

Plant Pathology 2020 - FGVC7

Identify the category of foliar diseases in apple trees

Under the guidance of : Prof. Ramkumar Hariharan

Xiaoli Zhou • Heyao Liu • Ming Wang • Zhan Wang

INFO 6105, Northeastern University, 2021



Introduction :

Identification of plant disease is usually done through visual inspection.

In this **kaggle** competition, participants were asked to train a machine learning model and diagnose plant diseases based on the image data set.

The categories include "healthy", "scab", "rust", and "multiple diseases".

Healthy



Scab



Rust



Multiple diseases



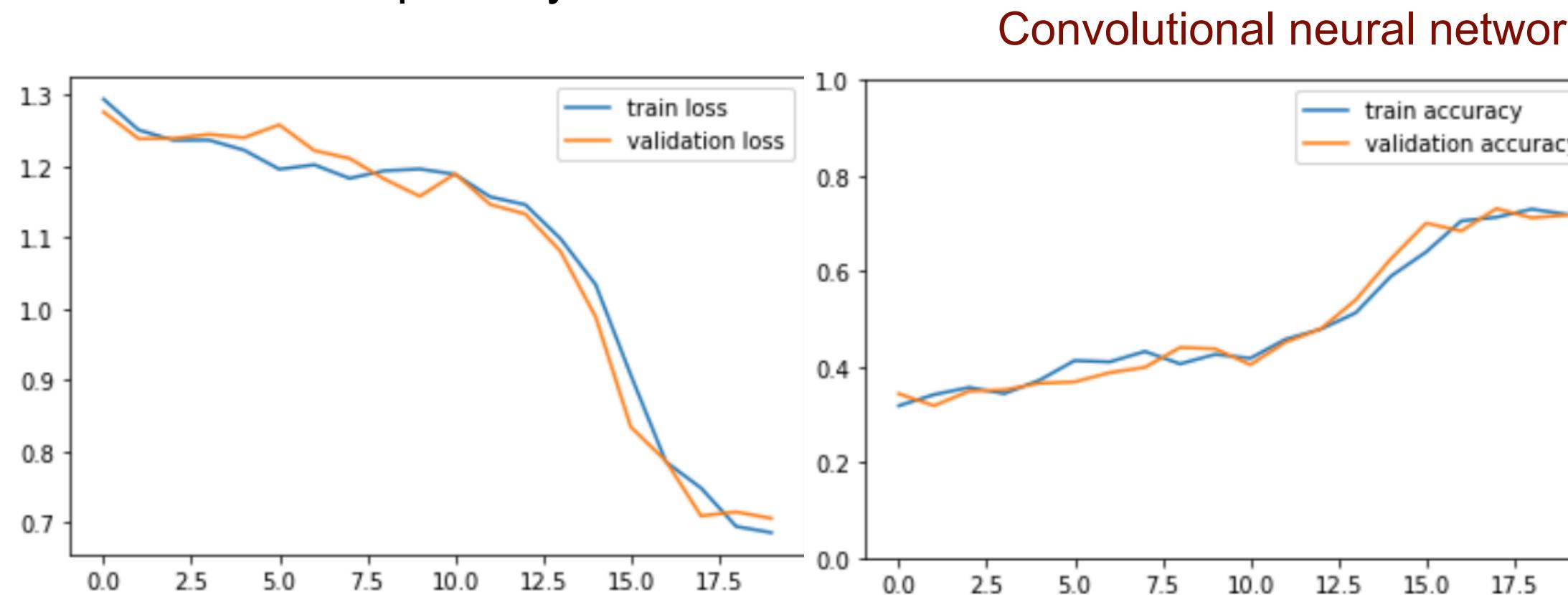
Data Augmentation :

Images are rescaled、rotated、flipped and changed brightness by ImageDataGenerator.
Apply our image augmentation to training and validation datasets.

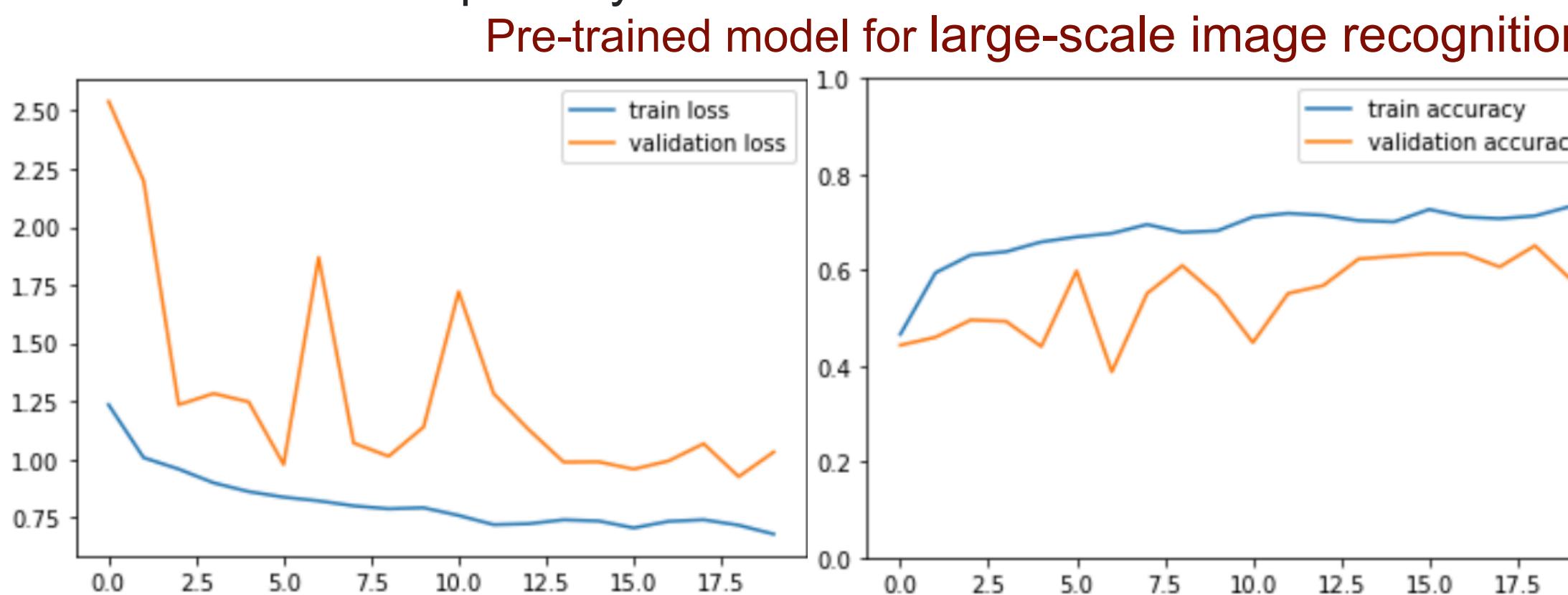


Image Classification Methods :

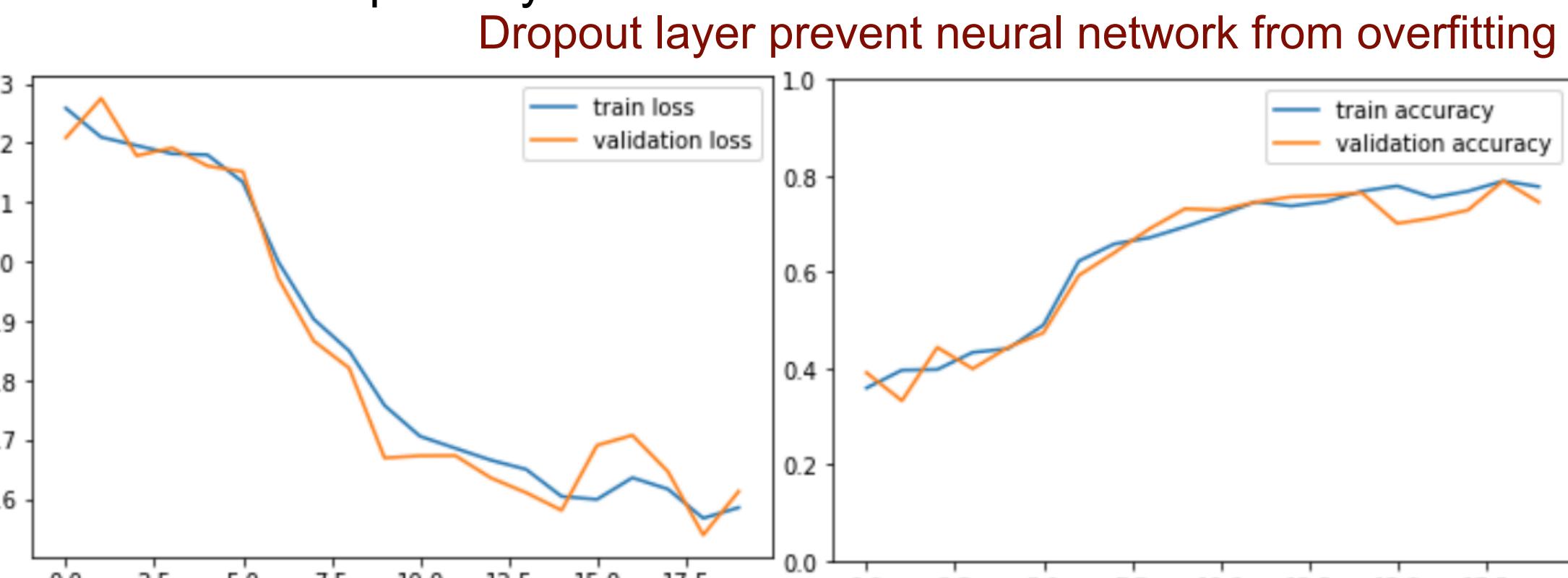
CNN w/o dropout layer :



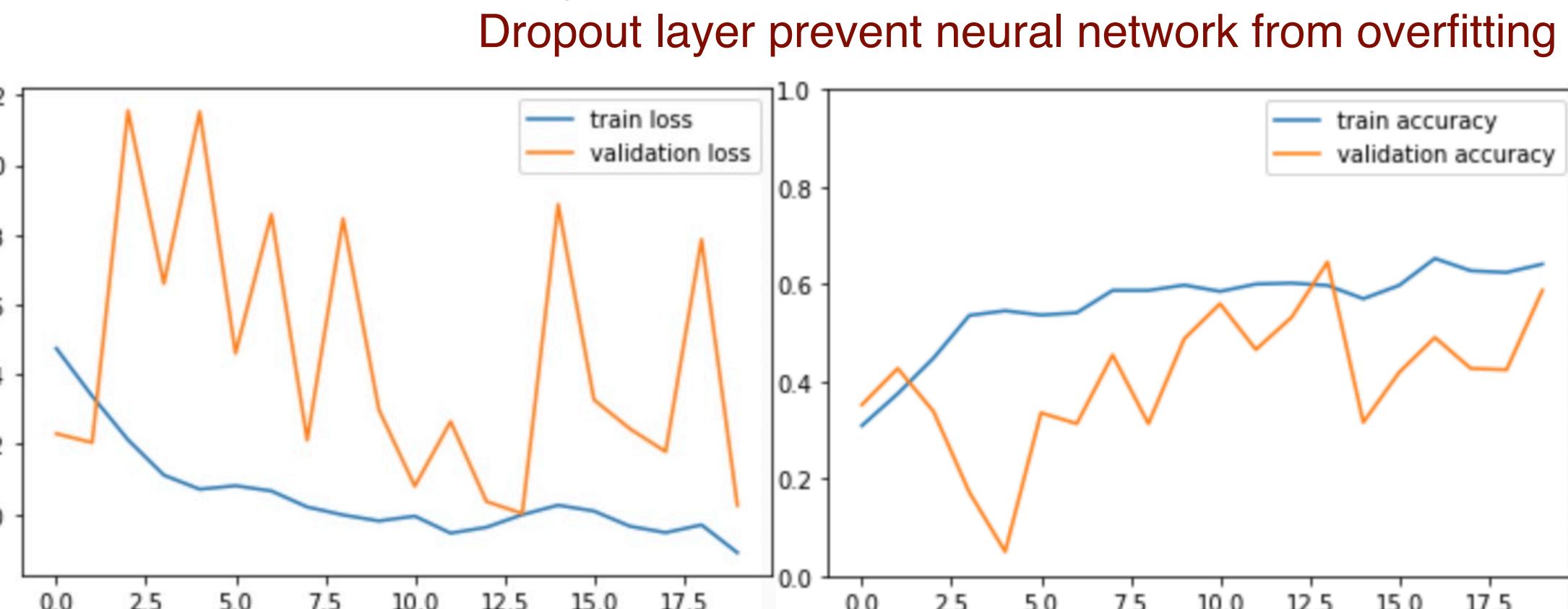
VGG16 w/o dropout layer:



CNN with dropout layer :



VGG16 with dropout layer :



Conclusion: CNN model with dropout layer has the best accuracy, which is about 79%.



Hyper-Parameter Tuning

CNN model hyper-parameters tuning with Keras Tuner and Tensorflow 2.0.
Change activation method, the number of filters and dense units.

```
Results summary
Results in hyperband/test
Showing 10 best trials
Objective(name='val_accuracy', direction='max')
Trial summary
Hyperparameters:
num_filters: 64
units: 192
dense_activation: sigmoid
tuner/epochs: 20
tuner/initial_epoch: 7
tuner/bracket: 2
tuner/round: 2
tuner/trial_id: fc86239a12bf79b9089d2effa6f2792e
Score: 0.7340720295906067
```



Test Predictions

Using the best model to generate predictions for the hidden test set. Submit file to Kaggle.

	image_id	healthy	multiple_diseases	rust	scab
0	Test_0	0.096149	8.757750e-03	0.887817	0.007277
1	Test_1	0.107260	7.343045e-03	0.880118	0.005279
2	Test_2	0.036778	3.412156e-01	0.008169	0.613837
3	Test_3	0.006763	8.441761e-01	0.004795	0.144266
4	Test_4	0.037938	2.949687e-07	0.962060	0.000001



References :

Plot training/validation set accuracy and loss curves:

<https://www.kaggle.com/michaelcripman/90-accuracy-vgg16-brain-tumor-diagnosis>

Keras Tuner for hyper-parameter tuning:

<https://www.sicara.ai/blog/en/hyperparameter-tuning-keras-tuner>

Make desire directories for raw image data:

<https://docs.python.org/3/library/os.html>

<https://www.geeksforgeeks.org/python-os-mkdir-method/>

Copy raw image data into desire directory :

<https://thispointer.com/python-how-to-copy-files-from-one-location-to-another-using-shutil-copy/>

<https://www.tutorialspoint.com/How-to-copy-files-from-one-folder-to-another-using-Python>

Preprocessing and modification of image data generation:

<https://towardsdatascience.com/data-preprocessing-and-network-building-in-cnn-15624ef3a28b>

