

ニューラルネットワーク（U-Net）による数式検出プログラム 実行説明書

0) 準備

1. Anaconda Prompt を立ち上げる：
Windowsメニュー → Anaconda3 (64-bit) → Anaconda Prompt
2. > cd C:\Users\admin\Desktop\MigratingU-net

1) 訓練用データの準備

ページ画像とアノテーションファイルからU-net学習用の画像セットを生成する

実行フォルダ

Desktop\MigratingU-net\1_MakeTrainData

実行プログラム

GTDB_DataCreation_Unet.py

プログラムの動作

指定されたアノテーションCSVファイルを読み込み、
ファイル内で指定されているページ画像ファイルを入力する。
画像の前処理、解像度変更、指定サイズへの切り出しを行い、
学習用の原画像・教師画像ペアを出力する。

実行方法

```
python GTDB_DataCreation_Unet.py [-h|—help] [—outimg-dir=<out image dir>]  
[—window-height=<height>] [—scale=<scale>] [—stride=<stride>] [—area-csv=  
<area_type_CSV>] <GTDB_csv_file> <GTDB_imageDir> <output Data Dir>
```

引数とオプションの意味

<GTDB csv file>：アノテーションファイル 1ジャーナルあたり1ファイル

<GTDB_imageDir>：ページ画像が格納されているフォルダ 1ページあたり1ファイル、PNG形式が望ましい

<ouput Data Dir>：学習用データ出力用ディレクトリ 指定されたディレクトリ内に、
images（入力画像）、groundtruth（教師データ）が存在する必要がある。

-h or --help : ヘルプメッセージ表示
--outimg-dir : 前処理後のページ画像を保存するフォルダを指定
--window-height : 学習用画像の縦横サイズを指定
--scale : 画像解像度の拡大・縮小率
--stride : 学習用小領域を切り出すステップ幅
--area-csv : 未使用

実行コマンドの例

```
> cd 1_MakeTrainData  
> python GTDB_DataCreation_Unet.py --outimg-dir="..\data\UnetData512_150dpi\imgs" --window-height="512" --scale="0.25" --stride="256" ..\data\GTDB_eng-2\Bergweiler83.csv ..\data\GTDB_eng-2 ..\data\UnetData512_150dpi
```

2) U-net の学習

実行フォルダ

Desktop\MigratingU-net\2_training

実行プログラム

train_Unet.py

プログラムの動作

指定されたディレクトリないの画像を訓練用データセットとして、
画像変換U-netを学習する。

実行方法

python train_Unet.py [dataDir] [window_size] [layer num] [output file]

引数とオプションの意味

dataDir : 1) で作成した学習用データの格納されているディレクトリ

Window_size : 学習する画像のサイズ

Layer_num : U-net の深さ 画像サイズ = $2^{\text{Layer_num}}$ が最大

Output file : 学習したU-netの重みを保存するファイル (.hdf5 の拡張子をつけてくだ

さい)

実行コマンドの例

```
> cd ..\2_training
> python train_Unet.py ..\data\UnetData512_150dpi 512 9 ..\model\unet.hdf5
* このコマンド例では、学習済みのニューラルネットワークは ..\model\unet.hdf5 に
  保存されます。
```

3) 数式検出 (画像変換)

学習したモデルを用いて数式検出用の画像変換を行う。

実行フォルダ

Desktop\MigratingU-net\3_predicting

実行プログラム

MathExtraction_Unet.py

プログラムの動作

指定された画像ファイルに対して、学習済みU-netを用いて画像変換を行う。

実行方法

```
python MathExtraction_Unet.py [input image] [output image] [window size]
[layer num] [U-net weight]
```

引数とオプションの意味

Input image: 入力画像ファイル

Output image: 変換後の画像ファイル

Window size: 切り出す小領域サイズ、 2) における window_size と同じ値を指定する

Layer num: U-net の深さ 2) における layer_num と同じ値を指定する

U-net weight: 2) で保存したU-netの重みファイル

実行コマンドの例

```
> cd ..\3_predicting
> python MathExtraction_Unet.py ..\data\GTDB-
```

1_MathText_Segmentation\ActaM_1970_038.png

..\data\Results\ActaM_1970_038.png 512 9 ..\model\unet.hdf5

＊ このコマンド例では、検出結果画像は ..\data\Results\ActaM_1970_038.png に保存されます。