ML Final

Human Protein Atlas Image Classification

NTU_r06922134_台北大冒險

Data Preprocessing/Feature Engineering

- RGBY (31072, 512, 512, 4)
- Read All train data (Not Data Generator) 30+ G
- Valid set (last 10%)

```
np.sum(valid_y, axis=0)

array([1284,  136,  390,  155,  176,  243,  108,  282,    3,    5,    3,
        117,   71,   50,  117,    4,   50,   19,   77,  149,   18,  371,
         74,  306,   38,  787,   29,    0], dtype=uint32)
```

Model Description(1)

- Dense > Conv2D
- Softmax > Sigmoid
- Best pretrained model : Densenet169

```
Layer (type)                 Output Shape              Param #
=================================================================
input_1 (InputLayer)         (None, 512, 512, 4)       0
_____
batch_normalization_1 (Batch (None, 512, 512, 4)       16
_____
dense_1 (Dense)              (None, 512, 512, 4)       20
_____
dense_2 (Dense)              (None, 512, 512, 3)       15
_____
densenet169 (Model)          (None, 1664)              12642880
_____
dense_3 (Dense)              (None, 28)                46620
=================================================================
```

Model Description(2)

- Focal_loss

- Batch_size = 4

- Adam(lr=1e-4)

- Class_weight

- ReduceLROnPlateau

Model Description(3)

- Pretrained model layers (trainable=False)
- Epoch = 4
- All layers trainable
- Epoch = 40
- save_best_val_f1 > save_lowest_val_loss

Test Time Augmentation(TTA)

- tta_classification(model, h_flip=True, rotation=(90, 270),
-                                   merge='mean')

- tta_model.predict(valid_x, batch_size=1, verbose=1)

Threshold

- Valid set
- np.linspace(0, 1, 1000)
- TTA

Predict

- Threshold
- TTA
- Fill nothing if NA

Ensemble

- **Inception_resnet_v2**
- **Densenet201**
- **Densenet169**
- **Densenet121**
- **InceptionV3**
- **ResNet50**
- vote=4