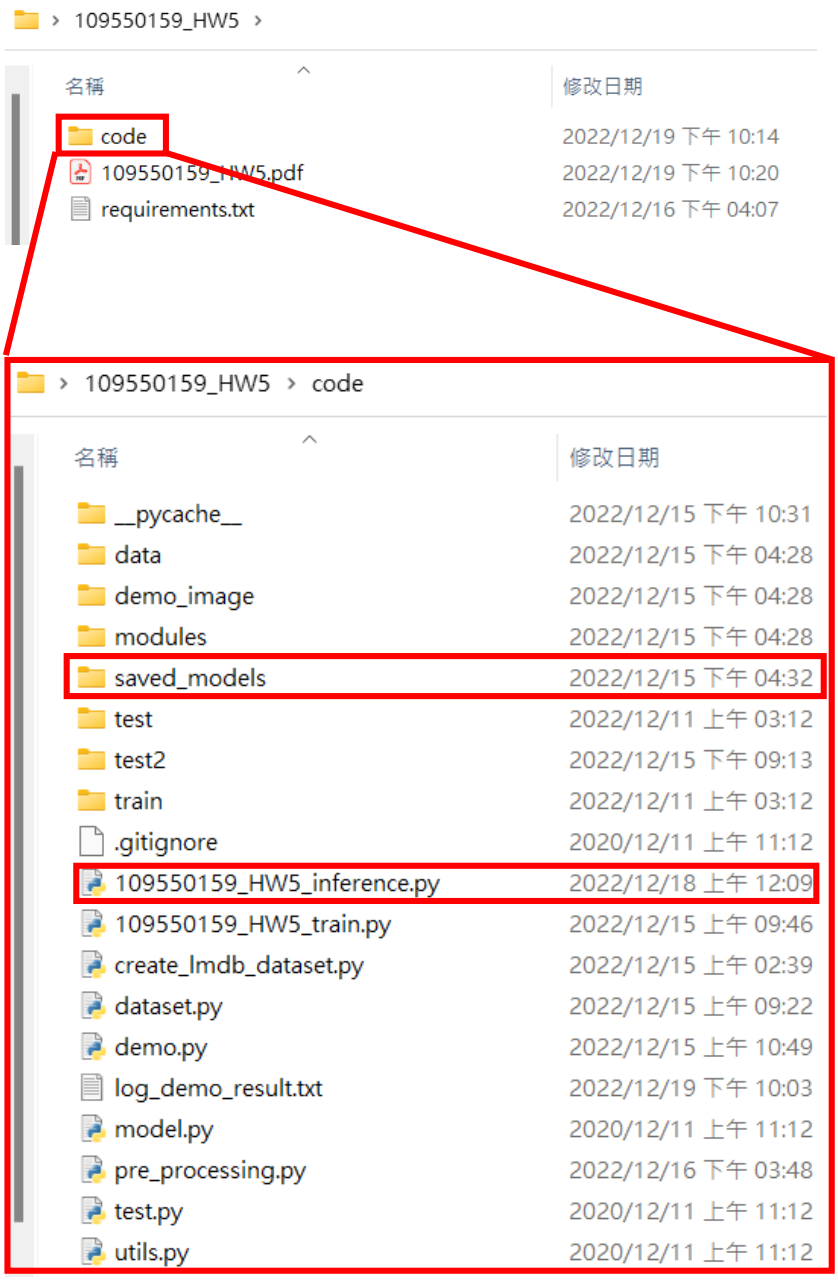
Introduction to Machine Learning HW5 Report

109550159 李驊恩

- Environment details
  - Download my model: https://drive.google.com/file/d/1l-4LrpZIrVku7PlgISTb08dsWJANha7a/view?usp=share_link

    Unzip saved_models.zip and **put the "saved_models" folder into the "code" folder** and **run 109550159_HW5_inference.py** to generate csv file.



Download on google drive and unzip the folder.

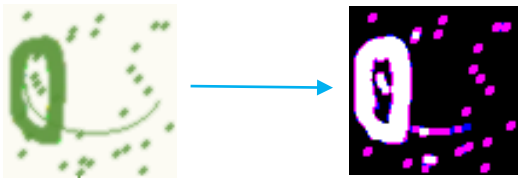
Run 109550159_HW5_inference.py to generate submission.csv.

Submission.csv will also be in the "code" folder.

- Python Version: 3.9.7
- Some important package must be installed : lmdb, pillow, torchvision, nltk, natsort (See more details in **requirements.txt**)

- Implementation details
  - Model Architecture

    There are three stages in the model architecture. First use pre_processing.py to process the image.

    After processing, use transformation.py to normalize the image to prevent it from tilting or being curved.

    Second, we can apply the neural network by feature_extraction.py to map the features of the image into character recognitions. And some interference information such as color or background will be eliminated.

    Last, output a sequence of characters according to the features.
  - Hyperparameters

    Hyperparameters are set in train.py. Some hyperparameters of the final model I use : batch size =192, iteration=300000, learning rate=1.0
  - Used deep learning framework

    CNN, RNN, ResNet

    References : https://github.com/clovaai/deep-text-recognition-benchmark

    https://arxiv.org/pdf/1904.01906.pdf

    Submission on Kaggle: