tissueloc Documentation

Release 2.0.1

Pingjun Chen

Feb 11, 2020

User Documentation

tissueloc is used for tissue localization in whole slide pathology image.

CHAPTER 1

Load Slide

1.1 select_slide_level

def select_slide_level(slide_path, max_size=2048): “”“Find the slide level to perform tissue localization

Parameters

slide_path [valid slide path] The slide to process.

max_size [int] Max height and width for the size of slide with selected level

Returns

level [int] Selected level.

d_factor: int Downsampling factor of selected level compared to level 0

Notes: The slide should have hierarchical storage structure.

“”“

1.2 load_slide_img

def load_slide_img(slide_path, level=0): “”“Load slide image with specific level

Parameters

slide_path [valid slide path] The slide to load.

level [int] Slide level to load.

Returns

slide_img [np.array] Numpy matrix with RGB three channels.

Notes:

Whole slide image can have more than 100,000 pixels in width or height, small level can be very big image.

“”“

CHAPTER 2

Locate Tissue

2.1 locate_tissue_cnts

def locate_tissue_cnts(slide_path,

max_img_size=2048, smooth_sigma=13, thresh_val = 0.80, min_tissue_size=10000):

“”” Locate tissue contours of whole slide image

Parameters

slide_path [valid slide path] The slide to locate the tissue.

max_img_size: int Max height and width for the size of slide with selected level.

smooth_sigma: int Gaussian smoothing sigma.

thresh_val: float Thresholding value.

min_tissue_size: int Minimum tissue area.

Returns

cnts: list List of all contours coordinates of tissues.

d_factor: int Downsampling factor of selected level compared to level 0

“””

2.2 rgb2gray

def rgb2gray(img): “””Convert RGB image to gray space.

Parameters

img [np.array] RGB image with 3 channels.

Returns

gray: np.array Gray image

“”“

2.3 thresh_slide

def thresh_slide(gray, thresh_val, sigma=13): “”” Threshold gray image to binary image

Parameters

gray [np.array] 2D gray image.

thresh_val: float Thresholding value.

smooth_sigma: int Gaussian smoothing sigma.

Returns

bw_img: np.array Binary image

“”“

2.4 fill_tissue_holes

def fill_tissue_holes(bw_img): “”” Filling holes in tissue image

Parameters

bw_img [np.array] 2D binary image.

Returns

bw_fill: np.array Binary image with no holes

“”“

2.5 remove_small_tissue

def remove_small_tissue(bw_img, min_size=10000): “”” Remove small holes in tissue image

Parameters

bw_img [np.array] 2D binary image.

min_size: int Minimum tissue area.

Returns

bw_remove: np.array Binary image with small tissue regions removed

“”“

2.6 find_tissue_cnts

def find_tissue_cnts(bw_img): “”” Fint contours of tissues

Parameters

bw_img [np.array] 2D binary image.

Returns

cnts: list List of all contours coordinates of tissues.

“”“

CHAPTER 3

About tissueloc

tissueloc is a python-based package used for whole slide pathology image tissue localization.

There are two main functions in this package:

1. Load the whole slide image from low level, which is set according to the maximum width or height that the loaded image could own.
2. Locate tissue regions through a series of image processing procedures.

This whole slide tissue localization is entirely based on basic image processing algorithms, which is extremely fast and could act as a preprocessing step for whole slide image automatic analysis. The parameters used in the processing may need to change according to specific application.