Summer 2021 CS 687 Capstone Project Progress Report
A Deep Learning Model for Multiple Apple Foliar Diseases Identification

Shanshan Yu
Advisor: Sam Chung, P.hd., Brian Maeng, P.hd.
MS in Computer Science
School of Technology and Computing (STC)
City University of Seattle (CityU)
yushanshan@cityunviersity.edu, maengjooyol@cityu.edu

Abstract

Apple orchards are facing threats of pathogens and insects, which costs millions of dollars lost every year in the United States. The current apple disease diagnose process is based on naked eyes. The time-consuming process requires apple farmers to have professional knowledge about various apple diseases. Incorrect diagnosis will cause chemical abuse, environmental pollution, and financial loss. Because of the involution of deep learning, fast and efficient plant diseases detection powered by a deep learning model is possible. In the past few years, many high accurate deep learning models achieved high performance on apple foliar diseases classification problems. But these models have a limitation that is they only can detect a single apple disease. The paper focuses on implementing a deep learning model for multiple apple foliar diseases identification problems according to the Plant Pathology 2021 competition on Kaggle. The model achieved high accuracy on both singly and multiply apple foliar diseases identification problems.

Keywords: apple orchards; computer vision; convolutional neural network; disease classification; deep learning.

1. INTRODUCTION

Apples are one of the most important fruit crops in the world (Statista, 2021). However, the productivity and quality of apple orchards are treated by apple foliar diseases. The current process of disease diagnosis in crops is highly dependent on manual work, which is time-consuming and expensive (FineGrainedVisualCat, 2021). The success of disease management is related to early disease detection, any delay and incorrect diagnosis can lead to either inadequate use of chemicals with increased costs and environmental impacts (Thapa, et al., 2020). Introducing machine learning models into apple orchards can decrease the loss of apple diseases and increase the productivity of apple orchards.

In 2020 and 2021, Fine-Grained Visual Cat hosted two Plant Pathology competitions on Kaggle. The purpure of the competition is to build an accurate machine learning model for plant disease identification. The first FGVCG competition in 2020 focus on single disease identification, and the winner's score was 0.98445 out of 1. However, the model had a limitation because it's designed for single disease identification. And for a special class "complex" – undefined apple foliar diseases or multiple apple disease, the model's performance dropped from 98% to 51%(Thapa, et al., 2021). Because of the limitation of the single-label classification model, the plant pathology 2021 competition asked participants to design a multi-labeling classification model. The new dataset introduced multi-labeled apple foliar diseases images classified by plant diseases experts. The winner of the competition obtained a score of 0.88336.

The paper will develop a multiple foliar diseases identification model based on a shared notebook (Allen, 2020) from the competition. The paper also will combine the advantages of the top winners' solutions in the Plant Pathology 2021 competition to achieve higher performance.

1.1 Problem Statement

The research problem is how to make a highly accurate multiple apple foliar diseases identification model.

1.2 Motivation

According to the United Nations survey (United Nations, 2021), 8.9% of the world's population suffers from hunger. AI can increase crop productivity by identifying crops diseases in the earlier stage. With higher crop productivities, we can feed more people in the world. Besides, the AI makes the crop production process trackable, let more people can eat high-quality, healthy foods.

1.3 Approach

The model will be trained on a pre-trained model efficientNetV2. The validation strategy is K-folds cross-validation with different k values. In the data preprocessing step, the unqualified images will be removed from the dataset. The string labels will be converted to numerical labels. In the training process, the model will introduce the winner's data argument strategy, which is to combine four random training set's images as multiple labels images (Hao, 2021). The prediction layer's activation function is the sigmoid function, and the unit number is equal to the classes number 5. The healthy class is excluded in the prediction classes because if all 5 classes are false, the prediction reveals the prediction result should be healthy.

1.4 Conclusion

The result shows the best model is trained on 2 cross-validation folders on Kaggle, which obtained 83% f1 scores. The models trained on Colab show lower f1 scores from 75% ~ 79%, and the one trained with data argument shows the lowest score 61%.

2. BACKGROUND

In the computer vision field, Convolutional Neural Network(CNN) has been widely used to solve image classification problems. The most common technique for image classification problems is transfer learning, which takes advantage of pre-trained models. If the training set is small, transfer learning regularly provides a better result compared to training a model from scratch. Transfer learning with CNN provides great performance. Many research papers utilized transfer learning on the single apple foliar disease classification problem and achieved a decent result.

However, CNN is not designed to solve multi-label classification problems. The CNN performance on the multi-label classification problem is not satisfied.

	Liu, B., Zhang, Y., He, D., & Li, Y. (2018).	Thapa, R., Zhang, K., Snavely, N., Belongie, S., & Khan, A. (2020).	The article's solution
Problem	Single disease identification for apple foliar disease	Single disease identification for apple foliar disease	Multiple diseases identification for apple foliar disease
Method	Transfer learning + softmax layer	Transfer learning + softmax layer	Transfer learning + sigmoid layer
Single Label Classification	Support	Support	Support
Multiple Label Classification	Not support	Partially support	Support

Even though many papers are researching the multi-labeled classification problem, the multi-label classification model for plant diseases hasn't been discovered yet.

3. RELATED WORK

The U.S apple industry is under constant threats from apple diseases caused by insects, fungal, bacterial, and viral pathogens. Apple diseases identification process highly depends on professional knowledge, and apple trees also can have complex symptoms caused by multiple apple diseases, which make the identification process more difficult. Currently, several deep learning models have been trained successfully for single plant disease classification. Many of the models can achieve 97~99 accuracy using CNN and transfer learning (Liu., et al, 2018). However, those models can't support multiple diseases classification. In the Plant Pathology 2020 challenge, the organizer asked participants to train a single-disease identification model for apple foliar diseases identification. The data instance with the label "complex" represented an apple leaf's instance with multiple apple diseases. Even though the competition's winner solution achieved overall 97% accuracy, it only achieved 51% accuracy on the "complex" label (Thapa., et al, 2020).

Literature Review

There are many pre-trained models available for transfer learning such as AlexNet, GoogLeNet, VGG, ResNet, InceptionNet. These models are trained on a big-scale dataset called ImageNet. For the apple disease identification problem, many researchers have trained single plant disease identification models and achieved 93 ~ 99% accuracy (Hungilo, Emmanuel & Emmanuel, 2019). However, there is little research about training models for multiple apple diseases identification problems.

Recently, there are several papers focus on the multi-labeling classification problem(Bashmal et al, 2021). The interested multi-labeling classification includes identifying objects on the landscape, UVC images. Bashmal mentioned — even though CNN can extract high-level features effectively, CNN is not originally designed for a multi-labeling classification task.
There are two classical approaches for multi-labeling classification problems (Karalas et al, 2016). The first approach is transformation, which reconstructs the multi-labeling classification problem into a multiple binary classification problem. But the disadvantage of this approach is it doesn't perform well on a large set of classes. The second approach is algorithm adaptation. This approach modifies multi-class classifiers so the classifiers can be applied to multi-label problems.
In Bazi's papers, he mentioned the start of the art approach transform net, which achieve a significant improvement in recent researches (Bazi, 2021). However, the paper will not discuss the details in this paper, since the paper will use another approach.
In 2021, Fine-Grained Visual Cat hosted the Plant Pathology competition on Kaggle. The competition focuses on multiple apple foliar diseases identification problems. The top3 solutions have their high-light points such as data argument strategy, pre-trained model, model embedding method.
The top3 scores of the competition are 0.88336, 0.87987, and 0.87560. The winner adopted transfer learning with efficientnetv2 and ResNet. The solution used a special multi-label label data argument strategy, in which the new instance was generated from multiple single-labeled instances using the mosaic approach. The second place combined the Plant Pathology 2020 and 2021 datasets as the training set. Its model is trained the model using SEResNeXt50 with SwinT. The third-place embedded models trained on ResNet50 and Seresnext50.

4. APPROACH

4.1 Data Collection

The training images are collected from Kaggle competition Plant Pathology 2021 – FGVC8(Thapa et al, 2020). There are 18,632 images available for training. Those images are taken by a high-resolution camera and have various angles, light illumination, and backgrounds. These images are labeled by apple plant disease experts. Each image can have single or multiple labels that indicate the leaf's diseases.

4.2 Data Analysis

The FGVC8 dataset consists of 18,632 4000 x 2672 RGB images, and it will be used for the training and cross-validation processes. A CSV file stores images' names and the images' labels. The diseases' labels include 5 common apple foliar diseases as scab, rust, frog eye leaf spot, powdery mildew. The complex class refers to an undefined disease or multiple diseases. The healthy class indicates healthy leaf, and the healthy class can't be combined with other labels. Figure 1 shows some examples of images in the FGVC8 dataset.

Figure 1. Example images in the FGVC8 dataset.

Figure 2 shows the multi-labeled images are much less than singly labeled images. 17,277 images are singly labeled and 1,355 images are multi-labeled.

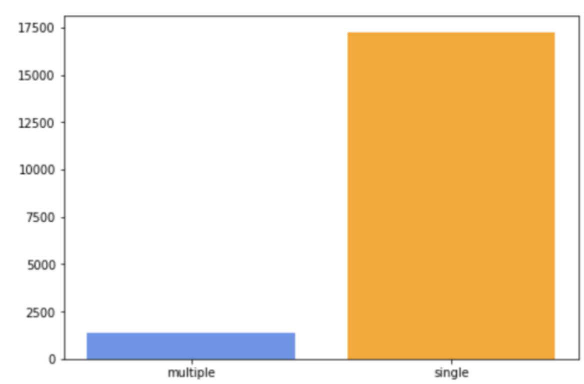

Figure 2. The ratio of multi-labeled images and singly labeled images

Figure 3 shows the number of labels in the dataset. The blue indicates the label belongs to a singly labeled image, and the orange means the label belongs to a multi-labeled image. 4,624 images are healthy and singly labeled, 5,712 images have scab disease, in which 4,826 images are singly labeled and 886 images are multi-labeled. 2,151 images have a complex disease, and 1,602 of them are singly labeled, 549 are multi-labeled. 4,352 images have frog eye leaf spot disease, 3,181 of them are singly labeled, 1,171 of them are multiple labeled. 2, 077 images have rust disease, 1,860 images of them are singly labeled, 217 of them are multi-labeled.

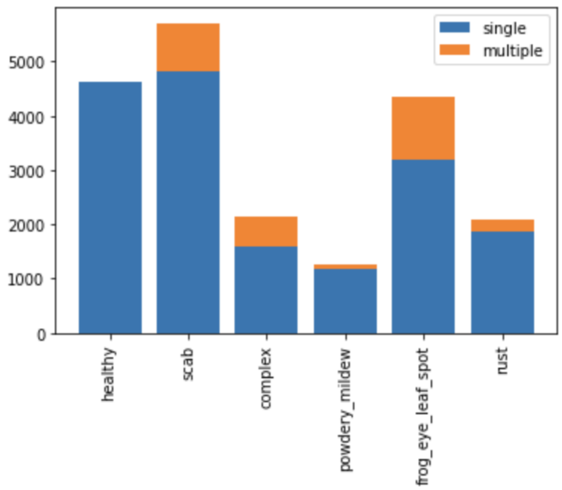

Figure 3. Labels Counts

The dataset also has 104 pairs of duplicate images (Kuzmenkov, 2021). Those images have identical pixel values but different labels. To make the model more accurate, those images have been removed from the training dataset. Figure 4 shows an example of duplicate pair with inconsistent labels.

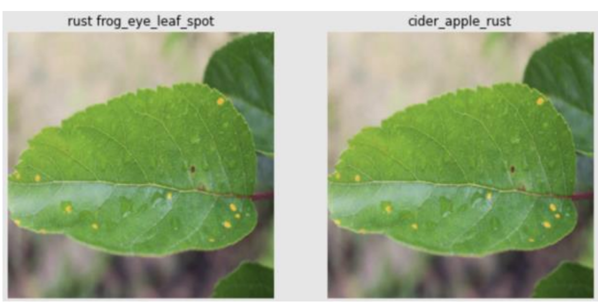

Figure 4. An example of duplicated image

4.3 Implementation

The models were trained on a Cloud notebook environment hosted by Kaggle and Colab. The model training and prediction code is modified from a shared notebook written by Allan's team (Allan, 2021). These notebooks are available via the links in the demo section (Shanshan, 2021). Due to the requirements of the competition rules that don't allow the submission notebook to run bigger than 9 hours, the notebooks are separated into two parts: training and prediction.

The deep learning model consists of an efficient net layer and a prediction layer. The efficient net layer is used to provide the extracted features for the input image, the prediction layer with an activation function sigmoid is used for converting extracted features to predictions.

Each input image in the training set have been normalized by the following formula:

$$Img = \frac{Img - \overline{IMG}}{\sigma(IMG)}$$

Because the training dataset is imbalanced, the author chose Focal Loss as the model's loss function. The Focal Loss formula is as follows:

$$FocalLoss(p_t) = -\alpha(1 - p_t)^{\gamma}\log(p_t)$$

The author has generated 3,000 images that concatenated from 3, 000 * 4 images. Figure 5 shows an example of a concatenated image.

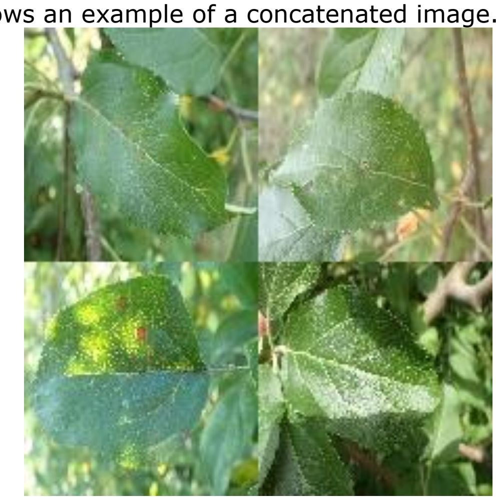

Figure 5. An example of a concatenated image

To compare the performance of the methods, the author trained the model without concatenated image first, then trained another model with a concatenated image. For each condition, the paper also uses different k values for cross-validation and compared the result. Due to Kaggle only allows 37 hours of GPU usage on notebooks, one of the models is trained on Colab.

The f1 scores comparison between the different models is shown in figure 6.

id	Platform	K	With Concatenated Image	CV F1 score	The public test set F1 score	The private test set F1 score
1	Kaggle	2	False	0.92	0.76	0.77
2	Colab	5	False	0.84	0.79	0.77
3	Kaggle	2	True	0.84	0.70	0.71

Figure 6. Models' comparison

8. RESULT ANALYSIS

The best result of these models is 0.79, and it was trained without concatenated images. According to the figure, the concatenated images decrease the model's performance. The reason may be the concatenated images are not similar to the images in the test set, including concatenated images may introduce noise into the training process, and the model learned unrelated patterns from the dataset with generated data.

9. CONCLUSION

The paper trained a deep learning model for a multi-labeled classification problem using transfer learning and a sigmoid layer with multiple output units. In the implementation process, the paper trained the models with and without the concatenated images. The best model achieves 79% f1 scores and it was trained without concatenated data. Concatenated images didn't improve the model performance as it's expected.

10. FUTURE WORK

Due to the limitation of time, the paper didn't find the optimal hyperparameters for the model. For the data argument strategy, the winner solution also provides other approaches such as ½ * 2 concatenation, and ¼ x 4 concatenation. In the future, the author will train models with different data argument approach, and check if the performance increases.

11. REFERENCE

Allan. (2021). PytorchLighting_EfficientNet_FocalLoss[Training]. Retrieved from: https://www.kaggle.com/crissallan/pytorchlightning-efficientnet-focalloss-training

Bashmal, L., Bazi, Y., Mohamad Mahmoud, A. R., Alhichri, H., & Ajlan, N. A. (2021). UAV Image Multi-Labeling with Data-Efficient Transformers. Applied Sciences, 11(9), 3974. http://dx.doi.org/10.3390/app11093974

Bazi, Y., Bashmal, L., Rahhal, M. M. A., Dayil, R. A., & Ajlan, N. A. (2021). Vision transformers for remote sensing image classification. *Remote Sensing* (Basel, Switzerland), 13(3), 516. https://doi.org/10.3390/rs13030516

FineGrainedVisualCat. (2021). *Plant Pathology 2021 – FGVC8.* Retrieved from: https://www.kaggle.com/c/plant-pathology-2021-fgvc8/overview

Hao, W. (2021). *1st place solution.* Retrieved from: https://www.kaggle.com/c/plant-pathology-2021-fgvc8/discussion/243042

Hungilo, G. G., Emmanuel, G., & Emanuel, A. W. (2019, April). Image processing techniques for detecting and classification of plant disease: a review. In Proceedings of the 2019 international conference on intelligent medicine and image processing (pp. 48-52). https://doi.org/10.1145/3332340.3332341

Karalas, K., Tsagkatakis, G., Zervakis, M., & Tsakalides, P. (2016). Land classification using remotely sensed data: Going multilabel. IEEE Transactions on Geoscience and Remote Sensing, 54(6), 3548-3563.

Kuzmenkov, N. (2021). *50 duplicate image pairs were revealed.* Retrieved from: https://www.kaggle.com/c/plant-pathology-2021-fgvc8/discussion/227829

Liu, B., Zhang, Y., He, D., & Li, Y. (2018). Identification of apple leaf diseases based on deep convolutional neural networks. *Symmetry*, 10(1), 11. http://doi.org/10.3390/sym10010011

Shanshan, Y. (2021). PytorchLighting_EfficientNet_FocalLoss[Training]. Retrieved from: https://www.kaggle.com/bcghost/pytorchlightning-efficientnet-focalloss-training

Shanshan, Y. (2021). Plant Pathology. Retrieved from: https://www.kaggle.com/bcghost/plant-pathology

Thapa, R., Zhang, K., Snavely, N., Belongie, S., & Khan, A. (2020). The Plant Pathology Challenge 2020 data set to classify foliar disease of apples. *Applications in Plant Sciences*, 8(9), e11390. https://doi.org/10.1002/aps3.11390

World Food Programme. (2020). Global Report on Food Crises. Retrieved from: https://www.wfp.org/publications/2020-global-report-food-crises

12. DEMO

Video link: https://youtu.be/MAjTME433ss