ConfigFile explanation (1 dataset = 1 Config file)

Overview : 1 dataset = 1 Config file

This page will help you to understand and set all the values of the config file
`DeepCellMap/code/config/dataset_new_fluorescence.py` or
`DeepCellMap/code/config/dataset_new_ihc.py` depending on the data on hand.

> Pour moi :
> Je vais voir les attributs les uns apres les autres
> Je les map dans le following table et je copie cole le “how to choose” en commentaire avant l’attribut du config file

- [ ] To add to the other config files
  - [ ] mapping_cells_colors (deja dans ihc)
  - [ ] limit_row_col_for_slide (deja dans ihc)

Common attributes

Attribute name	How to choose	IHC	Fluorescence
dataset_name	Use a descriptive name for your dataset.	✅	✅
mapping_img_name	Specify the names of the images in the dataset.	✅	✅
mapping_img_gender	Specify the genders of the individuals in the dataset.	✅	✅
mapping_img_disease	Specify the disease of the individuals in the dataset.	✅	✅
mapping_img_age	Specify the ages of the individuals in the dataset.	✅	✅
cell_class_names	List of cell/object names	✅	✅
mapping_cells_colors	Color associated to each cell	✅	✅
debug_mode	Set to `True` if you want to enable debugging mode.	✅	✅
data_type	Specify the type of data, such as whole slide images (WSI) or fluorescence	✅	✅
consider_image_with_channels	A fluorescent image can be considered sometimes as RGB (eg Cancer data) then,	✅	✅
data_format	Specify the file format of the data, such as TIFF.	✅	✅
conversion_px_micro_meter	Specify the conversion factor from pixels to micrometers. Ex : 1 px = 0.45 micro meter	✅	✅
channel_names	Specify the names of the channels in the data.	❌	✅
dim_position	In tensor of the image, position of C, X, Y and Z	❌	✅
has_Z	If there is a Z canal	❌	✅

tile_width	Set the width of each image tile.	✅	✅
tile_height	Set the height of each image tile.	✅	✅
roi_border_size	Set the size of the border around the region of interest (ROI).	✅	✅
border_size_during_segmentation	Set the size of the border during segmentation.	✅	✅
crop_size	Set the size of the crop for each image tile.	✅	✅
scale_factor	Downscaling factor	✅	✅
threshold_tissue	Set the threshold for tissue segmentation. If a tile contains at least (1-threshold_tissue)% of tissue the tile is considered to have tissue inside. (Save computation time)	✅	✅
preprocessing_config	Configure the preprocessing parameters for the data.	✅	✅
tissue_segmentation_param	Set the parameters for tissue segmentation.	✅	✅
tissue_extraction_accept_holes	If the segmented tissue can contains holes	✅	✅
save_tissue_segmentation_steps	Specify whether to save the intermediate steps of tissue segmentation.	✅	✅
channel_used_to_segment_tissue	Which channel is used to segment the tissue	❌	✅
use_imgs_as_channels	Some images can be split in different images (channel 1 is 1 image, channel 2 is one image and so on). Specify whether to use images as channels. So that image recombination is done. Use for Cancer dataset.	✅	✅
channels_cells_to_segment	List of the channels to segment during the process of cells detection (ex [1,2])	❌	✅
channels_of_interest	List channels of interest during the application of DeepCellMap	❌	✅
cells_from_multiple_channels	Only fluorescence : usually 1 channel = 1 celltype but cells can also result fromthe combination of several channels. Ex : dict({ "iba1_cd68" : [1,2] }) → new celltype iba1_cd68 emerges from the combination of channel 1 and 2. Cells that are double positive are considered iba1_cd68	❌	✅
association_cell_name_channel_number	dict specifying the association between cell names and channel number. Cells created from several channels are associated the channel number max_channel + key_position in cells_from_multiple_channels	❌	✅
param_best_cellpose	Set the parameters for the best Cellpose model (to detect ovoid nuclei)	✅	✅
cellpose_parameters	Dict that contains main Cellpose parameter like `downscale_factor_cellpose_tissue_segmentation`, `channel_nuclei` and `tile_subdivision_factor` used during the creation of the density heatmap	✅	✅
model_segmentation_name	Set the name of the segmentation model.	✅	✅
cell_segmentation_param	Set the parameters for cell segmentation.	✅	
cell_segmentation_param_by_channel	Set the parameters for cell segmentation but by channels (and can be image-specific) first key is		✅

	default parameters and other keys are the parameters for the other images		
classification_param	Set the parameters for cell classification.	✅	❌
cell_class_names_for_classification	Name of all the categories used during classification (include Background, and Detected)	✅	❌
tile_test_segmentation	Set the parameters for tile-level segmentation.	✅	✅
roi_test_tissue_border	Set the parameters for testing tissue border within ROIs.	✅	✅
roi_cool	Set the parameters for cool ROIs.	✅	✅
statistics_with_proba	Specify whether to compute statistics with probabilities. Used when DL did the classification and gave each cells a probability to belong to each celltype	✅	❌
colnames_table_cells_base	Specify the column names for the table of cells (created during detection segmentation and classification)	✅	✅
colnames_df_image	Specify the column names for the image dataframe.	✅	✅
colnames_df_roi	Specify the column names for the ROI dataframe.	✅	✅
path_cells	Set the path to the directory where cell images are stored.	✅	✅
physiological_regions_max_square_size	Set the maximum square size for each physiological region.	✅	❌
physiological_regions_group_for_comparaison	Set the group for comparison for each physiological region.	✅	❌
cell_cell_colocalisation_config	Set the parameters for cell-cell colocalisation analysis.	✅	✅
dbscan_based_analysis_config	Set the parameters for DBSCAN-based analysis.	✅	✅
neighbors_analysis_config	Set the parameters for neighbors-based analysis.	✅	✅
limit_row_col_for_slide	Limite row and col (on some slides the tissue is repeated)	✅	✅