

Using Deep Learning for Image-Based Plant Disease Detection

Sharada Prasanna Mohanty^{1,2}, David Hughes^{3,4,5}, and Marcel Salathé^{1,2,6}

¹Digital Epidemiology Lab, EPFL, Switzerland; ²School of Life Sciences, EPFL, Switzerland; ³Department of Entomology, College of Agricultural Sciences, Penn State University, USA; ⁴Department of Biology, Eberly College of Sciences, Penn State University, USA; ⁵Center for Infectious Disease Dynamics, Huck Institutes of Life Sciences, Penn State University, USA; ⁶School of Computer and Communication Sciences, EPFL, Switzerland

This manuscript was compiled on April 15, 2016

Crop diseases are a major threat to food security, but their rapid identification remains difficult in many parts of the world due to the lack of the necessary infrastructure. The combination of increasing global smartphone penetration and recent advances in computer vision made possible by deep learning has paved the way for smartphone-assisted disease diagnosis. Using a public dataset of 54,306 images of diseased and healthy plant leaves collected under controlled conditions, we train a deep convolutional neural network to identify 14 crop species and 26 diseases (or absence thereof). The trained model achieves an accuracy of 99.35% on a held-out test set, demonstrating the feasibility of this approach. When testing the model on a set of images collected from trusted online sources - i.e. taken under conditions different from the images used for training - the model still achieves an accuracy of 31.4%. While this accuracy is much higher than the one based on random selection (2.6%), a more diverse set of training data is needed to improve the general accuracy. Overall, the approach of training deep learning models on increasingly large and publicly available image datasets presents a clear path towards smartphone-assisted crop disease diagnosis on a massive global scale.

Deep Learning | Crop Diseases | Digital Epidemiology

Modern technologies have given human society the ability to produce enough food to meet the demand of more than 7 billion people. However, food security remains threatened by a number of factors including climate change[1], the decline in pollinators[2], plant diseases [3], and others. Plant diseases are not only a threat to food security at the global scale, but can also have disastrous consequences for smallholder farmers whose livelihoods depend on healthy crops. In the developing world, more than 80 percent of the agricultural production is generated by smallholder farmers [4], and reports of yield loss of more than 50% due to pests and diseases are common [5]. Furthermore, the largest fraction of hungry people (50%) live in smallholder farming households [6], making smallholder farmers a group that's particularly vulnerable to pathogen-derived disruptions in food supply.

Various efforts have been developed to prevent crop loss due to diseases. Historical approaches of widespread application of pesticides have in the past decade increasingly been supplemented by integrated pest management (IPM) approaches [7]. Independent of the approach, identifying a disease correctly when it first appears is a crucial step for efficient disease management. Historically, disease identification has been supported by agricultural extension organizations or other institutions such as local plant clinics. In more recent times, such efforts have additionally been supported by providing information for disease diagnosis online, leveraging the increas-

ing internet penetration worldwide. Even more recently, tools based on mobile phones have proliferated, taking advantage of the historically unparalleled rapid uptake of mobile phone technology in all parts of the world[8].

Smartphones in particular offer very novel approaches to help identify diseases because of their tremendous computing power, high-resolution displays, and extensive built-in sets of accessories such as advanced HD cameras. It is widely estimated that there will be between 5 and 6 billion smartphones on the globe by 2020. At the end of 2015, already 69% of the world's population had access to mobile broadband coverage, and mobile broadband penetration reached 47% in 2015, a 12-fold increase since 2007[8]. The combined factors of widespread smartphone penetration, HD cameras, and high performance processors in mobile devices lead to a situation where disease diagnosis based on automated image recognition, if technically feasible, can be made available at an unprecedented scale. Here, we demonstrate the technical feasibility using a deep learning approach utilizing 54,306 images of 14 crop species with 26 diseases (or healthy) made openly available through the project PlantVillage[9]. An example of each crop - disease pair can be seen in Figure 1.

Computer vision, and object recognition in particular, has made tremendous advances in the past few years. The PASCAL VOC Challenge[10], and more recently the Large Scale Visual Recognition Challenge (ILSVRC)[11] based on the ImageNet dataset[12] have been widely used as benchmarks for numerous visualization-related problems in computer vision, including object classification. In 2012, a large, deep convolutional neural network achieved a top-5 error of 16.4% for the classification of images into 1,000 possible categories[13]. In the following three years, various advances in deep convolutional neural networks lowered the error rate to 3.57% [13] [14] [15] [16] [17]. While training large neural networks can be very time-consuming, the trained models can classify images

Significance Statement

Crop diseases remain a major threat to food supply worldwide. This paper demonstrates the technical feasibility of a deep learning approach to enable automatic disease diagnosis through image recognition. Using a public dataset of 54,306 images of diseased and healthy plant leaves, a deep convolutional neural network is trained to classify crop species and disease status of 38 different classes containing 14 crop species and 26 diseases, achieving an accuracy of over 99%.

Corresponding author: marcel.salathe@epfl.ch

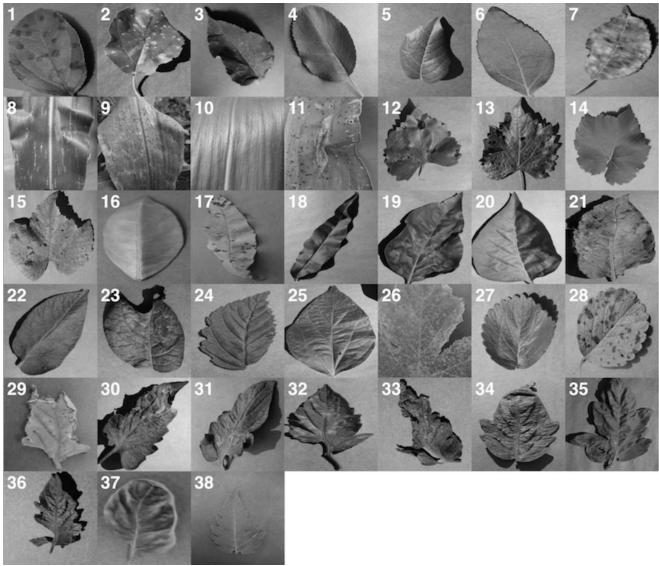


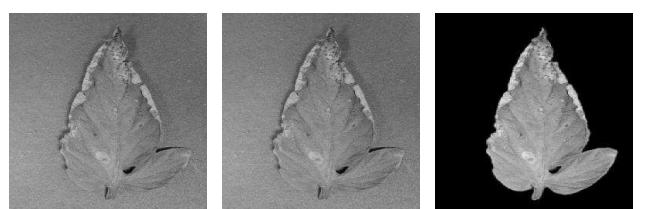
Fig. 1. Example of leaf images from the PlantVillage dataset, representing every crop-disease pair used. 1) Apple Scab, *Venturia inaequalis* 2) Apple Black Rot, *Botryosphaeria obtusa* 3) Apple Cedar Rust, *Gymnosporangium juniperi-virginianae* 4) Apple healthy 5) Blueberry healthy 6) Cherry healthy 7) Cherry Powdery Mildew, *Podosphaera spp.* 8) Corn Gray Leaf Spot, *Cercospora zeae-maydis* 9) Corn Common Rust, *Puccinia sorghi* 10) Corn healthy 11) Corn Northern Leaf Blight, *Exserohilum turicum* 12) Grape Black Rot, *Guignardia bidwellii*, 13) Grape Black Measles (Esca), *Phaeomoniella aleophilum*, *Phaeomoniella chlamydospora* 14) Grape Healthy 15) Grape Leaf Blight, *Pseudocercospora vitis* 16) Orange Huanglongbing (Citrus Greening), *Candidatus Liberibacter spp.* 17) Peach Bacterial Spot, *Xanthomonas campestris* 18) Peach healthy 19) Bell Pepper Bacterial Spot, *Xanthomonas campestris* 20) Bell Pepper healthy 21) Potato Early Blight, *Alternaria solani* 22) Potato healthy 23) Potato Late Blight, *Phytophthora infestans* 24) Raspberry healthy 25) Soybean healthy 26) Squash Powdery Mildew, *Erysiphe cichoracearum*, *Sphaerotheca fuliginea* 27) Strawberry Healthy 28) Strawberry Leaf Scorch, *Diplocarpon earlianum* 29) Tomato Bacterial Spot, *Xanthomonas campestris* pv. *vesicatoria* 30) Tomato Early Blight, *Alternaria solani* 31) Tomato Late Blight, *Phytophthora infestans* 32) Tomato Leaf Mold, *Fulvia fulva* 33) Tomato Septoria Leaf Spot, *Septoria lycopersici* 34) Tomato Two Spotted Spider Mite, *Tetranychus urticae* 35) Tomato Target Spot, *Corynespora cassiicola* 36) Tomato Mosaic Virus 37) Tomato Yellow Leaf Curl Virus 38) Tomato healthy

very quickly, which makes them also suitable for consumer applications on smartphones.

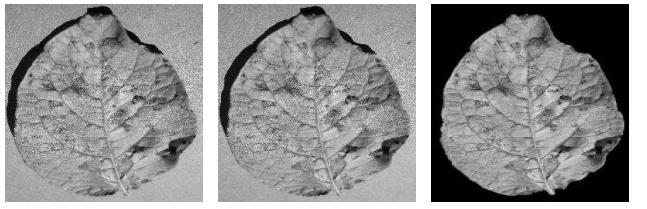
In order to develop accurate image classifiers for the purposes of plant disease diagnosis, we needed a large, verified dataset of images of diseased and healthy plants. Until very recently, such a dataset did not exist, and even smaller datasets were not freely available. To address this problem, the PlantVillage project has begun collecting tens of thousands of images of healthy and diseased crop plants [9], and has made them openly and freely available. Here, we report on the classification of 26 diseases in 14 crop species using 54,306 images with a convolutional neural network approach. We measure the performance of our models based on their ability to predict the correct crop-diseases pair, given 38 possible classes. The best performing model achieves a mean F_1 score of 0.9934 (overall accuracy of 99.35%), hence demonstrating the technical feasibility of our approach. Our results are a first step towards a smartphone-assisted plant disease diagnosis system.

Results

At the outset, we note that on a dataset with 38 class labels, random guessing will only achieve an overall accu-



(a) **Leaf 1:** Color (b) **Leaf 1:** Grayscale (c) **Leaf 1:** Segmented



(d) **Leaf 2:** Color (e) **Leaf 2:** Grayscale (f) **Leaf 2:** Segmented

Fig. 2. Sample images from the three different versions of the PlantVillage dataset used in various experimental configurations.

racy of 2.63% on average. Across all our experimental configurations, which include three visual representations of the image data (see Figure 2), the overall accuracy we obtained on the PlantVillage dataset varied from 85.53% (in case of *AlexNet::TrainingFromScratch::GrayScale::80-20*) to 99.34% (in case of *GoogLeNet::TransferLearning::Color::80-20*), hence showing strong promise of the deep learning approach for similar prediction problems. Table 1 shows the mean F_1 score, mean precision, mean recall, and overall accuracy across all our experimental configurations. All the experimental configurations run for a total of 30 epochs each, and they almost consistently converge after the first step down in the learning rate.

To address the issue of over-fitting, we vary the test set to train set ratio and observe that even in the extreme case of training on only 20% of the data and testing the trained model on the rest 80% of the data, the model achieves an overall accuracy of 98.21% (mean F_1 -Score of 0.9820) in the case of *GoogLeNet::TransferLearning::Color::20-80*. As expected, the overall performance of both AlexNet and GoogLeNet do degrade if we keep increasing the test set to train set ratio (see Figure 4(d)), but the decrease in performance is not as drastic as we would expect if the model was indeed over-fitting. Figure 4(c) also shows that there is no divergence between the validation loss and the training loss, confirming that overfitting is not a contributor to the results we obtain across all our experiments.

Among the AlexNet and GoogLeNet architectures, GoogLeNet consistently performs better than AlexNet 4(a), and based on the method of training, transfer learning always yields better results 4(b), both of which were expected.

The three versions of the dataset (color, gray-scale and segmented) show a characteristic variation in performance across all the experiments when we keep the rest of the experimental configuration constant. The models perform the best in case of the colored version of the dataset. When designing the experiments, we were concerned that the neural networks might only learn to pick up the inherent biases associated with the lighting conditions, the method and apparatus of

collection of the data. We therefore experimented with the gray-scaled version of the same dataset to test the model's adaptability in the absence of color information, and its ability to learn higher level structural patterns typical to particular crops and diseases. As expected, the performance did decrease when compared to the experiments on the colored version of the dataset, but even in the case of the worst performance, the observed mean F_1 score was 0.8524 (overall accuracy of 85.53%). The segmented versions of the whole dataset was also prepared to investigate the role of the background of the images in overall performance, and as shown in Figure 4(e), the performance of the model using segmented images is consistently better than that of the model using gray-scaled images, but slightly lower than that of the model using the colored version of the images.

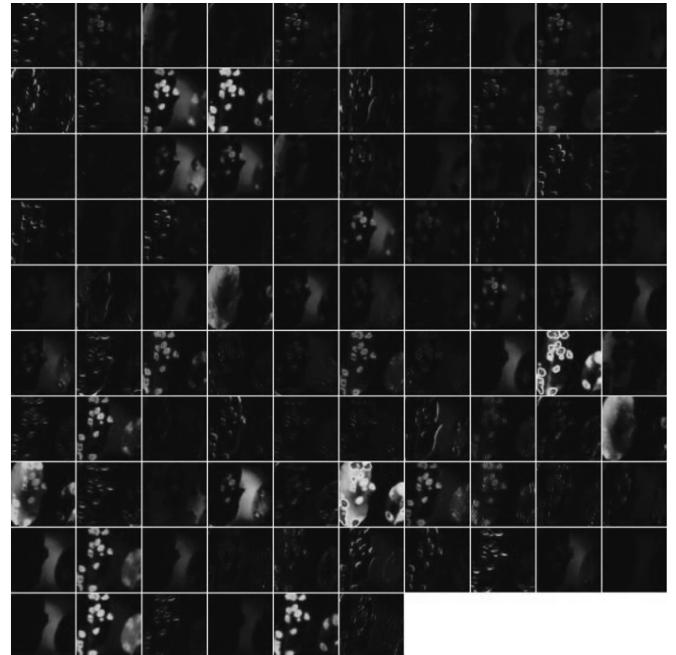
Finally, while these approaches yield excellent results on the PlantVillage dataset which was collected in a controlled environment, we also assessed the model's performance on images sampled from trusted online sources such as academic agriculture extension services. Such images are not available in large numbers, and using a combination of automated download from Bing Image Search with a visual verification step by one of us (MS), we obtained a small, verified dataset of 121 images (see Supplementary Material for a detailed description of the process). By using the model trained using *GoogLeNet:Segmented:TransferLearning:80-20*, we obtained an overall accuracy of 31.40% in successfully predicting the correct class label (i.e. crop and disease information) from among 38 possible class labels. We note that a random classifier will obtain an average accuracy of only 2.63%. When providing the information about the crop that the particular image belongs to, the accuracy increases to 47.93%. Across all images, the correct class was in the top-5 predictions in 52.89% of the cases.

Discussion

The performance of convolutional neural networks in object recognition and image classification has made tremendous progress in the past few years. [13] [14] [15] [16] [17]. Previously, the traditional approach for image classification tasks has been based on hand-engineered features such as SIFT[18], HoG[19], SURF[20], etc., and then to use some form of learning algorithm in these feature spaces. This led to the performance of all these approaches depending heavily on the underlying predefined features. Feature engineering itself is a complex and tedious process which needed to be revisited every time the problem at hand or the associated dataset changed considerably. This problem has occurred in all traditional attempts to detect plant diseases using computer vision as they leaned heavily on hand-engineered features, image enhancement techniques, and a host of other complex and labour-intensive methodologies. A few years ago, AlexNet[13] showed for the first time that end-to-end supervised training using a deep convolutional neural network architecture is a practical possibility even for image classification problems with a very large number of classes, beating the traditional approaches using hand-engineered features by a substantial margin in standard benchmarks. The absence of the labor-intensive phase of feature engineering and the generalizability of the solution makes them a very promising candidate for a practical and scaleable approach for computational inference of plant diseases.



(a) Example image of a leaf suffering from Apple Cedar Rust, selected from the top-20 images returned by Bing Image search for the keywords "Apple Cedar Rust Leaves" on April 4th, 2016. Image Reference : *Clemson University - USDA Cooperative Extension Slide Series, Bugwood.org*



(b) Visualization of activations in the first convolution layer(*conv1*) of an AlexNet architecture trained using *AlexNet:Color:TrainFromScratch:80-20* when doing a forward pass on the image in Figure 3(a)

Fig. 3. Visualization of activations in the initial layers of an AlexNet architecture demonstrating that the model has learnt to efficiently activate against the diseased spots on the example leaf

Using the deep convolutional neural network architecture, we trained a model on images of plant leaves with the goal of classifying both crop species and the presence and identity of disease on images that the model had not seen before. Within the PlantVillage data set of 54,306 images containing 38 classes of 14 crop species and 26 diseases (or absence thereof), this goal has been achieved as demonstrated by the top accuracy