transfer Learning Architectures

The application of deep learning models means the used of large of data for training, nonetheless, when it is about images can be difficult to obtain, on the ground of this, it is a helpful approach to work with a model already trained whit the objective to transfer the learned knowledge of features to the next model. An efficient solution is to implement Transfer Learning improving the performance of state-of-the-art CNN model.

i **VGG-16:** VGG is an acronym of Visual Geometry Group, which was created by researchers at the University of Oxford in England. It is considered one of the best computer vision in image classification and image recognition at the moment. This architecture works with a small convolutional filter size 3x3, which allows to increase the depth of the layers and deal with the large scale images (Sarkar, 2018). Even though, there is a series of VGG models, the most popular has been the VGG-16, it is because it can classify an amount of images of various classes achieving an accuracy of 92.7%.This model has 16 learnable parameters layers, where thirteen are convolutional layers and three fully connected layers, however, the real structure of the model is composed of thirteen convolutional layers + five Max Pooling layers + three Dense layers, it says, 21 layers (G, 2021).

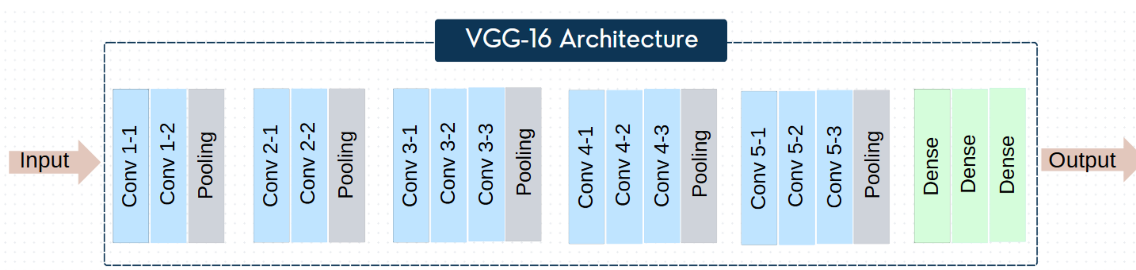

Fig. 6. VGG-16 Model Architecture

ii **ResNet-50:** It is called ResNet due to the term residuals, it is considered that residuals bring an easier learning than unmodified feature maps. The main idea with this model was to address the issue of the vanishing and exploding gradient, the way to achieve this is throughout to skip connections between some levels, which connect layer activations to the following layers, creating the residuals, these will provide learning to the next convolutional layers, which will skip more connections and create more residuals (Chin, 2023). ResNet-50 architecture is based on four parts (Kundu, 2023), first, **convolutional layers** that extract features such as edges, textures and shapes from the input image using batch normalization and ReLU activation, followed by max pooling, helping to keep the most important features. After, **the identity block and the convolutional block** process and transform the features, the identity block is the way where is learned residual functions from the input to the desired output, additionally the convolutional block has a 1x1 convolutional layer, which decrease the number of filters. And finally, the **fully connected layers** makes possible the classification using probabilities for each class using a softmax function.

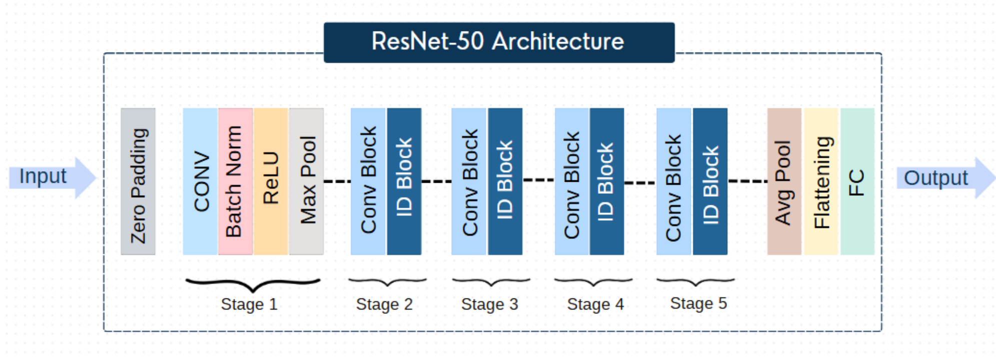

Fig. 7. ResNet-50 Model Architecture

iii **DenseNet-121:** Dense Convolutional Networks (DenseNet) architecture skip connections on steroids, which make to skip a plenty of them. Instead of using residuals as ResNet, DenseNet use concatenation, directly adding earlier information to the outputs of future layers without modifying neuron numbers, this require additional neurons but allows the model to process raw data effectively. DenseNet employs dense blocks, where each layer accesses the outputs of all previous layers through feature map concatenation, promoting feature reuse and maintaining parameter efficiency. Each dense block uses zero-padding to keep spatial dimensions consistent for concatenation. The number of filters in each layer within a block is constant, called the growth rate, which ensure a structures increase in channels. The fill network stacks multiple dense blocks, combined convolutional and pooling layers, to gradually reduce the spatial dimensions of the feature map (Chin, 2023).

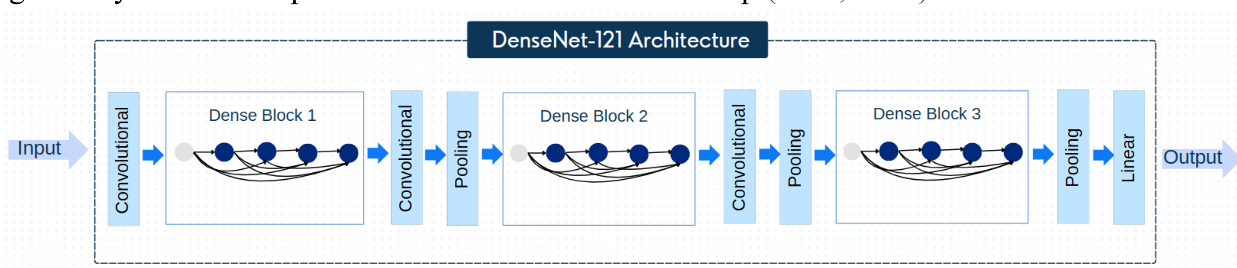

Fig. 8. DenseNet-121 Model Architecture split into Dense Blocks and Layers.

iv **MobileNet:** It was designed with the aim to guarantee efficiency and speedy on mobile and embedded devices, as a consequence of lower level of computational resources. MobileNet is built with a depth-wise separable convolutional layer, nonetheless, the first layers has only 28 layers, which is factorized into two layers, the **depth-wise convolutional** and **point-wise** convolutional layers (Chin, 2023). The purpose of using this depth-wise separable convolutions is to reduce the number of parameters and operations when it is processing images. The depth-wise convolution take an input of feature maps for convolving them with a group of filter to a single input channel, this single channel is convolved again and achieved an output of feature maps with the same number of channels similar to the input (AI, 2023). Once is done this step, it is implemented a set of 1x1 convolutions known as the point-wise convolution, where each filter is distributed in each input channel, improving the number of channels, at the same time that their special dimensions reduce.

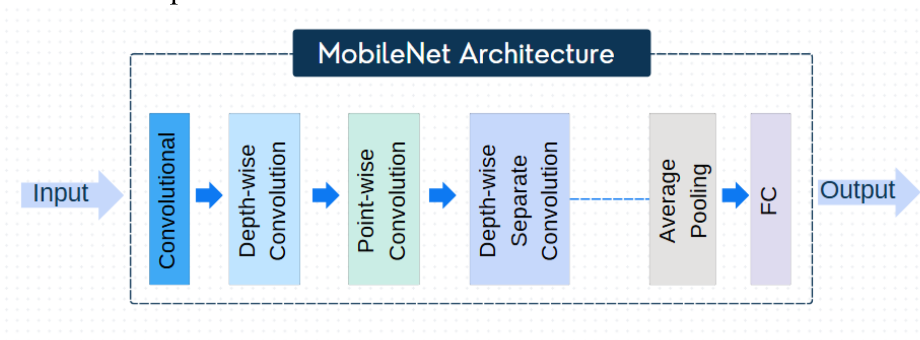

Fig. 9. The MobileNet Architecture

- **Inception:** or called GoogLeNet as well, this architecture works with sub-networks called inception modules, instead of having a high number of parameters, Inception handles few parameters but they are more efficient than another architectures. An Inception model typically includes several convolutional layers with different sizes such as 1x1, 3x3, and 5x5, and a max pooling, both of them organize in parallel. This structure helps to capture features at different scales simultaneously, enhancing efficiency and mitigating the vanishing gradient problem (Géron, 2018). Each step in inception model are responsible for: **1x1 convolution** for reducing the number of channels and decreasing computational complexity. **3x3 convolution** is on charge to extract standard features from the input image. **5x5 convolutional** make the capture of features with large-scales from the input image and **max pooling** reduces the size in each 2x2 window by keeping the maximum value. Finally, all outputs are concatenated and followed by 1x1 convolution to reduce the number of channels and improve accuracy.

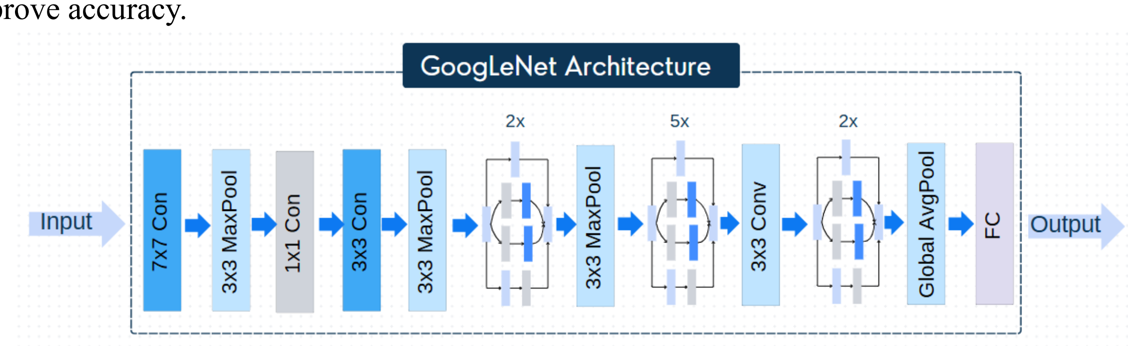

Fig. 10. The GoogLeNet Architecture

3.4 Evaluation

Evaluation metrics are essential for assessing the quality of deep learning models, and their accuracy for the prediction in each model. It is important to use multiple evaluation metrics, as a model might perform well according to one metric but poorly according to another metric. Therefore, in this research is assessed the accuracy and effectiveness of the proposed coffee plant disease classification models using various evaluation metrics.

a. **Accuracy:** it is the ratio of the number of correct predictions to the total number of predictions made for a dataset. Accuracy provides a general idea of the reliability of predictions of model. However, it does not distinguish between classes or types of errors, making it insufficient for imbalanced datasets.

$$\boldsymbol{Accuracy = \frac{Number\ of\ Correct\ predictions}{Total\ number\ of\ predictions\ made}}$$

b. **Confusion Matrix:** known as well as error matrix, displays the number of correct and incorrect predictions made by a model compared to the actual classifications in the test set, highlighting the types of error made. It describes a classification performance of model on test data with known true values (Vujovic, 2021). Confusion Matrix is created after predicting the test data. The possible outputs could be presented prediction are explained such as:
   - **True Positive (TP):** Number of outputs that are actually positive and are predictive positive.
   - **True Negative (TN):** Number of outputs that are actually negative and are predictive negative.
   - **False Positive (FP):** Number of outputs that are actually negative but predictive positive. (Type 1 Errors).
   - **False Negative (FN):** Number of outcomes that are actually positive but predicted negative. (Type 2 Errors).

Positives and negatives refer to the predicted outputs, while true and false indicate whether the predictions are correct. From the confusion matrix it can obtain the succeeding four classification metrics:

1. **Recall (Sensitivity):** or True Positive Rate (TPR) is the ratio of true positives to the total number of actual positives in the dataset. It evaluates the effectiveness of the model in identifying positive samples. A higher number of false negatives results in a lower recall:

$$\boldsymbol{Recall = \frac{TP}{TP + FN}}$$

2. **Precision:** or Positive Predictive Value (PPV) refers to the ratio of true positives to the total number of positives predicted by the model. It is particularly useful for imbalance datasets. A higher number of false positives leads to lower precision.

$$\boldsymbol{Precision = \frac{TP}{TP + FP}}$$

3. **F1 Score:** or F-measure is the weighted mean of precision and recall. A classifier achieves a high F1 score only when both precision and recall are high.

$$\boldsymbol{F1\ Score = 2\frac{Precision\ x\ Recall}{Precision + Recall}}$$

4. Results and Discussion

The dataset used for this paper, contained less than 300 images in each category (see fig. 11), at the beginning, as consequence, it has been applied data augmentation with the main aim to increase the dataset size, which can lead to obtain a better performance, and to prevent overfitting at the moment to apply the purposed models.

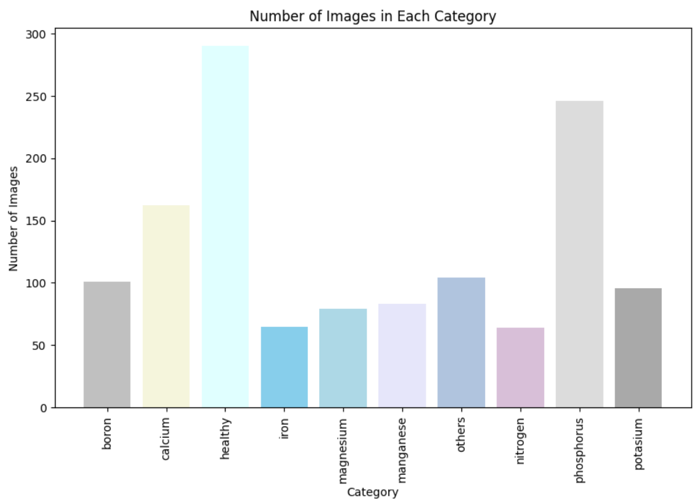

Fig. 11. Distribution of images in each category

The techniques implemented for data augmentation were rotation of 40 degrees, translation of the images horizontally and vertically by up to 20% of width and height respectively, moreover, zooming and flipping images, resulting in 2,543 images per class. Following by training and validation of the dataset, where the dataset was rescaled to normalize the pixel values. For the optimization of the models was shuffled with a buffer size of 1000 to ensure that batches are randomized and also, it was used a TensorFlow AUTOTUNE setting, which the batches are preloaded to optimize GPU utilization during training.

The purpose of this paper is to classify images of leaves in coffee plants according to nutritional deficiencies, including healthy leaf, this was done by the following architectures, which are pre-trained models as a base. The main output of shape used in each model is height 224, width 224 and RGB (Red, Green Blue) color space of 3. The customization of the top layers consists of a Global Average Pooling, batch normalization, a dense layer, which is fully connected with 1024 units, using a ReLU activation, the next layer was a dropout with a 50% rate, which help to minimize overfitting and an output layer using softmax function with 10 units. In the training process, it was implemented a learning rate schedule, which was gradually reducing as training progresses in Adam optimizer; categorical cross-entropy was chosen as the function since is a multiple-class classification, for evaluating the performance of the model was used accuracy, with all this parameter included in the model, it was trained for 50 epoch initially. After the initial training, layers of the base model from layer 300 onward are unfrozen and fine-tuned, this process was trained for an additional 10 epochs with a much lower learning rate (0.00001), with the objective to performance the weights of the model. Upcoming is described a brief summary of result through accuracy and loss curves obtained in each model.

4.1 VGG-16

Overall, the model seem to be performing well, however, there is room for improvement validation performance. At first instance, the model is performing well on training data, but this performance does not entirely translate to validation data, it says, while training accuracy keeps increasing, the validation accuracy does not follow at the same rate (Fig: 12 Top-Left), suggesting overfitting. In training and validation loss happens the same (Fig 12: Bottom-Left), the model performs well on the training data but has limited improvements in the validation data. Once the model is fine-tuned, it appears that the model generalizes well to the data validation set, with only a minor gap between training and validation accuracy, it means that the process has improved the model, reducing overfitting (Fig 12: Top-Right), while in training and validation, also the gap between them remaining relatively small due to the model is learning general patterns instead of memorizing the training data (Fig12: Bottom-Right).

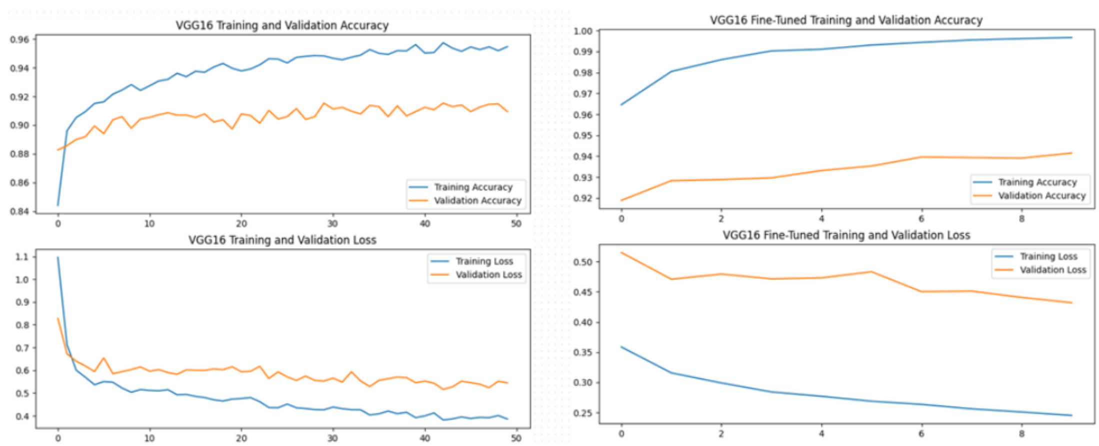

Fig. 12. Training and validation Accuracy (Top-Left), Training and validation Loss (Bottom-Left), Fine-Tuned Training and Validation Accuracy (Top-Right), Fined-Tunned Training and Validation Loss (Bottom-Right) of VGG-16 Model.

4.2 ResNet-50

For ResNet-50 model, the training accuracy shows a strong increase with minor fluctuations, indicating the model is learning and improving over epochs, validation accuracy also increase, though it is slightly more variable compares to the training; seeing the figure 13 (Top-Left), it can perceive that the training accuracy is consistently higher than validation, indicating possible overfitting slightly. On the other hand, training and validation losses decrease, showing that the model is learning and improving performance on both datasets (Fig 13: Bottom-Left), the fluctuations in validation loss suggest some instability in generalization performance, but its trend is positive. If the model is fine-tunned shows a high performance on training data bit not as strongly on validation data, due to the training accuracy obtained 1.0 and the training loss compared to validation metrics suggest that the model has learned to memorize the training data. In general, the fine-tuned model (Fig 13: Right side) shows that validation accuracy and loss improved, however, the improvements are less pronounced than the training metrics. Despite some overfitting, the validation accuracy shows a steady increase and the loss decrease, resulting in a good generalization progress.

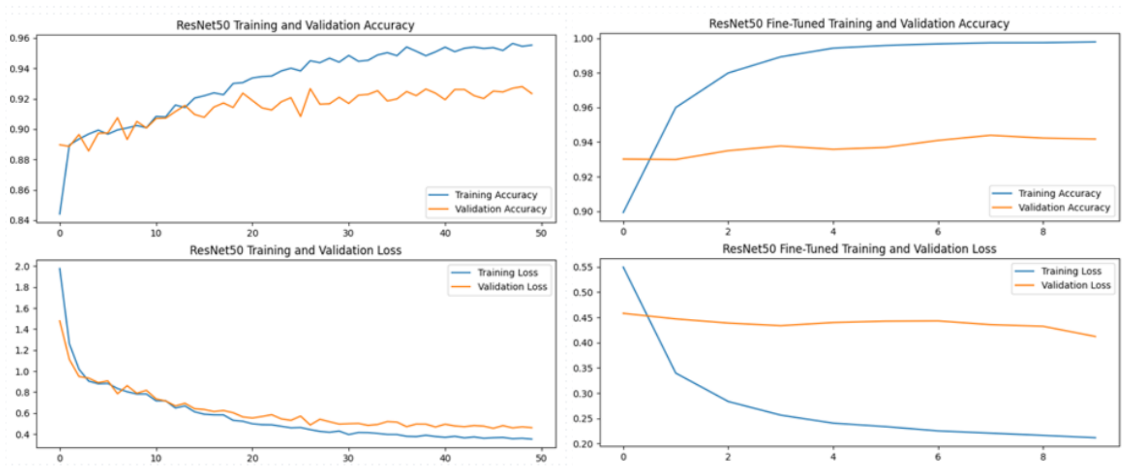

Fig. 13. Training and validation Accuracy (Top-Left), Training and validation Loss (Bottom-Left), Fine-Tuned Training and Validation Accuracy (Top-Right), Fined-Tunned Training and Validation Loss (Bottom-Right) of ResNet-50 Model.

4.3 DenseNet-121

Overall, the DenseNet-121 model shows that there are a noticeable gap between training and validation accuracy, the model performs better on training data, which could indicate overfitting. However, the gap is not excessively large, which is a positive sign (Fig 14: Top-Left). In training and validation loss decrease, but the training loss drops faster and lower than the validation loss (Fig14: Bottom-Left), suggesting that while the model is learning well, there still be some overfitting for improvements. Fine-Tuning DenseNet-121 model, there is a gap between accuracy (92.24%) versus validation accuracy (90.88%), performing much better on the training data (Fig14: Top-Right) and suggesting possible overfitting. In training and validation losses, both are decreasing over epochs, which indicates that the model is learning well, nonetheless, the faster and more significant decrease in training loss compared to validation loss, reaffirming overfitting issue to the training data (Fig 14: Bottom-Right).

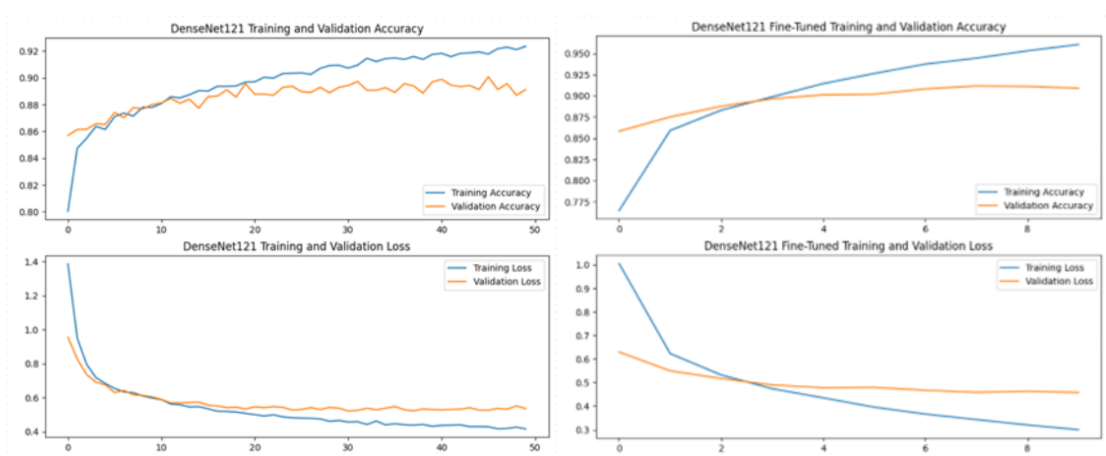

Fig. 14. Training and validation Accuracy (Top-Left), Training and validation Loss (Bottom-Left), Fine-Tuned Training and Validation Accuracy (Top-Right), Fined-Tunned Training and Validation Loss (Bottom-Right) of DenseNet-121 Model.

4.4 MobileNet-V2

MobileNet-V2 shows good results, with steady improvements in both accuracy and loss. The training is consistently higher than validation accuracy, this gap between them, still indicating overfitting slightly, as it performs better on the training than on unseen validation data (Fig 15: Top-Left), nevertheless, the validation accuracy seems to stabilize and improve steadily. Reviewing the loss in the figure 15 (Bottom-Left), the validation loss does not drop as faster as the training loss, reinforcing the idea of presence of overfitting. Running a fine-tuned to MobileNet-V2, the model shows a good improvement due to the training accuracy increase quickly, which is typical as the model starts learning the data patterns and validation accuracy improves steadily, suggesting that the model is learning to generalize well, nevertheless, it does not reach to training accuracy (Fig 15: Top-Right). For both training and validation loss are decreasing, it says, the model reduces errors on both datasets, training and validation.

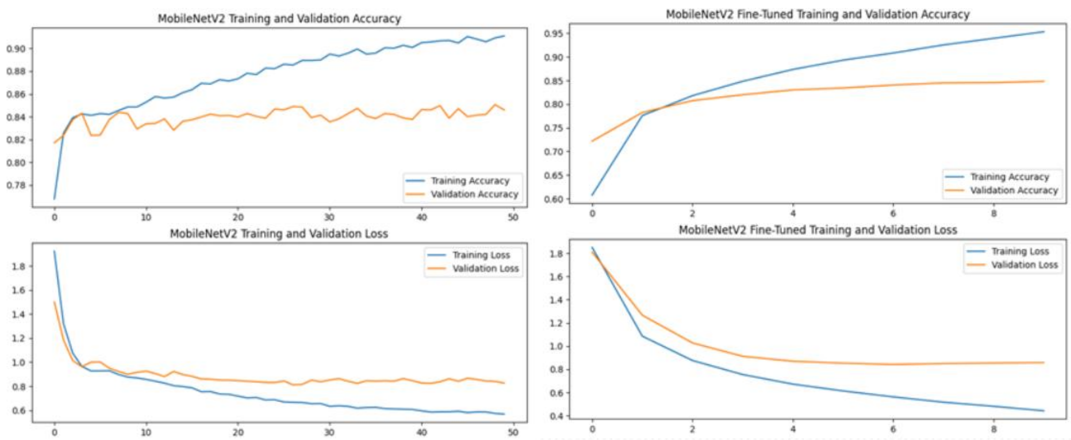

Fig. 15. Training and validation Accuracy (Top-Left), Training and validation Loss (Bottom-Left), Fine-Tuned Training and Validation Accuracy (Top-Right), Fined-Tunned Training and Validation Loss (Bottom-Right) of MobileNet-V2 Model.

4.5 Inception-V3

Overall, seeing both accuracy and loss plots (Figure 16: Left side), shows strong learning capabilities, however may require adjustments to address the fluctuating validation loss. The accuracy of the model consistently improved while the loss decreased, indicating effective learning, while validation accuracy generally improved, the validation loss has periods of fluctuation, that can see from 40 onwards, suggesting overfitting. The fine-tuned Inception-V3 model shows strong learning capabilities on the training data but exhibits signs of overfitting as training progresses, i.e., while training accuracy continues to increase and training loss decreases, validation loss starts to plateau and even slightly increase, it is when starts to overfit the training data (Fig 16: Right side).

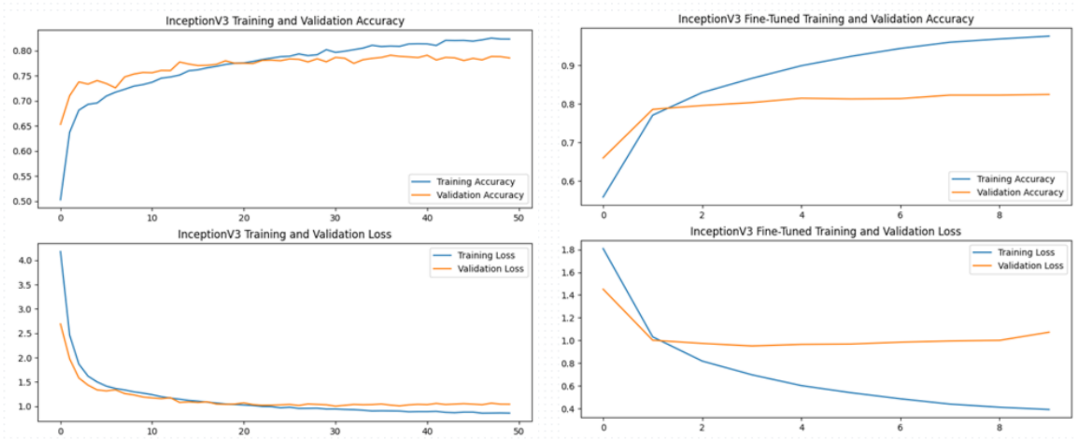

Fig. 16. Training and validation Accuracy (Top-Left), Training and validation Loss (Bottom-Left), Fine-Tuned Training and Validation Accuracy (Top-Right), Fined-Tunned Training and Validation Loss (Bottom-Right) of Inception-V3 Model.

4.6 Overall Evaluation

The following table helps to concentrate the accuracy of each model, the five models proposed for this research display an accuracy over 90%, however, the best performance was done by VGG-16 architecture, which obtained an accuracy of 96.16%, followed by ResNet-50 with an accuracy of 96.16%. Due to all of the models showed possible overfitting using only 50 epochs, they were fine-tuned, applying 10 epochs, where ResNet-50 obtained and improvement of 99.83%, whilst VGG-16, its improvement just achieved an accuracy of 99.69%. On the other hand, must of the models showed that validation dataset with training dataset suggests overfitting, increasing the size of the data and applying regularization techniques can help to reduce overfitting and improve the validation performance further. Despite that validation dataset reaffirms that ResNet-50 shows the best performance with and without fine-tuning the model, with an accuracy of 92.34% and 94.17% respectively as the table 2 illustrate. In addition, the five model were evaluated on test data, once again, ResNet-50 achieved a high accuracy compared with the rest of the models, with 94.23% of accuracy (Table 3).

	Training		**Validation**	
Models	Accuracy	Accuracy Fine-Tuned	Accuracy	Accuracy Fine-Tuned
VGG-16	**96.16%**	99.69%	90.94%	94.15%
ResNet-50	95.85%	**99.83%**	**92.34%**	**94.17%**
DenseNet-121	92.48%	96.24%	89.13%	90.88%
MobileNet-V2	91.68%	95.65%	84.60%	84.82%
Inception-V3	82.29%	97.89%	87.56%	82.42%

Table 2. Accuracy with and without Fine-Tuned on Training and Validation data.

Models	**Accuracy**	**Loss**
VGG-16	94.05%	0.4404
ResNet-50	**94.23%**	**0.4246**
DenseNet-121	91.06%	0.4430
MobileNet-V2	85.29%	.8387
Inception-V3	83.24%	1.04

Table 3. Accuracy and Loss on Test datasets.

Owing to ResNet-50 provide the best performance, it was evaluated on the test dataset the performance of each class, it say, the table 4 concentrate the evaluation metrics like precision, recall and F1-score for the architecture mentioned before, predicting the classes related to nutrients deficiencies in coffee plants such as Boron, Calcium, Iron, Magnesium, Manganese, Nitrogen, Phosphorus, Potassium, other deficiencies and includes for healthy leaf. The table reflects an overall strong performance of the model (seeing as well Confusion Matrix in the figure 18), particularly in detecting healthy leaf coffee plants and major nutrient deficiencies like Potassium and Phosphorus though there is a margin for improvement in detecting Magnesium deficiencies and other less common deficiencies. The results that be related to the symptoms associated to Potassium or Phosphorus deficiencies are more remarkable than associated to the other nutrients, moreover, the symptoms in other nutrients class could be similar to the other deficiencies of class (Fig 17).

Classes	Precision	Recall	F1-Score
Boron (B)	96%	97%	97%
Calcium (Ca)	93%	95%	94%
Iron (Fe)	92%	97%	94%
Magnesium (Mg)	91%	90%	90%
Manganese (Mn)	94%	90%	92%
Nitrogen (N)	95%	94%	95%
Phosphorus (P)	97%	96%	97%
Potassium (K)	98%	94%	96%
Other deficiencies	87%	90%	88%
Healthy	**99%**	**98%**	**99%**

Table 4. Metrics of ResNet-50 Model

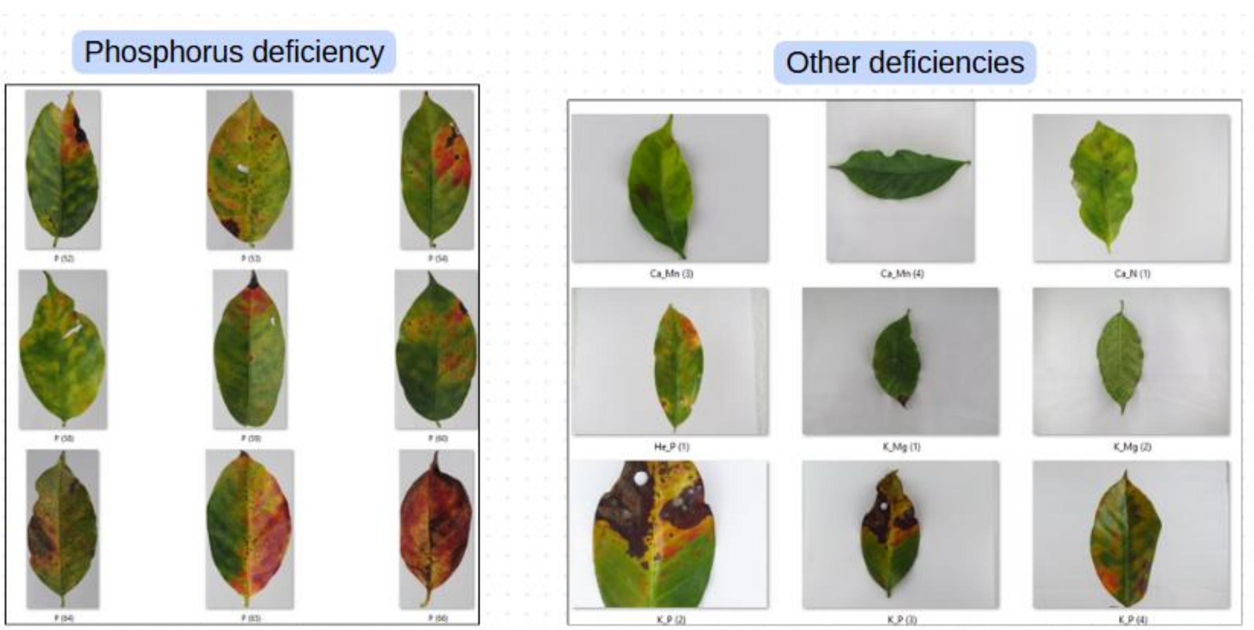

Fig. 17. Phosphorus deficiency class versus other deficiencies class

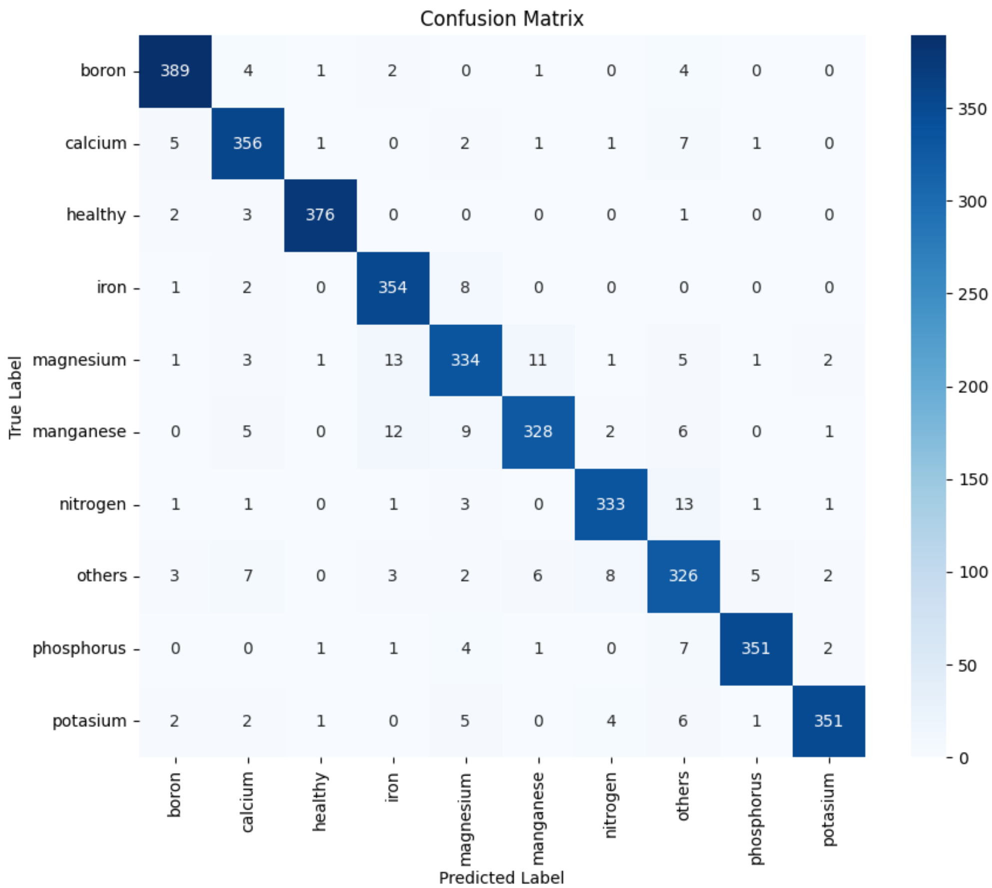

Fig. 18. ResNet-50 Confusion Matrix.

5. Conclusion

The focus of this research is to apply deep learning models to classification of coffee plant leaves, must of previous investigations are based on symptoms that affect to the leaves as consequence of a disease or pests. One of the limitation on analyse different factors or parts of the plants is the availability of datasets, so, in CoLeaf dataset was the opportunity to apply different models of deep learning for classification of leaves according to nutritional deficiencies. It was proposed five models using transfer learning in VGG-16, ResNet-50, DenseNet-121, MobileNet-V2 and Inception-V3, where the higher performance was done by ResNet-50 tune-fined model in training and validation, moreover, test datasets, with an accuracy of 99.83%, 94.17% and 94.23% respectively.

References

AI, N. (2023, 01 20). "*MobileNet*". (Medium) Retrieved 08 02, 2024, from https://medium.com/@nocodingai/mobilenet-fc34af9f58a5

Albawi, S. M.-Z. (2017). "*Understanding of a Convolutional Neural Network*". 2017 International Conference on Engineering and Technology (ICET). Antalya.

Attri, I. K. (2023). "*A review od Deep Learning Techniques used in Agriculture"*. Ecological Informatics, 77.

Bera, A. B. (2024). *"PND-Net: pnat nutrition deficiency and disease classification using graph convolutional network"*. Scientific Reports, 14(1), 15537.

Binney, E. R. (2022). *"Coffee Leaf Disease Classification and the Effect of fine-Tuning on Deep Convolutional Neural Networks"*. International Journal for Multidisciplinary Research, 4(5), 861.

Boa Sorte Ximenes, L. F. (2019). *"Coffee Leaf Disease Recognition Based on Deep Learning and Texture Attributtes"*. Procedia Computer Science, 159, 135-144.

Brodzicki, A. P.-K. (2020). *"Transfer Learning Methods as a New Approach in Computer Vision Tasks with Small Datasets"*. Foundations of Computing and Decision Sciences, 179-193.

Chaubey, H. K. (2023). *"Coffee Leaf Disease and Severity Prediction Using Deep Learning"*. TENCON 2023 - 2023 IEEE Region 10 Conference (TENCON) (pp. 117-1180). Chiang Mai, Thailand: IEEE.

Chin, E. K. (2023). "*The Deep Learning Architect's Handbook"*. Birmingham, UK: Packt Publishing Ltd.

Cong Pham, T. N.-D. (2023). *"Artifical Intelligence-based solutions for coffee leaf disease classification"*. International Conference on Marine Sustainable Development and Innovation 2023.

E Umbaugh, S. (2023). "*Digital Image Enhancement, Restoration and Compression, 4th Edition"*. Abingdon, Oxon: CRC Press.

G, R. (2021). *"Everything you need to know about VGG16. Great Learning"*.

Gaurav, D. (2023, 05 04). "*Kaggle"*. Retrieved 06 14, 2024, from https://www.kaggle.com/datasets/gauravduttakiit/coffee-leaf-diseases

Géron, A. (2018). "*Chapter 3. Convolutional Neural Networks"*. O'Reilly Media, Inc.

Haba, D. (2023). "*Data Augmentation with Python"*. Birminham, UK: Packt Publishing Ltd.

Hitimana, E. S. (2023). *"An Intelligent System-Based Coffee Plant Leaf Disease Recognition using Deep Learning Techniques on Rwandan Arabica Dataset"*. Technologies, 11(5), 116.

J., A. E. (2022). *"Deep Learning-Based Leaf Disease detection in Crops Using Images for Agricultural Applications"*. Agronomy, 12(10), 2395.

Jean Ayikpa, K. M. (2022). *"Experimental Evaluation of Coffee Leaf Disease Classification and Recognition based on Machine Learning and Deep Learning Algorithms"*. Journal of Computer Science, 18(12), 1201-1212.

Jepkoech, J. M. (2021). *"Arabica Coffee Leaf Images dataset for Coffee leaf disease detection and classification"*. Data in Brief, 36, 107-142.

Jepkoech, J. M. (2021, 03 26). "*JMuBEN*". (Mendeley Data) Retrieved 04 15, 2024, from https://data.mendeley.com/datasets/t2r6rszp5c/1

Jepkoech, J. M. (2021, 03 26). "*JMuBEN2*". (Mendeley Data) Retrieved 4 15, 2024, from https://data.mendeley.com/datasets/tgv3zb82nd/1

Kolhar, S. J. (2024). *"Deep Neural Networks for Classifying Nutrient Deficiencies in Rice Plants Using Leaf Images"*. International Journal of Computing and Digital Systems, 305-314.

Krohling, R. A. (2019, November 6). "*BRACOL - A Brazilian Arabica Coffee Leaf images dataset to identification and quantification of coffee diseases and pests"*. (Mendeley Data) Retrieved 07 29, 2024, from https://data.mendeley.com/datasets/yy2k5y8mxg/1

Kumar, M. G. (2020).*" Disease Detection in Coffee Plants using Convolutional Neural Networks"*. 2020 5th International Conference on Communication and Electronics Systems (ICCES). Coinmbatore, India.

Kundu, N. (2023, 01 23). "*Exploring ResNet50: An In-Depth Look at the Model Architecture and Code Implementation"*. (Medium) Retrieved 08 01, 2024, from https://medium.com/@nitishkundu1993/exploring-resnet50-an-in-depth-look-at-the-model-architecture-and-code-implementation-d8d8fa67e46f

Kusanur, V. C. (2021). *"Using Transfer Learning for Nutrient Deficiency Prediction and Classification in Tomato Plant"*. International Journal of Advanced Computer Science and Applications.

Lewis, K. P. (2020). *"Classification And Detection Of Nutritional Deficiencies In Coffee Plants Using Image Processing And Convolutional Neural Network (CNN)"*. International Journal of Scientifc & Technology Research.

Lisboa, E. L. (2021). *"Coffee Leaf Diseases Identifiaction and Severity Classification using Deep Learning"*. Anais Estendidos da XXXIV Conference on Graphics, Patterns and Images (SIBRAPI Estendido 2021*)*. Brazil.

Liu, J. W. (2021). *"Plant diseases and pest detection based on deep learning: A review"*. Plant Methods, 17(1), 22.

Mridha, K. T. (2023). *"Explainable Deep Learning for Coffee Leaf Disease Classification in Smart Agriculture: A Visual Approach"*. 2023 International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE) (pp. 1-8). Ballar, India: IEEE.

Mulugeta Abuhayi, B. A. (2023). *"Coffee disease classification using Convolutional Neural Network based on feature concatenation"*. Informatics in Medicine Unlocked, 39, 101245.

N. R. Gatkal, S. M. (2023). *"Machine Learning in Agriculture"*. Just Agriculture, *3*(7).

Nayak, A. C. (2023). *"Application of smartphone-image processing and transfer learning for rice disease and nutrient deficiency detection".* Smart Agricultural Technology, 100-195.

Novtahaning, D. S.-M. (2022). *"Deep Learning Ensemble-Based Automated and High-Performing Recognition of Coffee Leaf Disease"*. Agriculture, 1909.

Paulos, E. B. (2022). *"Detection and Classification of Coffee Leaf Disease using Deep Learning"*. 2022 International Conference on Information and Communication Technology for Development for Africa (ICT4DA) (pp. 1-6). Bahir Dar, Ethiopia: IEEE.

Ramos-Ospina, M. G.-T. (2023). *"Deep Transfer Learning for Image Classification of Phosphorus Nutrition States in Individual Maize Leaves"*. Electronics, 16.

S, R. M. (2022). *"Explainable AI for Crop disease detection"*. 2022 4th International Conference on Advances in Computing, Communication Control and Networking (ICAC3N). Greater Noida, India.

Salehi, A. W. (2023). *"A Study of CNN and Transfer Learning in Medical Imaging: Advantages, Challenges, Future Scope"*. Sustainability, 15(7), 5930.

Sarkar, D. B. (2018). "*Hands-On Transfer Learning with Python"*. Birmingham, UK.: Packt Publishing Ltd.

Sewak, M. K. (2018). "*Practical Convolutional Networks"*. Birminham, UK: Packt Publishing.

Sharma, M. N. (2022). *"Ensemble Averaging of Transfer Learning Models for Identification of Nutritional Deficiency in Rice Plant"*. Electronics, 148.

Singh Patel, G. R. (2021). "*Smart Agriculture: Emerging Pedadogies of Deep Learning, Machine Learning and Internet of Things"*. London, UK: CRC Press Taylor & Francis Group.

Taglione, J. (2021). "*Coffee leaf Disease"*. (Kaggle) Retrieved 07 20, 2024, from https://www.kaggle.com/code/jtaglione/coffee-leaf-diseases/data

Thakker, D. M. (2020). *"A Novel Application of Deep Learning with Image Cropping: A Smart City use Case for Flood Monitoring"*. Journal of Reliable Intelligent Enviroments, *6*(1), 51-61.

Tharsanee, R. M. (2021). *"Deep Convolutional Neural Network-based image Classification for COVID-19 diagnosis"*. Data Science for COVID-19, 117-145.

Tuesta-Monteza, V. A.-C.-D. (2023). *"CoLeaf-DB: Peruvian coffee leaf images dataset for coffee leaf nutritional deficiencies detection and classifcation"*. Data in Brief, 48, 109226.

Tuesta-Monteza, V. M.-C.-D. (2023, 05 30). "*CoLeaf Dataset"*. Retrieved 07 30, 2024, from https://data.mendeley.com/datasets/brfgw46wzb/2

Vujovic, Ž. Đ. (2021). *"Classification Model Evaluation Metrics"*. International Journal of Advanced Computer Science and Applications, *12*(6).

Xu, Z. G. (2020). *"Using Deep Convolutional Neural Networks for Image-Based Diagnosis of Nutrient Deficiencies in Rice"*. Computational Intelligence and Neuroscience, 1-12.