Challenge ANNDL

The ReLUtants

Anno Accademico 2023/2024

Jacopo Piazzalunga, Gabriele Puglisi, Davide Salonico, Denis Sanduleanu

1 Introduction

The goal of this challenge is to address a binary classification problem, distinguishing plants into two categories: those in optimal health (class 0) and those exhibiting some form of disease (class 1). As part of the model evaluation, four key metrics will be considered: accuracy, F1 score, recall, and precision. Accuracy, being the primary criterion, represents the percentage of correct predictions relative to the total predictions, providing an overall assessment of the model's performance. Subsequently, the F1 score, recall, and precision will be analyzed, offering an other perspective to evaluate the model's ability to balance correctness, sensitivity, and precision. This report will delve into the implementation of a deep learning model to tackle this binary classification challenge, examining the dataset e the adopted methodology.

2 Dataset

2.1 Analysis

Initially, we conducted a thorough examination of the dataset, revealing the presence of 5200 images related to plants. However, upon closer inspection, we identified additional images within the dataset that were unrelated to plants. Initially, we retained these images, but after several iterations, we decided to exclude them using only 5004 images for our study. This decision was motivated by their lack of relevance to the primary task, and their presence was found to hinder the accurate extraction of features crucial for the model's performance.

2.2 Data Augmentation

To enhance the robustness and diversity of our training dataset, we adopted the data augmentation technique. This involved applying various transformations to the existing data to generate new variations of the original images. The objective was to improve the model's ability to generalize across different conditions. Our data augmentation strategy included conventional techniques such as rotation and zoom. Furthermore, we implemented other advanced augmentation methods, like CutMix and MixUp.

3 Models Architecture

In our exploration of plant disease classification, we drew inspiration from recent papers in the field, where a prevailing strategy involved transfer learning. This technique utilizes pre-existing neural networks for initial feature extraction, followed by fine-tuning on the specific task. Motivated by the success of this approach, we chose to implement a similar methodology

3.1 Choose of the pre-existing neural network

Our goal from the beginning has been to try different known nets in order to measure which one performed best. Our first attempts where with relatively small models such as VGG, ResNet and

MobileNet. To be precise these are 'families' of nets, so for everyone of them different variants and configurations where tried. ConvNeXtLarge reveald to be the best one overall, even the number of parameters is significantly larger than the one of the previous nets. The operation of trying different nets implied using different preprocessing functions and some tweaking of the hyperparameters. We managed to automatize this process using Optuna, a hyperparameters optimization framework, using a random search approach. This has been possible only for smaller networks due to our limited resources, but theoretically it could have been used for every model.

3.2 Transfer Learning and Fine Tuning

Transfer learning and fine tuning proved to be a key factor during the training of the model. The weights have been always imported by the IMAGENET dataset to initialize our networks (according to the competition rules). To perform fine tuning we noticed that bigger network perform better with a reduced number of final unlocked layers so that's the overall direction we took. At first we tried to perform two separate training phases. In the first one we freezed all the layers of the imported network and attached at the end a non-imported single layer classifier. In the second phase we unlocked some layers of the convolutional network to perform the real fine tuning method. With practise we realized that this approach permitted us to save precious training time (and reducing the number of hyper-parameters) respect to train everything from scratch due to the high complexity of the architecture we chose. In the last models we decided to perform only fine tuning phase because we saw from the cross validation that this approach performed better.

3.3 Solutions implemented in our model

- The best data augmentation techniques were zoom, rotation e flip in our case. We tried to use brightness and traslation but they didn't provide any improvements.

- We decided to use a training, validation, test split of 0.8, 0.1 and 0.1. We tried different splits and this brought the best results.

- Early Stopping had an enormous impact on preventing overfitting and saving time during training phase

- The inclusion of dropout helped prevent the model from relying too heavily on specific neurons, promoting a more robust and generalized network.

- Across different model architectures, our experiments consistently demonstrated that AdamW, when fine-tuned with optimized initial learning rates and weight decay parameters, consistently outperformed other widely used optimizers such as standard Adam, SGD, and RMSprop.

- For the fine tuning phase we decided to adopt a low learning rate with a callbalck that reduce learning rate when a metric has stopped improving, in our case the validation accuracy.

4 Experiments

During this challenge various techniques were experimented:

- Batch Normalization didn't bring a clear advantage in this sense

- The introduction of CutMix and MixCut techniques did not yield substantial alterations in our results. Despite their implementation, the impact on the outcomes remained relatively modest

- Ensemble wasn't really effective. We inspected why and probably the models we trained inferred similar characteristics even being from different families of network. Trying to run test with the same splits we noticed that the misclassified images were more or less the same.

- Applying homomorfic filtering or any kind edge detection operation to the images didn't improve performances and in some cases worsened them. This evidence is totally according to current literature [THT+] for plant images classification tasks.

- We found an imbalance in the plant disease classification dataset (3101 healthy, 1903 unhealthy)and we tried to resolve this issue using Keras' ImageDataGenerator. After some submissions in the second phase of the challenge we saw a drop of performance so we decided to work with the unbalance dataset provided by the competition

- We explored self-supervised learning, a method known for autonomously learning representations from unlabeled data. Despite careful task selection and experimentation, our attempt did not yield expected improvements

- We tried to augument the dataset doubling its size. We used different trasformation in order to that. The usage of this enlarged dataset didn't provide any improvement, specially in the second phase. According to that we dediced to maintain the standard dataset provided and perfoming other data augmentation techniques.

5 Results

1,2 shows the values of accuracy and loss during the training in the fine tuning phase. In 1 we can see the results obtained on our notebook making inference with the test set.

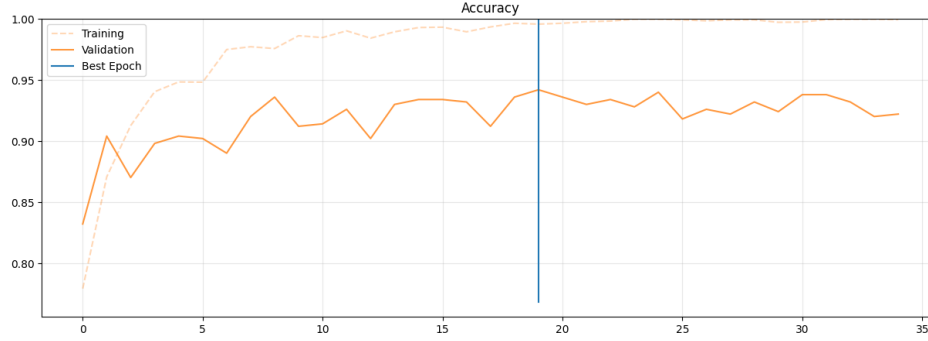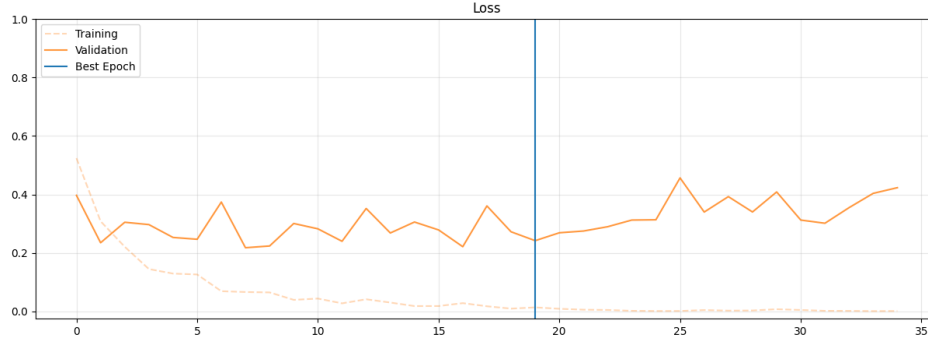Figure 1: Model accuracy in the first phase during the fine tuning operation

Figure 2: Model loss in the first phase during the fine tuning operation

Accuracy	Precision	Recall	F1
0.9421	0.9417	0.9351	0.9382

Table 1: Results on test set

6 Contributions

In our collaborative project, each team member played a crucial role. Gabriele, with a focus on literature review, delved into various papers to provide a comprehensive overview of state-of-the-art

models, guiding our model selection process. Jacopo and Davide led the model development, spending their time on the choosing of the right hyperparameters working with Gabriele for the choice of the net. Denis, on the other hand, played a pivotal role in the data augmentation domain, offering valuable insights and support to enhance the diversity of our dataset. Together, our collective efforts resulted in a well-rounded project, benefitting from the diverse skills and contributions of Jacopo, Denis, Davide, and Gabriele.

References

[AFdBV] Andre S. Abade, Paulo Afonso Ferreira, and Flavio de Barros Vidal. Plant diseases recognition on images using convolutional neural networks: A systematic review.

[Gre93] George D. Greenwade. The Comprehensive Tex Archive Network (CTAN). *TUGBoat*, 14(3):342–351, 1993.

[Kal] Anjaneya Teja Sarma Kalvakolanu. Plant disease detection from images.

[Ker] Keras. [kerascv] semi-supervised image classification using contrastive pretraining with simclr.

[PGL+23] Adrien Pavao, Isabelle Guyon, Anne-Catherine Letournel, Dinh-Tuan Tran, Xavier Baro, Hugo Jair Escalante, Sergio Escalera, Tyler Thomas, and Zhen Xu. Codalab competitions: An open source platform to organize scientific challenges. *Journal of Machine Learning Research*, 24(198):1–6, 2023.

[RIA] Hasin Rehana, Muhammad Ibrahim, and Md. Haider Al. Plant disease detection using region-based convolutional neural network.

[RUMR] J. Anitha Ruth, R. Uma, A. Meenakshi, and P. Ramkumar. Meta-heuristic based deep learning model for leaf diseases detection.

[THT+] Vo Hoang Trong, Yu Gwang Hyun, Nguyen Huy Toan, Lee Ju Hwan, Dang Thanh Vu, and Kim Jin Young. A study on applying homomorphic filter and deep neural network for apple trees diseases classification.

[YTSG] Jianping Yao, Son N. Tran, and Samantha Sawyer Saurabh Garg. Deep learning for plant identification and disease classification from leaf images: Multi-prediction approaches.