Plant Pathology: Detecting Apple Plant Leaf diseases using EfficientNet B4

Jihwan Lee
Dept. of Computer Science
Sogang University, Seoul, Korea
jihwan2854@naver.com

Hamza Bashir
Dept. of Computer Science
Sogang University, Seoul, Korea
hamzakips8@gmail.com

Hamsa Priya P
Dept. of Computer Science
Sogang University, Seoul, Korea
phamsapriya@gmail.com

***Abstract*— Plant Pathology is the science that studies the causes of plant diseases. It investigates the factors behind the failure of plants to reach their genetic potential, and develops interventions to protect plants, reduce crop losses and improve food security. It can be challenging for the farmers where the plant diseases may attack them, destroy parts or all of the plants, and reduce much of their produce. However, science has a practical and noble goal of protecting the food available for humans and animals. Alike from the above goal, there was a competition hosted by Kaggle "PLANT PATHOLOGY 2021 – FGVC8" to identify the category of foliar diseases in apple trees. This competition was part of the Fine-Grained Visual Categorization FGVC8 workshop at the Computer Vision and Pattern Recognition Conference CVPR 2021. Our aim as part of the project is to help find the apple disease with better accuracy using Image classification.**

Keywords—Plant Pathology, Image classification, Fine-Grained Visual Categorization

I. INTRODUCTION

Misdiagnosis of numerous plant diseases has an influence on agricultural crops, and poor management measures result in higher input costs, more outbreaks, and substantial economic and environmental consequences Plant Pathology is the study of how diseases occur in plants and how to control them. Control methods are contingent on accurate disease identification. As a result, one of the most important aspects of Plant Pathology is diagnosis. The standard system of disease diagnosis, which involves human scouting, is time-consuming and costly. As a result of this being highlighted as one of the significant difficulties in the agriculture industry, this notion has been chosen as one of the subjects for the CVPR competition.

The competition Plant Pathology 2020 – FGVC7 was organized last year as part of CVPR 2020. This year's competition, "PLANT PATHOLOGY 2021 – FGVC8," has been updated with more photos and diseases insights and lessons learned from last year's competition.

The quantity and diversity of data obtained in recent years has been primarily contributed to recent improvements in deep learning models. Without actually gathering additional data, data augmentation allows practitioners to greatly expand the diversity of data available for training models. Cropping, padding, and horizontal flipping are examples of data augmentation procedures.

A. PROBLEM OVERVIEW

Foliar (leaf) diseases are a significant threat to apple orchard productivity and quality. This competition is being launched to find a better method using image classification to reduce the cost and time spent on manual scouting (data science). However, we encounter a few problems when employing computer vision-based disease diagnosis, such as significant variations of visual symptoms of a similar disease among different cultivars and new types that have emerged under cultivation. The differences in natural and image-capturing environments are the cause of these variations: leaf color and morphology, the age of infected tissues, non-uniform image background, different light illumination during imaging, and so on.

B. DATASET OVERVIEW

As aforementioned, a competition on Plant Pathology was held last year (2020) as part of CVPR 2020, with a dataset of 3651 RGB photos of apple foliar disease. This year's competition expands on previous year's by asking you to deal with more diseases and to provide specific information about leaves with multiple infections.

For Plant Pathology 2021-FGVC8, they have significantly increased the number of foliar disease images and added additional disease categories. The dataset contains approximately 23,000 high-quality RGB images of apple foliar diseases, including a large expert-annotated disease dataset. To overcome the problem aforementioned in previous section, the images in the current dataset were taken from real-field scenarios with non-uniform backgrounds in different maturity stages and at different times of day under different focal camera settings.

The train set metadata contains:

- **image** - the image ID
- **labels** - the target classes, a space delimited list of all diseases found in the image. Unhealthy leaves with too many diseases to classify visually will have the complex class, and may also have a subset of the diseases

identified. { 0:'powdery_mildew', 1: 'scab', 2: 'complex', 3: 'frog_eye_leaf_spot', 4: 'rust', 'healthy' }

II. RELATED WORK

Plant Leaf Disease Detection and Classification using Conventional Machine Learning and Deep Learning: Machine learning-based techniques have shown to be effective in a variety of image processing applications in recent years. Learning that is based on artificial intelligence applications has yielded positive results. Machine learning approaches train the system to learn autonomously and improve its output based on its own experiences. Fungal-like organisms have been identified to infect 85% of plants. In comparison to standard image processing approaches, machine learning and deep learning algorithms for detecting plant diseases are more accurate and take less time. In the field of plant disease, researchers face substantial challenges such as the lack of data sets for each disease, background noise in collected photos, poor resolution images, and the textural quality of plant leaves changing with the environmental changes.

Traditional methods, machine learning, and deep learning strategies for detecting and classifying plant diseases are covered in this survey. Image pre-processing, segmentation, feature selection, and classification are the four primary processes for detecting and classifying illnesses. According to the results of the survey, K-means for segmentation, SVM, and ANN are the most effective approaches for detecting and classifying diseased plants. Following a review of many deep learning research publications, it can be determined that CNN performs best in the detection and classification of plant diseases. All of the comparisons between traditional machine learning methods and deep learning methods show that deep learning is clearly superior to traditional methods. Because certain datasets were obtained in a typical setting (i.e., without noise), it's likely that when noise enters the scene, the algorithm's performance will deteriorate. After reviewing hundreds of papers, one significant issue was discovered: many academics created their own datasets that were not available to other researchers, preventing new algorithm development from being tested on datasets that were not publicly accessible. The creation of a hardware algorithm that can assist farmers in detecting and classifying diseases is a future direction.

Deep Residual Learning for Image Recognition: Deeper neural networks are more difficult to train. A residual learning framework is presented to ease the training of networks that are significantly deeper than those previously used. Rather than learning unreferenced functions, the layers are explicitly reformulated as learning residual functions with reference to the layer inputs. As a result, it presents extensive empirical evidence that these residual networks are easier to tune and can benefit from increasing depth. It analyzes residual nets with a depth of up to 152 layers on the ImageNet dataset, which is 8x deeper than VGG nets but still has lesser complexity. On the ImageNet test set, an ensemble of these residual nets achieves 3.57 % error. On the ILSVRC 2015 classification task, this result took first place. We also provide CIFAR-10 analysis with 100 and 1000 layers. For many visual recognition tasks, the depth of representations is important. It achieves a 28 percent relative improvement on the COCO object identification dataset, only due to our extremely deep representations. Deep residual nets were the backbone of their submissions to the ILSVRC and COCO 2015 competitions1, where they also took 1st place in the ImageNet detection, ImageNet localization, COCO detection, and COCO segmentation tasks.

Rethinking the Inception Architecture for Computer Vision: Since 2014, very deep convolutional networks have been popular, with significant gains in a variety of benchmarks. Although increased model size and computational cost tend to transform to immediate quality gains for most tasks (as long as enough labeled data is provided for training), computational efficiency and low parameter count are still enabling factors for a variety of use cases, including mobile vision and big-data scenarios. Exploring approaches to improve existing networks in ways that use appropriately factorized convolutions and aggressive regularization to make the extra processing as efficient as possible. On the ILSVR 2012 classification, the highest quality version of Inception-v2 achieves 21.2 percent top-1 and 5.6 percent top-5 error for single crop evaluation, setting a new state of the art. When compared to the network described by Ioffe et al., this is accomplished with a relatively small (2.5x) increase in computational cost. Training high-quality networks on small training sets is achieved due to the combination of lower parameter count and further regularization with batch-normalized auxiliary classifiers and label-smoothing.

EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks: Convolutional Neural Networks (ConvNets) are often built with a fixed resource budget and then scaled up for higher accuracy when more resources become available. This paper examines model scaling in detail and discovers that carefully balancing network depth, width, and resolution can enhance performance. Based on this finding, a new scaling method is introduced that uses a simple yet very effective compound coefficient to equally scale all depth/width/resolution dimensions. It shows how effective this method is at scaling up MobileNets and ResNets. Furthermore, neural architecture search is utilized to create a new baseline network and scale it up to create the family of models, called EfficientNets, which outperform earlier ConvNets in terms of accuracy and efficiency. In particular, EfficientNet-B7 achieves state-of-the-art 84.3% top-1 accuracy on ImageNet, while being 8.4x smaller and 6.1x faster on inference than the best existing ConvNet. With an order of magnitude fewer parameters, these EfficientNets also transfer well and attain state-of-the-art accuracy on CIFAR-100 (91.7%), Flowers (98.8%), and three other transfer learning datasets. It proves that a mobilesize EfficientNet model can be scaled up very successfully, surpassing state-of-the-art accuracy with an order of magnitude fewer parameters and FLOPS, on both ImageNet and five frequently used transfer learning datasets, due to this compound scaling method.

EfficientNets have shown a very good accuracy in multiple reallife problems like it was used for the automatic diagnosis of COVID-19, skin lesion classification, breast cancer detection, fruit recognition and classification of hematoxlin from images [7][8][9][10][11].

III. METHODOLOGY

In this section, the model architecture is discussed in detail including the dataset description, data augmentation and model architecture diagram. The whole architecture is shown by the following diagram.

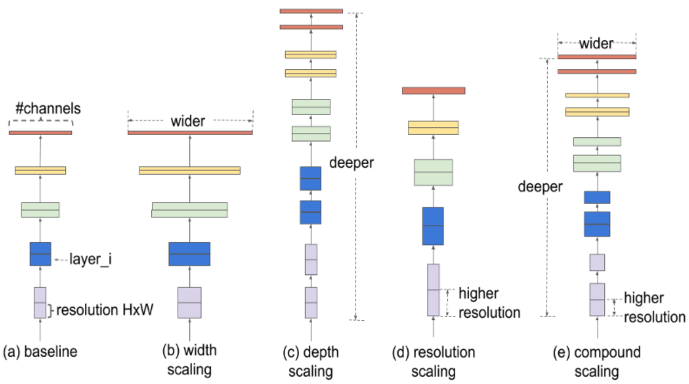


Figure 1: Architecture diagram of efficientnetb4

C. Dataset

The Dataset consist of images of plants diseases that we need to classify. The competition named as Plant Pathology is actually an update to the previous year challenge on Kaggle. The goal of the challenge is to identify the apple diseases by using the images of the leaves.it consist of 5 different classes. The total dataset has 23,000 RGB images of foliar disease in apple. The class distribution of the dataset can be shown in the following figure.

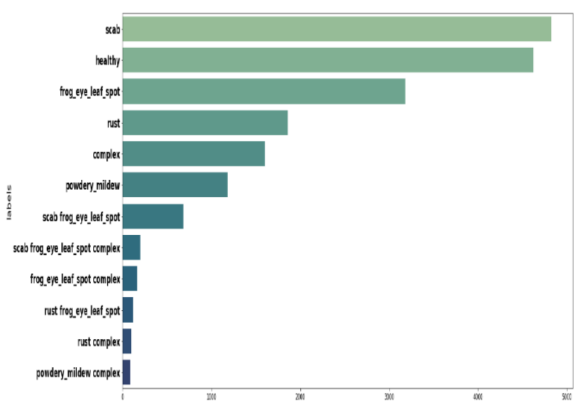


Figure 2: class distribution

D. Data Preprocessing

In order to train the model, we have done a lot of data processing so that we can get make our dataset in required format and make the model run on TPU more efficiently. We have done the following processing on the plant's dataset.

- Removed duplicates from the dataset.
- Label formatting using Multi-label Binarizer.
- Make stratified folds.
- Resized the images to 600x600.
- One hot encoding
- Make TFRecords.

	complex	frog_eye_leaf_spot	powdery_mildew	rust	scab	healthy
image						
800113bb65efe69e.jpg	0	0	0	0	0	1
8002cb321f8bfcdf.jpg	1	1	0	0	1	0
80070f7fb5e2ccaa.jpg	0	0	0	0	1	0
80077517781fb94f.jpg	0	0	0	0	1	0
800cbf0ff87721f8.jpg	1	0	0	0	0	0

One-Hot Encoding

Figure 3: label Formatting

The following figure 4 shows the distribution of the classes when we divided them into 5 folds so, we can apply cross validation and test our trained model on all the sets.

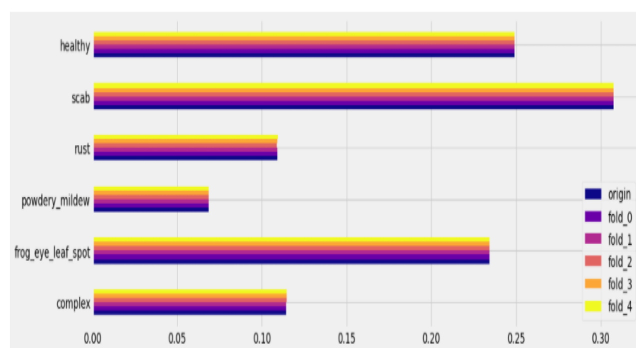
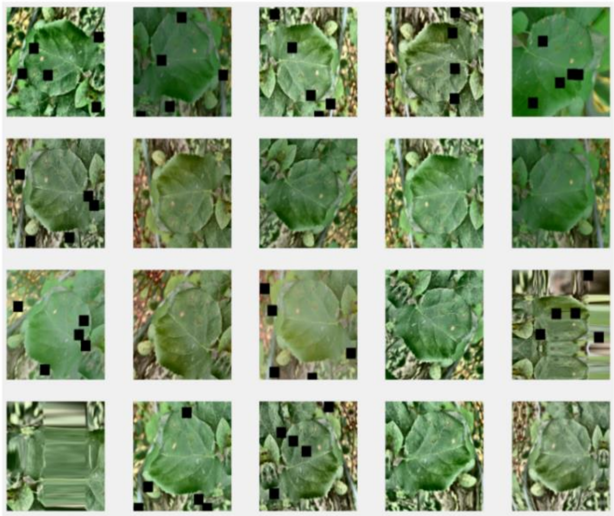


Figure 4: 5 folds TFRecords

E. Data Augmentation

In order to achieve high accuracy. Data augmentation played a major role in it [6]. As we can see from the class distribution diagram the dataset was not balanced so, there were some chances that model would converge to particular class but, we need a more generalized model [5]. Applying data augmentation with image data generator class in keras was also an option but it was really time consuming while training the model and can affect the TPU performance so, we have used Albumentation library to apply the data augmentation like vertical flip, horizontal flip, random rotate, shift scale rotates, median blur etc. following diagram, shows some data augmentation results.



F. *EfficientNet b4 Architecture*

The basic architecture of the model that shows the scaling is given in figure 1. we used a new method for a scaling up of the model, which makes use of a simple but effective CNN factor analysis in a more structured way. In contrast to the traditional methods used in any network, efficientnet set up the dimensions, such as width, height, and resolution, the model scale each dimension is evenly spread, with a fixed set of scaling factors. On the basis of the new scaling method, and the recent advances in AutoML, our method belongs to a family of models is referred to as EfficientNets that 10 times smaller and faster.

Compound Scaling:

To understand the effect of the network on the scale, they have investigated the effect of scale model sizes. At the scale of the individual dimensions of the model's performance, we observed that the consideration of all the dimensions of the network width, depth, and clarity with the best available resources, to enhance the overall system performance. The first step in a full zoom and a method for performing a grid-search on the relationship between a variety of scalable dimensions, the underlying network is under constant resource constraints. This will determine the appropriate zoom level for each of the dimensions mentioned above.

The range of the efficiency of the model is highly dependent on the underlying network. Therefore, in order to further improve efficiency, they developed a new core network, the search for a neural architecture with the AutoML MNAS framework, which allows for optimal accuracy and efficiency. This architecture allows for the use of a mobile Inverted bottleneck convolution (MBConv), which is comparable with MobileNetV2 and MnasNet, but it is slightly larger due to the increase in the project on the FLOP.

IV. Experiments and Evaluation

G. *Training details and accuracy*

We developed our model on Keras framework and choose Adam Optimizer for training of our model. We trained our model on machine with two Intel Xeon E5-2620v3 CPUs and TPU. We used batch size of 64 and trained the model for 20 epochs on TPU.

H. *Accuracy on Dataset*

The total dataset has 23,000 RGB images of foliar disease in apple. There are total 5 different classes so we added one dense layer on top of the efficient model with 5 output nodes. We have trained the model and evaluate it using 5 fold cross validation so, we achieved 0.86 F1 score on the test data. We have just shown the fold 0 and fold 4 below.

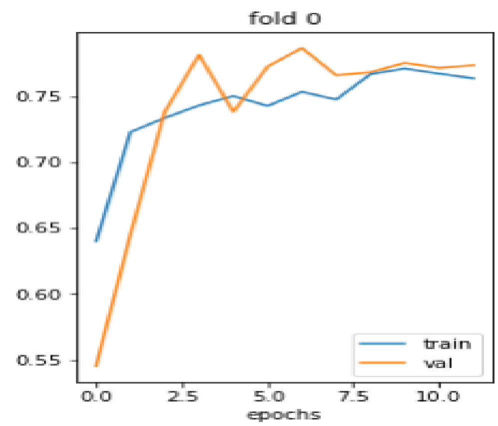


Figure 5: Fold 0 accuracy

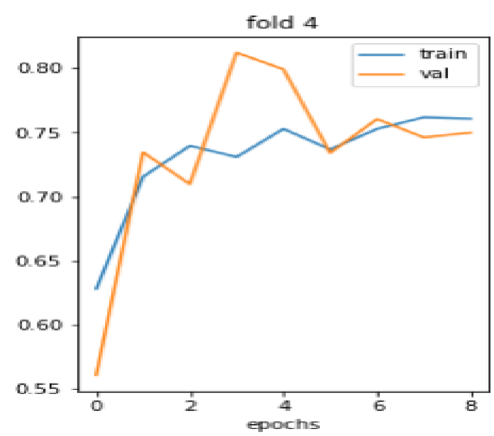


Figure 6: Fold 4 accuracy

V. Conclusion and discussion

Apples are one of the most important temperate fruit crops around the world. Needles and diseases are a serious threat to the productivity and the quality of the grounds. The modern process of diagnosing diseases such as gardens, it is based on the handbook of human exploration, that is, it is a time-consuming and expensive. Even though computer vision-based models have shown promise in the identification of plant pathogens, there are a few limitations that need to be resolved. We have used the dataset provided by the Kaggle community. That dataset consists of 23,000 RGB images. We have used applied data augmentation to increase out training data and then we trained Efficient net model on our dataset. The model shown a great performance on the test data with an F1 score of 0.86. the trained model was well tested with the help of 5-fold cross validation technique. So, we can use this computer vision method to detect the disease accurately.

VI. References

[1] Hardikkumar S. Jayswal and Jitendra P. Chaudhari. "Plant Leaf Disease Detecetion and Classification using Conventional Machine Learning and Deep Learning." *International Journal on Emerging Technologies* 11(3): 1094-1102(2020).

[2] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep Residual Learning for Image Recognition"

[3] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. " Rethinking the Inception Architecture for Computer Vision".

[4] Mingxing Tan, Quoc V. Le. "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks " *International Conference on Machine Learning, 2019*, ICML 2019.

[5] Van Dyk, David A., and Xiao-Li Meng. "The art of data augmentation." *Journal of Computational and Graphical Statistics* 10.1 (2001): 1-50.

[6] Perez, Luis, and Jason Wang. "The effectiveness of data augmentation in image classification using deep learning." *arXiv preprint arXiv:1712.04621* (2017).

[7] Marques, Gonçalo, Deevyankar Agarwal, and Isabel de la Torre Díez. "Automated medical diagnosis of COVID-19 through EfficientNet convolutional neural network." *Applied Soft Computing* 96 (2020): 106691.

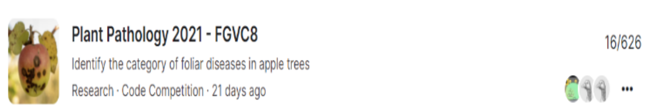
[8] Gessert, Nils, et al. "Skin lesion classification using ensembles of multi-resolution EfficientNets with meta data." *MethodsX* 7 (2020): 100864.

[9] Wang, Jun, et al. "Boosted efficientnet: detection of lymph node metastases in breast cancer using convolutional neural networks." *Cancers* 13.4 (2021): 661.

[10] Duong, Linh T., et al. "Automated fruit recognition using EfficientNet and MixNet." *Computers and Electronics in Agriculture* 171 (2020): 105326.

[11] Munien, Chanaleä, and Serestina Viriri. "Classification of Hematoxylin and Eosin-Stained Breast Cancer Histology Microscopy Images Using Transfer Learning with EfficientNets." *Computational Intelligence and Neuroscience* 2021 (2021).

VII. Leaderboard and Contribution



We ranked in 2.6%. (16/626)

Contribution:

Jihwan Lee: 40%
Hamza Bashir: 28%
Hamsa Priya P: 32%