Deep Transfer Learning Models for Tomato Disease Detection

Maryam Ouhami¹✉, Youssef Es-Saady¹, Mohamed El Hajji¹,
Adel Hafiane², Raphael Canals², and Mostafa El Yassa¹

¹ IRF-SIC Laboratory, Ibn Zohr University, Agadir, Morocco
maryam.ouhami@edu.uiz.ac.ma

² PRISME Laboratory, INSA CVL University of Orléans, Orléans, France

Abstract. Vegetable crops in Morocco and especially in the Sous-Massa region are exposed to parasitic diseases and pest attacks which affect the quantity and the quality of agricultural production. Precision farming is introduced as one of the biggest revolutions in agriculture, which is committed to improving crop protection by identifying, analyzing and managing variability delivering effective treatment in the right place, at the right time, and with the right rate.

The main purpose of this study is to find the most suitable machine learning model to detect tomato crop diseases in standard RGB images. To deal with this problem we consider the deep learning models DensNet, 161 and 121 layers and VGG16 with transfer learning. Our study is based on images of infected plant leaves divided into 6 types of infections pest attacks and plant diseases. The results were promising with an accuracy up to 95.65% for DensNet161, 94.93% for DensNet121 and 90.58% for VGG16.

Keywords: Precision agriculture · Plant disease detection · Transfer learning · DensNet

1 Introduction

With a surface area of nearly 8.7 million hectares, the Moroccan department of agriculture assumes that agricultural field produces a very wide range of products and generates 13% of the gross domestic product (GDP) [1]. This sector has experienced a significant evolution of the GDP due to the exploitation of fertilization and plant protection systems [1]. Despite the efforts, it still faces important challenges, such as diseases. Pathogen is the factor that causes disease in the plant; it is induced either by physical factors such as sudden climate changes or chemical/biological factors like viruses and fungi [2].

Market gardening and especially tomato crops in the Sous-Massa region are ones of the crops that are exposed to several risks which increase the quantity and quality of the agriculture products. The most important damages are caused by pests' attacks (leaf-miner flies, *Tuta absoluta* and Thrips) in addition to cryptogamic pathogens infections (early blight, late blight and Powdery Mildew).

Since diagnosis can be performed on plant leaves, our study is conducted as a task of classification of symptoms and damages on those leaves. Ground imaging with RGB

camera presents an interesting way for this diagnosis. However, robust algorithms are required to deal with different acquisition conditions: light changes, color calibration, etc. For several years, great efforts have been devoted to the study of plant disease detection. Indeed, feature engineering models [3–6] on one side with Convolutional Neural Networks (CNN) [7–10] on the other side; are carried out to solve this task.

In [6], the study is based on a database of 120 images of infected rice leaves divided into three classes bacterial leaf blight, brown spot, and leaf smut (40 images for each class). Authors have converted the RGB images to an HSV color space to identify lesions, with a segmentation accuracy up to 96.71% using k-means. The experiments were carried out to classify the images based on multiple combinations of the extracted characteristics (texture, color and shape) using Support Vector Machine (SVM). The weakness of this method consists in a moderate accuracy obtained of 73.33%. In fact, the image quality was decreased during the segmentation phase, during which some holes were generated within disease portion, which could be a reason for the low classification accuracy. In addition, leaf smut is misclassified with an accuracy of 40%, which requires other types of features to improve the results.

In the same context, in [4], the authors have proposed an approach for diseases recognition on plant leaves. This approach is based on combining multiple SVM classifiers (sequential, parallel and hybrid) using color, texture and shape characteristics. Different preprocessing have been performed, including normalization, noise reduction and segmentation by the k-means clustering method. The database of infected plant leaves contains six classes including three types of insect pest damages and three forms of pathogen symptoms. The hybrid approach has outperformed the other approaches, achieving a rate of 93.90%. The analysis of the confusion matrix for these three methods has highlighted the causes of misclassification, which are essentially due to the complexity of certain diseases for which it is difficult to differentiate their symptoms during the different stages of development, with a high degree of confusion between powdery mildew and thrips classes in all the combination approaches.

In another study, the authors have used a maize database acquired by a drone flying at a height of 6 m [11]. They have selected patches of 500 by 500 pixels of each original image of 4000 by 6000, and labelled them according to the fact that the most central 224 by 224 area contained a lesion. For the classification step between healthy and diseased, a sliding window of 500 by 500 on the image is used and was introduced in the convolutional neural network (CNN) ResNet model [8]. With a test precision of 97.85%, the method remains non generalizable since the chosen disease has important and distinct symptoms compared to other diseases and the acquisitions are made on a single field. For that reason, it is not clear how the model would perform to classify different diseases with similar symptoms.

Another work that uses aerial images, with the aim of detecting disease symptoms in grape leaves [9]. Authors have used CNN approach by performing a relevant combination of image features and color spaces. Indeed, after the acquisitions of RGB images using UAV at 15 m height. The images were converted into different colorimetric spaces to separate the intensity information from the chrominance. The color spaces used in this study were HSV, LAB and YUV, in addition to the extracted vegetation indices (Excessive Green (ExG), Excessive Red (ExR), Excessive Green-Red (ExGR), Green-Red Vegetation Index (GRVI), Normalized Difference Index

(NDI) and Red-Green Index (RGI)). For classification, they have used the CNN model Net-5 with 4 output classes: soil, healthy, infected and susceptible to infection. The model has been tested on multiple combinations of input data and three patch sizes. The best result was obtained by combining ExG, ExR & GRVI with an accuracy of 95.86% on 64×64 patches.

In [10], the authors tested several existing state-of-the-art CNN architectures for plant disease classification. The public PlantVillage database [12] was used in this study. The database consists of 55,038 images of 14 plant types, divided into 39 classes of healthy and infected leaves, including a background class. The best results were obtained with the transfer learning ResNet34 model, achieving an accuracy of 99.67%.

Several works have been carried out to deal with the problem of plant diseases detection using images provided by remote sensing materials (smartphones, drones...). Nevertheless, CNN have demonstrated high performances to solve this problem compared to models based on the classic feature extracting methods.

In the present study we took advantages from the deep learning and transfer learning approaches to address the problem of the most important damages caused by pests' attacks and cryptogamic pathogens infections in tomato crops. The rest of the paper is organized as follows. Section 2 presents a comparative study and discusses our preliminary results. The conclusion and perspectives are presented in Sect. 3.

2 Materials and Methods

2.1 Data Description

The study was conducted on a database of images of infected leaves, developed and used in [3–5]. The images were taken with a digital camera, Canon 600D, in several farms in the area of Sous Massa, Morocco. Additional images are collected from the Internet in order to increase the size of the database. The dataset is composed of six classes, three of damage caused by insect pests (leafminer flies, thrips and tuta absoluta), and three classes of cryptogamic pathogens symptoms (Early blight, late blight and powdery mildew). The dataset is validated with the help of an agricultural experts. Figure 1 depicts the types of symptoms on tomato leaves, Table 1 presents the composition of the database and the symptoms of each class. The images were resized in order to put the leaves in the center of the images.

2.2 Architecture Model

The motivation behind using deep learning for computer vision is the direct exploitation of image without any hand-crafted features. In plant disease detection field, many researchers have chosen deep models DensNets and VGGs for their high performance in standard computer vision tasks.

DensNets.

The idea behind the DensNet architecture is to avoid creating short paths from the early layers to the later layers and to ensure maximum information flow between the layers of the network. Therefore, DensNet connects all its layers (with corresponding feature

Fig. 1. Images from the dataset where: (a) Early blight, (b) late blight, (c) Powdery mildew, (d) leafminer fly, (e) Thrips and (f) Tuta absoluta.

Table 1. Dataset distribution.

Name	Early blight	Late blight	Powdery mildew	Leaf miner flies	Thrips	Tuta absoluta	Total
Size	111	111	111	111	111	111	666

map sizes) directly to each other. In addition, each layer obtains additional inputs from all previous layers and transmits its own characteristic maps to all subsequent layers [13]. Indeed, according to [13] DensNets require substantially fewer parameters and less computation to achieve state-of-the-art performances. Figure 2(a) gives an example of a five dense layers convolutional model. In this study DensNet was used with 121 layers and 161 layers.

Fig. 2. (a) Architecture of a five dense convolutional layers model [13], (b) Architecture of VGG16 model.

VGG.
Very deep convolutional networks or VGG ranked second in ILSVRC-2014 challenge [14]. The model is widely used for image recognition task especially in the crop field [15, 16]. Consequently, we used VGG with 16 layers. The architecture has 16

convolutional and 5 pooling layers, followed by 3 fully connected layers. The filters are of size $3 \times 3 \times m$ where m is the number of feature maps. Figure 3(b) illustrates the architecture of VGG16.

Fig. 3. (a) Evolution of the loss during training for DensNet161, DensNet121 and VGG16, (b) Evolution of accuracy during training for DensNet161, DensNet121 and VGG16.

Fine-Tuning the Models.

Fine-tuning is a transfer learning method that allows to take advantage of models trained on another computer vision task where a large number of labelled images is available. Moreover, it reduces the needs on having a large dataset and computation power to train the model from scratch [16]. Fine-tuned learning experiments are much faster and more accurate compared to models trained from scratch [10, 15, 17, 18]. Hence, we fine-tuned all network layers of the 3 models based on learned features on the ImageNet dataset [19]. The idea is to take pre-trained weights from the VGG16, Densnet121 and Densnet161 trained on the ImageNet dataset, use those weights as first step of our learning process, then we keep them for every convolutional layer in all the iterations and update only the weights of the linear layers.

2.3 Results and Discussion

Experimental Setup

Experiments were run on a Google Compute Engine instance named Google Colaboratory (Colab) [20] as well as a local machine LenovoY560 with 16 GB of RAM. Colab notebooks are based on Jupyter and work as a Google Docs object. In addition to that, the notebooks are pre-configured with the essential machine learning and artificial

intelligence libraries, such as TensorFlow, Matplotlib, and Keras. Colab operates under Ubuntu 17.10 64 bits and it is composed of an Intel Xeon processor and 13 GB RAM. It is equipped with a NVIDIA Tesla K80 (GK210 chipset), 12 GB RAM, 2496 CUDA.

We implemented and executed the experiments in Python, using PyTorch library [21], which performs automatic differentiation over dynamic computation graphs. In addition, we used the PyTorch model Zoo, which contains various models pretrained on the ImageNet dataset [19]. The model architecture is train with stochastic gradient descent (SGD) optimizer with learning rate of 1e-3 (0.005) and a total of 20 epochs. The dataset is divided into 80% for training and 20% for evaluation.

Performance Evaluation.

The evaluation of loss during the training phase illustrated in Fig. 3(a). Based on the graph we can observe that the training loss converged for all models. A big reduction of loss started from the first 5 epochs, after 20 iterations all the models were optimized with low losses reaching a score of 0.12 for DensNet161, 0.14 for DensNet121 and 0.15 for VGG16.

Figure 3(a) shows the training set accuracy score for epoch of 1 to 20. The training set accuracy score at epoch 20 reach 96.4%, 95.27% and 94.7% for DensNet161, DensNet121 and VGG16 respectively.

After the 14th epoch the training start to converge as well as the training accuracy. In addition, after testing with higher learning rate and increasing number of epochs, the best training scores were achieved using DensNets models. Which means that the models performed better with less parameters. Besides, DensNet161 performed better than Densnet121 due to the deeper architecture of the model.

We can observe in Table 2 that DensNets performed better than VGG model during the test even if their losses reached a score around 0.14 during. Note that DensNet with deeper layers had better test score. In the test phase DensNet161 outperformed DensNet121 and VGG16 with an accuracy 95.65%, 94.93% and 90.58% respectively. We can clearly see from the Table 2 that DensNet161 outperformed the other models in classifying leafminer fly, thrips and powdery mildew with an accuracy up to 100%, 95.65% and 100% respectively. Furthermore, DensNet121 had the best classification rate for early blight, late blight and tuta absoluta with an accuracy of 100%, 95.65% and 95.65 respectively.

Table 2. Accuracy on test set for each class and average accuracy overall classes for DensNet161, DensNet121 and VGG6

Accuracy	DensNet161	DensNet121	VGG16
Leafminer fly	100	91.30	78.26
Thrips	95.65	91.30	91.30
Powdery mildew	100	95.65	95.65
Early blight	95.65	100	100
Late blight	91.30	95.65	86.96
Tuta absoluta	91.30	95.65	91.30
Over all	95.65	94.93	90.58

In order to compare the two models having the best accuracies, we calculated the confusion matrix of the testing dataset for those models. Figure 4 represents the confusion matrix for the DensNet classification models with 161 and 121 layers. More images were misclassified for DensNet121 compared to DensNet161. Moreover, the most confused classes for the DensNet121 are leafminer fly and thrips with two thrips images classified as leafminer fly and one leaf miner classified as thrips. In DensNet161 model, the confusion is more likely between early blight and late blight with one early blight image classified as late blight and one image late blight classified as early blight.

Fig. 4. Confusion matrix of the DensNet architectures. (a) DensNet161 (b) DensNet121

Thrips and early blight are the most misclassified classes for both models, which is due to the similarity between the symptoms making it difficult to differentiate between these classes.

Table 3. State of art comparison.

Authors	Crop	Data size	Method	Accuracy
Brahimi [10]	PlantVillage	55,038 images	Inception_v3	99.76%
Hanks [8]	Maize	8766 UAV Images	ResNet	97.85%
Kerkech [9]	Vine	70,560 patch, UAV images	Le Net-5 on YUV and ExGR	95.92%
El Massi [4]	Tomato	600 images	SVM combination	93.90%
Harshadkumar [6]	Rice	120 images	k-means	73.33%
Our model	Tomato	666 images	DensNet161	95.65%

State of Art Comparison

Table 3 describes the studies we cited earlier in Sect. 1, aligned with our model results. Each approach is using different dataset. Nevertheless, according to the accuracies listed, the approaches based on deep learning models outperformed the approaches based on feature engineering. The results of our model are promising starting with a dataset with the size of 666 images – see Sect. 2.1 and achieving an accuracy of 95.65% using DensNet161 model.

3 Conclusion

In this paper we have studied three deep learning models in order to deal with the problem of plant disease detection. The best test accuracy score is achieved with DenseNet161 with 20 training epochs, outperforming the tested architectures. From the study that has been conducted it is possible to conclude that DensNet has a suitable architecture for the task of plants disease detection based on crop images. Moreover, we realized that DensNets requires less parameters to achieve better performances. The preliminary results are promising. In our future works we will try to improve the results, increase the dataset size and address more challenging diseases detection problems.

Acknowledgment. This study was supported by Campus France, Cultural Action and Cooperation Department of the French embassy in Morocco, in the call of proposals, 2019 campaign, under the name of “Appel à Projet Recherche et Universitaire”.

References

1. MAPM du développement rural et des eaux et forêts, L'agriculture en chiffre_Plan Maroc vert, L'agriculture en chiffre (2018)
2. Jullien, A., Huet, P.: Agriculture de précision. In: *Agricultures et territoires*, Editions L, pp. 1–15 (2005)
3. Es-Saady, Y., El Massi, I., El Yassa, M., Mammass, D., Benazoun, A.: Automatic recognition of plant leaves diseases based on serial combination of two SVM classifiers. In: *Proceedings of the International Conference on Electrical and Information Technologies (ICEIT 2016)*, pp. 561–566 (2016)
4. El Massi, I., Es-Saady, Y., El Yassa, M., Mammass, D., Benazoun, A.: A hybrid combination of multiple SVM classifiers for automatic recognition of the damages and symptoms on plant leaves. In: Mansouri, A., Nouboud, F., Chalifour, A., Mammass, D., Meunier, J., ElMoataz, A. (eds.) *ICISP 2016. LNCS*, vol. 9680, pp. 40–50. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-33618-3_5
5. El Massi, A., Es-Saady, I., El Yassa, Y., Mammass, M., Benazoun, D.: Automatic recognition of the damages and symptoms on plant leaves using parallel combination of two classifiers. In: *13th International Conference on Computer Graphics, Imaging and Visualization (CGiV), Morocco*, pp. 131–136. IEEE (2016)
6. Prajapati, H.B., Shah, J.P., Dabhi, V.K.: Detection and classification of rice plant diseases. *Intell. Decis. Technol.* **11**(3), 357–373 (2017)

7. El, I., Es-Saady, Y., El, M., Mammass, D., Benazoun, A.: Automatic recognition of vegetable crops diseases based on neural network classifier. *Int. J. Comput. Appl.* **158**(4), 48–51 (2017)
8. Wu, H., et al.: Autonomous detection of plant disease symptoms directly from aerial imagery. *Plant Phenom. J.* **2**(1), 1–9 (2019)
9. Kerkech, M., Hafiane, A., Canals, R.: Deep learning approach with colorimetric spaces and vegetation indices for vine diseases detection in UAV images. *Comput. Electron. Agric.* **155** (July), 237–243 (2018)
10. Brahimi, M., Arsenovic, M., Laraba, S., Sladojevic, S., Boukhalfa, K., Moussaoui, A.: Deep learning for plant diseases: detection and saliency map visualisation. In: Zhou, J., Chen, F. (eds.) *Human and Machine Learning. HIS*, pp. 93–117. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-90403-0_6
11. Wiesner-Hanks, T., et al.: Image set for deep learning: field images of maize annotated with disease symptoms. *BMC Res. Notes* **11**(1), 10–12 (2018)
12. Hughes, D., Salathe, M.: An open access repository of images on plant health to enable the development of mobile disease diagnostics (2015). [arXiv:1511.0806](https://arxiv.org/abs/1511.0806)
13. Huang, G., Weinberger, Q.K.: Densely connected convolutional networks. In: *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4700–4708 (2018). [arXiv:1608.06993](https://arxiv.org/abs/1608.06993)
14. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. In: *3rd International Conference on Learning Representations, ICLR 2015*, pp. 1–14 (2015)
15. Chebet, E., Yujian, L., Njuki, S., Yingchun, L.: A comparative study of fine-tuning deep learning models for plant disease identification. *Comput. Electron. Agric.* **161**, 272–279 (2019)
16. Abas, M.A.H., Ismail, N., Yassin, A.I.M., Taib, M.N.: VGG16 for plant image classification with transfer learning and data augmentation. *Int. J. Eng. Technol. (UAE)* **7**, 90–94 (2018)
17. Bah, M.D., Hafiane, A., Canals, R.: Deep learning with unsupervised data labeling for weed detection in line crops in UAV images. *Remote Sens.* **10**(11), 1–22 (2018)
18. Mohanty, S.P., Hughes, D.P., Salathé, M.: Using deep learning for image-based plant disease detection. *Front. Plant Sci.* **7**(9), 1–10 (2016)
19. Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L.: Imagenet: a large-scale hierarchical image database. In: *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 248–255. IEEE (2009)
20. Carneiro, T., Da Nobrega, R.V.M., et al.: Performance analysis of Google Colaboratory as a tool for accelerating deep learning applications. *IEEE Access* **6**, 61677–61685 (2018)
21. Ketkar, N.: Introduction to PyTorch. *Deep Learning with Python*, pp. 195–208. Apress, Berkeley, CA (2017). https://doi.org/10.1007/978-1-4842-2766-4_12