Plant Pathology

Qianwen Luo
//Affiliation
//Address
qwluo@umich.edu

Abstract

The plant pathology is an important problem of research interest. The plant pathology challenge 2020 focuses on training a model to classify foliar diseases of apples given the photos of the apple leaves. In this task, data insufficiency and data imbalance should be addressed. To tackle with the above issues, we implement a series of data augmentation operations including illumination, contrast adjustment, flipping, rotation, cropping and blurring to enrich the dataset. We use ResNet as the model framework and use 5-fold cross validation to train the model. Experiment results show that the proposed methods can achieve good results.

1 Introduction

Plant pathology is an issue of significant importance. In fields and orchards, crops and fruit trees are constantly under threat from various diseases which affect the plant growth and may result in economic loss. To prevent the diseases from outbreaking and causing great damage, it is important to spot, detect and diagnose the diseases in the early stage. Misjudge or delayed diagnosis can lead to misuse of chemicals, increased costs and environmental side-effects. Therefore, timeliness and accuracy are both required in this task. Relying purely on human effort to detect plant diseases, though highly accurate, falls short in terms of time efficiency and is very costly.

Recently, with the development and wide use of the internet and artificial intelligence, researches focusing on combing AI and agriculture have also experienced a rapid boost. Models based on deep learning and computer vision technology have been studied and developed to solve the plant pathology detection problem. They often take the photo of a plant as the input and predict whether it is healthy or what kind of disease it is infected with. This approach can save manual labor, increase efficiency and reduce cost. However, there are still some remaining challenges. Due to the difficulty of gathering photos and classify labels, data insufficiency and data imbalance is very common in practical scenarios. Besides, variance in symptoms, genetic variations in plants, light conditions when taking photos all pose great challenge to the accurate detection of the model. Efforts are still needed to further the current researches.

‘Plant Pathology Challenge’ is a task in the above field. This task focuses on classifying foliar diseases of apples by training a model. Given a photo of one apple leaf or a bunch of leaves, the model should be able to accurately assess its/theirs health condition, and distinguish its label between four classes including healthy, rust, scab and combination (which means there are more than one disease). This task poses several challenges, and it requires the model to be able to not only assign a given image to a category based on the training set, but also to handle rare categories and new symptoms and can tackle the different angle, light condition, depth perception of the photos. Mean column-wise ROC AUC which is the average ROC AUC for each label, is used as the evaluation metric to measure the performance of the model.

In this task, 1821 samples are provided in the training set, with a class distribution of healthy: rust: scab: multiple = 516: 622: 592: 91, an approximate of 6: 6: 6: 1. With the fourth class containing significantly less amount of data, this dataset is an imbalanced dataset. Besides, through the observations of the training set, we find that there are duplicate images with different labels, meaning the dataset is also noisy. In order to address the data insufficiency and data imbalance problem and to increase the robustness of the model, we apply data augmentation to expand and enhance the original dataset. We find that many images in the dataset contain the background area with the main leaf not in focus, so rescaling and resizing are needed. Besides, the light intensity varies in different photos, which can be mitigated by changing illumination and contrast of the photos. Taking those concerns into consideration, we employ a series of operations to augment the dataset including random illumination enhancement, random contrast enhancement, up-and-down flipping, left-right flipping, random rotation and cropping and several blurring operations.

In the model training, we chose ResNet as the model backbone and use cross-entropy as the training loss. We perform a 5-fold cross-validation to improve the model performance and robustness. The experiment result shows the effectiveness of our methods. 5-fold cross-validation yield result of 96.97%

2 Method

In this section, we give a more formal and detailed description of the methods we use in the project.

2.1 data augmentation

Data augmentation is a set of techniques frequently used in deep learning that involves generating new, synthetic data to enlarge the size of the original training dataset. Data augmentation introduces variation and diversity into the dataset, thus serves the purpose of reducing overfitting, increasing the robustness of the model and improving its generalization performance. This technique is crucial and usually applied to deal with data insufficiency and imbalanced data distribution. In computer vision, the artificial data is usually generated by modifying or synthesizing the existing data, common operations including flipping, scaling, rotating the picture and adjust its colors.

In our project, we use the python library Albumentations to transform images for data augmentation. We employ several augmentation operations to the apple leaf dataset including random illumination enhancement, random contrast enhancement, up-and-down flipping, left-right flipping, random rotation and cropping and blurring operations that involve applying various blurring techniques, such as gaussian blurring, median blurring, and so on.

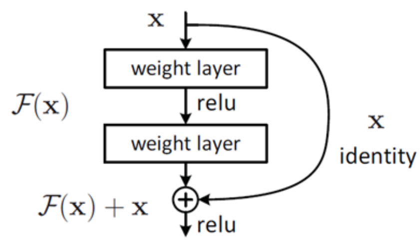
By applying these data augmentation operations, we expand and enhance the original dataset.

2.2 Resnet 18

Convolutional neural networks are widely used in computer vision, employed to tackle various problems including image processing, classification, segmentation and other important tasks. Composed of convolutional layers that essentially slide a filter over the input image, convolutional neural networks are enable to learn the features within the receptive field. With stacked convolutional layers, CNNs can effectively learn hierarchical features from local regions to the whole picture, yielding great performance in image-related tasks.

In our project, we choose the deep residual neural network ResNet18 as our model backbone.

ResNet18, proposed by "Deep Residual Learning for Image Recognition", is a deep convolutional neural network with residual design for image recognition tasks. Deep neural networks face the problem of degradation. When the number of layers continues to increase, the performance increases for a while then begins to decrease. The gradient in deep neural networks may vanish or explode, impeding the proper training of the networks. To deal with this problem, the residual network introduces the design of residual learning. The core building block of the ResNet architecture is the residual module, which consists of two or more convolutional layers and a shortcut connection or skip connection that directly connects the input to the output. With the residul connection, instead of learning the neural function directly, the neural layers only have to learn the residual between the input and the output. The output of the module F(x)+x is the sum of the input x and the output of the convolutional layers F(x). This design allows for the gradients to flow more easily through the network, preventing the banishing problem and enabling the network to learn more complex and deeper representations, drastically improving model robustness and performance. The design of the residual module is shown below:



ResNet18 is comprised of multiple layers among which there are 18 layers with weights including 17 convolutional layers and 1 fully-connected layer. ResNet18 has been pre-trained on image classification task on ImageNet data and is widely utilized as a backbone network in various computer vision applications, including image classification.

In this project, we fine-tuned the ResNet18 model to fit our specific task of classifying foliar diseases of apples. This involved replacing the final layer of the ResNet18 architecture with a new layer that outputs the four-class prediction, and training the entire model using the augmented dataset and cross-entropy loss. Fine-tuning the pre-trained ResNet18 model leverages the knowledge learned from the pre-trained task and adapts it to our specific needs, providing a strong initialization for the model and enhancing its performance.

2.3 k-folds cross validation

Cross validation splits the original training dataset into k subsets or folds. During training, one fold will be the validation set while the remaining k-1 folds will be the training set. This process is repeated k times, then the average result of the k-fold cross-validation is used as a more robust estimate of the model's performance. Stratified k-folds cross-validation stratifies the dataset such that each fold contains roughly the same proportions of each class in the original dataset, ensuring that the model is trained and evaluated on diverse samples. This technique is particularly useful when the dataset is imbalanced. Therefore, in our project, we apply stratified k-folds cross-validation with k=5 to tackle with the imbalanced data distribution challenge.

3 Experiments

In this section, we first describe the details of experiment settings, and then give the experiment results and corresponding analysis.

3.1 Experiment setting

We carry out experiments on the given plant pathology dataset. We train the model on the augmented training set. The sample submission dataset is used as the validation set. We record the loss and the model performance during training and testing.

For the dataset, we carry out previously mentioned augmentations using python package albumentation and adjust the input image to size 512x521.

Throughout our experiments, we choose the default resnet18 in package torchvision with pre-trained parameters as our model backbone and finetune the model on the given dataset. For hyperparameters and training configurations, we use cross entropy as the model loss and choose Adam as the optimizer with a learning rate of 5e-5. The model is trained for 10 epochs with batch size of 32. The random seed is set to 42. We applied a scheduler of linear warmup in the first 2 epochs and linear decay in the following steps. We implement vanilla k-fold and stratified k-fold cross validation with k=5 during training during training respectively and compare their results.

For evaluation, since the task is a multi-label classification problem, we use the macro average of area under the ROC curve as the evaluation metric.

3.2 Experiment result

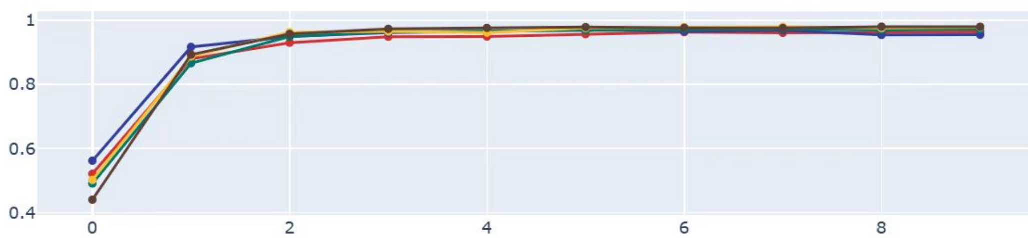
When using vanilla k-fold cross validation, the model yields a score of 96.86%.

Table 1 presents the results of the 5 folds when using stratified k-fold cross validation and Fig 1 presents the valid score during training for the 5 folds. We observe that the model converges quickly with no overfitting or underfitting. The final score of 5-folds cross validation is 96.97%.

Table 1

Fold	Valid score
1	96.35%
2	95.46%
3	97.21%
4	97.84%
5	97.94%

Fig1



4 Analysis

In above experiments, comparing the results yielded from using standard K-fold cross validation and stratified K-fold cross validation. We can observe that stratified cross validation surpasses the standard one because it takes into consideration the proportions of different classes in the given dataset and is more suitable for datasets with imbalanced data distributions like the given plant pathology dataset.

5 Conclusion

Plant pathology is an important issue that requires timely and accurate detection. Deep learning models based on computer vision technology have been developed to detect plant diseases from photos. The "Plant Pathology Challenge" task focuses on classifying foliar diseases of apples on a dataset of 1821 samples with an imbalanced class distribution. To address these issues, we apply data augmentation and use ResNet as the model backbone with cross-entropy as the training loss. We perform 5-fold cross-validation and the results show that the methods used are effective, achieving a result of 96.97%. In summary, the paper demonstrates the potential of using AI and machine learning techniques to improve plant pathology detection and highlights the importance of addressing data insufficiency and imbalance to improve model robustness. The paper provides insights into the challenges and methods and propose a effective solution.