PyHistopathology

Release 1.0

Feb 10, 2020

Contents

1 Features **3**

2 Table of Contents **5**

2.1 Reading . . . 5
2.2 Usage . . . 5
2.3 Arguments . . . 5
2.4 Return Type . . . 6
2.5 Example . . . 6
2.6 Denoising . . . 6
2.7 Usage . . . 6
2.8 Arguments . . . 6
2.9 Return Type . . . 7
2.10 Extracting . . . 7
2.11 Usage . . . 7
2.12 Arguments . . . 7
2.13 Return Type . . . 8

PyHistopathology is a Python API for pre-processing Whole-Slide Images for use in a machine learning algorithm.

CHAPTER 1

Features

Reading Whole-Slide Images with WSI_Scanning.readWSI()

- Input: WSI path or directory
- Output: Numpy array of WSI with data type int32

Denoising Whole-Slide Images with Denoising.denoising()

- Input: WSI Path or directory
- Output: Numpy array of WSI Image (After denoising) with dtype int32

Patch Extraction for Whole-Slide Images with Extractingpatches.extractingPatches()

- Input: WSI Path or directory
- Output: Fills up the outputpath with patches extracted from the WSI

CHAPTER 2

Table of Contents

2.1 Reading

Use WSI_Scanning.readWSI() to read Whole-Slide Images

2.2 Usage

readWSI(WSI_path, magnification_level, annotation_file, annonated_level)

2.3 Arguments

2.3.1 WSI_path

Path or location of WSI.

2.3.2 magnification_level

Level of zoom, for example 40, 20, 10, or 5. Default magnification level is 20. - Note: if magnification 40x for max zoom level of 20x image an error will be raised.

2.3.3 annotation_file

Default annotation = None. If annotation are available in a xml file, set annotation_file to be the xml file path.

2.3.4 annonated_level

If annotation is given then set annotated_level equal to the z-axis of the annotations. Default annotatedlevel is 0.

2.4 Return Type

Numpy array of WSI Image (After denoising) with dtype int32

2.5 Example

```
from WSI_Preprocessing.Preprocessing import WSI_Scanning
import cv2
img,slide_dim = WSI_Scanning.readWSI("example.svs")
cv2.imwrite("example.png",img)
```

2.6 Denoising

Use Denoising.denoising() to remove stains, folds and other background noise in Whole-Slide Images

2.7 Usage

denoising(inputsvs, magnification, filtering, patch_size, upperlimit, lowerlimit, red_value, green_value, blue_value)

2.8 Arguments

2.8.1 inputsvs

Path or location of WSI.

2.8.2 magnification

Level of zoom, for example 40, 20, 10, or 5. Default magnification level is 20. - Note: if magnification 40x for max zoom level of 20x image an error will be raised.

2.8.3 filtering

GuassianBlur, RGBThersholding, or None

GuassianBlur: Homogeneity calculations based on image smoothing and Gaussian blur equations. We compute sum of square differences between two consecutive Gaussian blurred images as score for homogeneity.

- Upper limit: Upper threshold of homogeneity score. Default value is 9500 with kernel size of 1111
- Lower limit: lower threshold of homogeneity score. default value is 1500 with kernel size of 1111

- Patch size: Not significant parameters for GuassianBlur filtering

RGBThersholding: Validated patches based on RGB values of patches

- red_value: Red threshold
- green-value: Green threshold
- blue_value: Blue Threshold

None: Only removes Background

Note that our default is GuassianBlur technique. GuassianBlur is highly effective and requires more computational power (RAM). RGBThersholding is less effective which needs less computational power

2.9 Return Type

Numpy array of WSI Image (After denoising) with dtype int32

2.10 Extracting

Use Extractingpatches.extractingPatches() to extract patches from Whole-Slide Images

2.11 Usage

extractingPatches(inputsvs, outputpath, magnification, patch_extraction_creatia, number_of_patches, filtering, patch_size, upperlimit, lowerlimit, red_value, green_value, blue_value, Annotation, Annotationlevel, Requiredlevel, reconstructionimagepath)

2.12 Arguments

2.12.1 inputsvs

Path or location of WSI.

2.12.2 magnification

Level of zoom, for example 40, 20, 10, or 5. Default magnification level is 20. - Note: if magnification 40x for max zoom level of 20x image an error will be raised.

2.12.3 filtering

GuassianBlur, RGBThersholding, or None

GuassianBlur: Homogeneity calculations based on image smoothing and Gaussian blur equations. We compute sum of square differences between two consecutive Gaussian blurred images as score for homogeneity.

- Upper limit: Upper threshold of homogeneity score. Default value is 9500 with kernel size of 1111
- Lower limit: lower threshold of homogeneity score. default value is 1500 with kernel size of 1111

- Patch size: Not significant parameters for GuassianBlur filtering

RGBThersholding: Validated patches based on RGB values of patches

- red_value: Red threshold
- green-value: Green threshold
- blue_value: Blue Threshold

None: Only removes Background

Note that our default is GuassianBlur technique. GuassianBlur is highly effective and requires more computational power (RAM). RGBThersholding is less effective which needs less computational power

2.12.4 patch_extraction_creatia

random, or None. Default is None. For extracting a fixed number of patches for WSI we can use random.

2.12.5 number_of_patches

Default number of patches is 2000

2.12.6 outputpath

Folder to store the extracted patches

2.12.7 reconstructionimagepath

If you want to compare the patches with WSI we can mention the reconstructionimagepath. Default is None. Note: it only works with patch_extraction_creatia = None.

2.13 Return Type

None, fills up output path with images directly instead of returning a Numpy array.