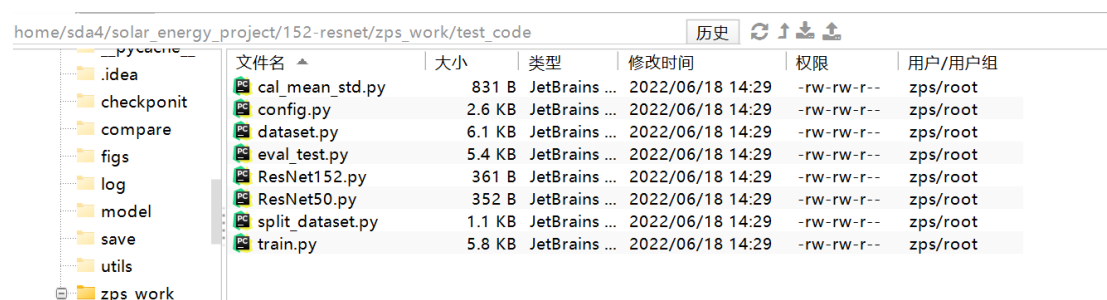


Solar project code user manual

First, move to the directory where the current code is located.



1. Calculate the mean and standard deviation of the original training set

Command:

```
python cal_mean_std.py --cal_mean_std_path /home/sda4/data/0min-300s-regular/DNI/LR/train --mask 0
```

Where: cal_mean_std_path is followed by the path of the original data training set, and mask is followed by whether the image needs to be masked (0 is not masked, 1 is masked).

```
(lab) zps@lgy-lab:/home/sda4/solar_energy_project/152-resnet/zps_work/test_code$ python cal_mean_std.py --cal_mean_std_path /home/sda4/data/0min-300s-regular/DNI/LR/train
mean:
[tensor(0.4922), tensor(0.5424), tensor(0.5440)]
std:
[tensor(0.2952), tensor(0.3117), tensor(0.3254)]
```

Figure 1. calculate mean and std values

2. Modify mean and std of dataset.py file

Enter dataset.py, modify the mean and std in the MyDataSet and MyDataSetTest classes to the values shown in Figure 1. Remember that both classes need to be modified.

```
# Standardize data sets
self.normalize = transforms.Normalize(
    mean=[0.4922, 0.5424, 0.5440],
    std=[0.2952, 0.3117, 0.3254]
)
```

Figure 2. location of dataset.py to be modified

3. Re divide the data set


A folder “data” for storing new data sets can be created in the current directory, and

three subfolders of train, val and test can be created in the “data” folder to store new training, verification and test sets.

Command:

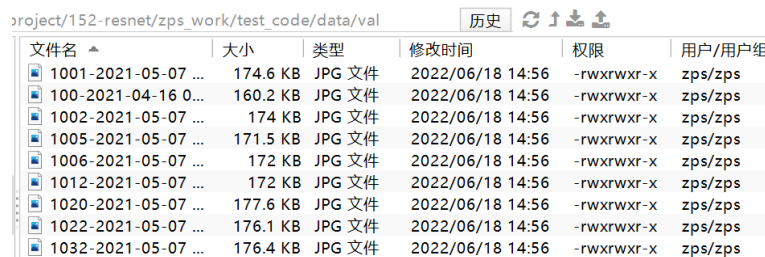
```
python split_dataset.py --source_train_path /home/sda4/data/0min-300s-regular/DNI/LR/train --source_test_path /home/sda4/data/0min-300s-regular/DNI/LR/test --target_train_path data/train --target_val_path data/val --target_test_path data/test
```

Where: source_train_path followed by the location of the original data training set, source_test_path followed by the location of the original data test set (the original data has only two folders of training and test sets), target_train_path followed by the path where the new training set is stored, target_val_path followed by the path where the new verification set is stored, target_test_path followed by the path where the new test set is stored.



```
(lab) zps@zy-lab:/home/sda4/solar_energy_project/152-resnet/zps_work/test_code$ python split_dataset.py --source_train_path /home/sda4/data/0min-300s-regular/DNI/LR/train --source_test_path /home/sda4/data/0min-300s-regular/DNI/LR/test --target_train_path data/train --target_val_path data/val --target_test_path data/test
Data set partition completed!
(lab) zps@zy-lab:/home/sda4/solar_energy_project/152-resnet/zps_work/test_code$
```

Figure 3. dividing data sets



文件名	大小	类型	修改时间	权限	用户/用户组
1001-2021-05-07 ...	174.6 KB	JPG 文件	2022/06/18 14:56	-rwxrwxr-x	zps/zps
100-2021-04-16 0...	160.2 KB	JPG 文件	2022/06/18 14:56	-rwxrwxr-x	zps/zps
1002-2021-05-07 ...	174 KB	JPG 文件	2022/06/18 14:56	-rwxrwxr-x	zps/zps
1005-2021-05-07 ...	171.5 KB	JPG 文件	2022/06/18 14:56	-rwxrwxr-x	zps/zps
1006-2021-05-07 ...	172 KB	JPG 文件	2022/06/18 14:56	-rwxrwxr-x	zps/zps
1012-2021-05-07 ...	172 KB	JPG 文件	2022/06/18 14:56	-rwxrwxr-x	zps/zps
1020-2021-05-07 ...	177.6 KB	JPG 文件	2022/06/18 14:56	-rwxrwxr-x	zps/zps
1022-2021-05-07 ...	176.1 KB	JPG 文件	2022/06/18 14:56	-rwxrwxr-x	zps/zps
1032-2021-05-07 ...	176.4 KB	JPG 文件	2022/06/18 14:56	-rwxrwxr-x	zps/zps

Figure 4. new data set

4. Model training

Command:

```
python train.py --data_type ghi --train_path data/train --val_path data/val --model_name ResNet152 --save_path result/train_info --mask 0 --num_workers 4
```

Where: data_type followed by the label to be trained (GHI, DHI, DNI), save_path followed by the path to save the model and training information, mask followed by whether the image needs to be masked (0 is not masked, 1 is masked), num_workers is followed by how many threads are used to read data. The more threads are used, the

faster the reading speed is.

```
(lab) zps@lgz-lab:/home/sda4/solar_energy_project/152-resnet/zps_work/test_code$ python train.py --data_type ghi --train_path data/train --val_path data/val --model_name ResNet152 --save_path result/train_info
epoch: 1 train loss: 0.08131609112024307 epoch_time: 70.71676325798035 MAE: 0.1996738165616989 MBE: -0.19781643152236938 MSE: 0.04576191306114197 RMSE: 0.21392033994197845 MAPE: 49.7734260559082 Corr: 0.96142828464508
06
epoch: 2 train loss: 0.025165816769003868 epoch_time: 70.9731297492981 MAE: 0.15891237556934357 MBE: -0.15769319236278534 MSE: 0.029125044122338295 RMSE: 0.1706606149673462 MAPE: 39.79888153076172 Corr: 0.974160313606
2622
```

Figure 5. model training

5. Model test (prediction)

In the model saving path (such as result/train_info), find the latest saved model (the one with the highest round), such as result/train_info/ResNet152_model_5.pt (stop training after only 7 times, for demonstration purposes).

project/152-resnet/zps_work/test_code/result/train_info 历史

文件名	大小	类型	修改时间	权限	用户/用户组
ResNet152_model_...	230.5 MB	PT 文件	2022/06/18 15:09	-rw-rw-r--	zps/zps
ResNet152_model_...	230.5 MB	PT 文件	2022/06/18 15:11	-rw-rw-r--	zps/zps
ResNet152_model_...	230.5 MB	PT 文件	2022/06/18 15:12	-rw-rw-r--	zps/zps
ResNet152_model_...	230.5 MB	PT 文件	2022/06/18 15:15	-rw-rw-r--	zps/zps
train_info.txt	1.1 KB	文本文档	2022/06/18 15:15	-rw-rw-r--	zps/zps

Figure 6. The later the model is saved, the better.

Command:

```
python eval_test.py --load_model result/train_info/ResNet152_model_5.pt --data_type ghi --test_path data/test --model_name ResNet152 --save_path result/test_info --mask 0 --num_workers 4
```

Where, load_model followed by the path of the model saved, data_type followed by the predicted label (consistent with the training), save_path followed by the path where the prediction results and indicators are saved, test_path is followed by the path of the test set, mask is followed by whether the image needs to be masked (the same as the setting during training), num_workers followed by the number of threads used to read data.

```
(lab) zps@lgz-lab:/home/sda4/solar_energy_project/152-resnet/zps_work/test_code$ python eval_test.py --load_model result/train_info/ResNet152_model_5.pt --data_type ghi --test_path data/test --model_name ResNet152 --s
ave_path result/test_info
/home/sda4/solar_energy_project/152-resnet/zps_work/test_code/eval_test.py:26: FutureWarning: The pandas.datetime class is deprecated and will be removed from pandas in a future version. Import from datetime module in
stead.
time = pd.datetime.strptime(timestamp, '%Y-%m-%d %H:%M:%S')
MAE: 64.812934075481820 MBE: 35.51176071166992 MSE: 6504.0292966875 RMSE: 80.6475601196289 MAPE: 15.762832641601562 Corr: 0.9810340749160767
```

Figure 7. model test (prediction)

y_project/152-resnet/zps_work/test_code/result/test_info 历史

文件名	大小	类型	修改时间	权限	用户/用户组
ghi_result.csv	28.7 KB	Microsof...	2022/06/18 15:26	-rw-rw-r--	zps/zps
test_info.txt	139 B	文本文档	2022/06/18 15:26	-rw-rw-r--	zps/zps

Figure 8. stored results