Classification of Foliar Diseases in Apple Leaves using Natural Language Processing

Qianwen Luo, Nathan Fox

Undergraduate Research Opportunity Program, Michigan Institute for Data Science

ABSTRACT

This project 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.

INTRODUCTION

- Given a photo of one apple leaf or a bunch of leaves, the model should be able to accurately assess its health condition, and distinguish its label between four classes including healthy, rust, scab and combination (which means there are more than one disease).
- 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.
- The light intensity varies in different photos, which can be mitigated by changing illumination and contrast of the photos.

METHODS

Resnet 18

- When the number of layers continues to increase, the performance increases for a while then begins to decrease.
- 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.

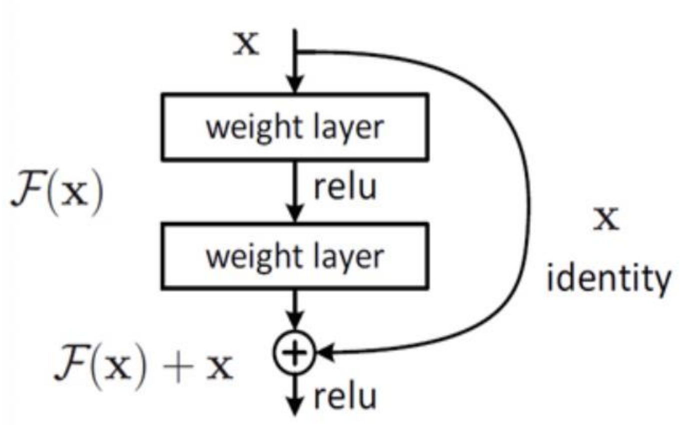


Fig 2

- The output of the module F(x)+x is the sum of the input x and the output of the convolutional layers F(x). (Fig 2)

Data Augmentation

- 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 operation.

k-folds cross validation

- Splits the original training dataset into k subsets or folds.
- 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.

EXPERIMENTS & DISCUSSION

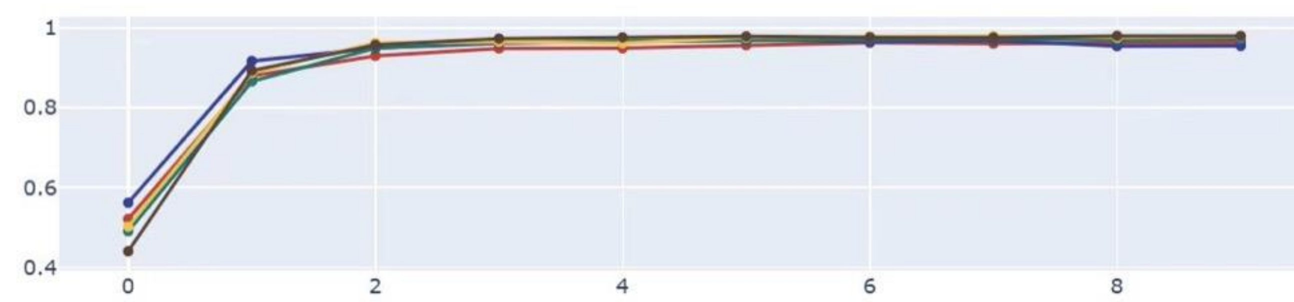
- Throughout our experiments, we choose the default resnet18 in package torchvision with pre-trained parameters as our model backbone and finetune the model on the given dataset.
- For hyperparameters and training configurations, we use cross entropy as the model loss and choose Adam as the optimizer 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.

	fold	epoch	train_loss	valid_loss	valid_score
0	0	0	1.438875	1.404425	0.484913
1	0	1	1.056218	0.730558	0.882492
2	0	2	0.564259	0.378210	0.928703
3	0	3	0.416162	0.318794	0.943165
4	0	4	0.371767	0.259038	0.956969
5	0	5	0.326739	0.244183	0.958639

Table 1

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

Fig1



- Table 1 presents the results of the 5 folds when using stratified k-fold cross validation.
- 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%.
- Stratified cross validation surpasses the standard one as 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.

CONCLUSION

- Develop deep learning models based on computer vision technology to detect plant diseases from photos.
- Use data augmentation and ResNet as the model backbone with cross-entropy as the training loss.
- Perform 5-fold cross-validation and the results show that the methods used are effective, achieving a result of 96.97%.
- Demonstrate the potential of using AI and machine learning techniques to improve plant pathology detection.
- Highlight the importance of addressing data insufficiency and imbalance to improve model robustness.

ACKNOWLEDGEMENTS

[1] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.

[2] Buslaev A, Iglovikov V I, Khvedchenya E, et al. Albumentations: fast and flexible image augmentations[J]. Information, 2020, 11(2): 125.

[3] Arlot S, Celisse A. A survey of cross-validation procedures for model selection[J]. 2010.

[4] Thapa R, Zhang K, Snavely N, et al. The Plant Pathology Challenge 2020 data set to classify foliar disease of apples[J]. Applications in plant sciences, 2020, 8(9): e11390.