Purpose

This SOP guides IMC users in the process of obtaining single cell data from raw IMC files. It provides users with cell mask(s) and .tiff files that can be used in histoCAT, where both single cell and neighborhood analysis can be performed, or extracts the single cell target expression and position information for further R analysis.

This workflow is free and open source and provided by the Bodenmiller lab, please read their description at: https://github.com/BodenmillerGroup/ImcSegmentationPipeline

At the bottom of the above webpage it is stated "This pipeline was presented at the 2019 Imaging Mass Cytometry User Group Meeting. The slides can be downloaded here."

Click on the link, download the slides and towards the end of the slides, there is the same pipeline with images that can further aid with the following steps.

Reference:

Vito RT Zanotelli, Bernd Bodenmiller, ImcSegmentationPipeline: A pixel classification based multiplexed image segmentation pipeline
Sept 2017, 10.5281/zenodo.3841961, Zenodo, https://doi.org/10.5281/zenodo.3841961

Prerequisites

I. Software requirements

Please download the following software:

1. CellProfiler 3 version 3.1.8: https://CellProfiler.org/previous_releases/
   Note: this is the only version that works
2. Ilastik 1.3.3 post1: https://www.ilastik.org/download.html
3. Anaconda (Python v3.7, Graphical Installer): https://www.anaconda.com/distribution/
4. histoCAT: Software requirements for histoCAT are available on: https://bodenmillergroup.github.io/histoCAT/ and there is also a link to the user manual and the github page there

II. Data requirements

1. IMC data:

a zipped folder containing:

Address: INO-F608A, Inselspital, CH-3010 Bern
Contact: info.imc@dbmr.unibe.ch
Tel.: +41 31 632 88 95 or +41 31 632 22 94
Website: www.imc.unibe.ch

- MCD file(s)
- associated TXT files

2. Antibody panel metadata:

Please create an antibody panel CSV file containing the following column information:

- "Metal tag": metal isotope, format: In113
- "Full": boolean (0,1): which channels to use for single cell measurements
- "ilastik": boolean (0,1): which channels to use for ilastik training

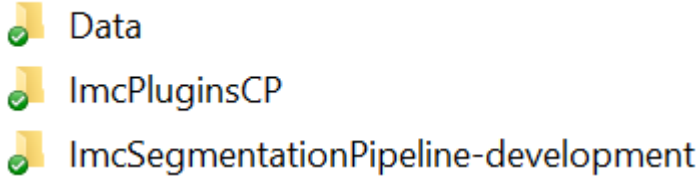
III. Download IMC Segmentation Pipeline supporting files and scripts

- Navigate to https://github.com/BodenmillerGroup/ImcSegmentationPipeline

1. Download the zip file (click the green 'Code' button and scroll down to download ZIP) ImcSegmentationPipeline-development and put the file in your Project folder as shown below in 'Data organization.' Extract the ZIP file contents and then delete the ZIP file, leaving just the extracted folder in your Project folder.
2. CellProfiler Plugins: These are modules that are added to CellProfiler to facilitate handling, processing as well as measurement of multiplexed data.
   - Navigate to https://github.com/BodenmillerGroup/ImcPluginsCP.
   - Download the plugins, put the file in your Project folder, extract the ZIP file contents and delete the ZIP file as before.
   - Load them into CellProfiler 3.1.8 like so: File > Preferences > CellProfiler plugins directory: Browse (navigate to your new plugins folder at ImcPluginsCP/plugins folder) and select OK and restart CellProfiler.
3. Load the IMC file conversion tool conda_imctools.yml into Anaconda Navigator:
   - Open Anaconda, click on Environments and then Import > Browse to ImcSegmentationPipeline-development/ setup folder and select the file conda_imctools.yml (this will take a couple minutes to load).

IV. Data organization

We suggest making a Project folder that contains:



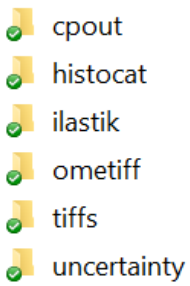
Within the folder Data, put the IMC data zip folder (MCD file and its associated TXT files) and antibody panel metadata (CSV file).

Begin IMC segmentation pipeline

Launch a Jupyter notebook:

1. In Anaconda select Environment and activate imctools.
2. Launch a Jupyter notebook by returning to the home screen.
   Note: If it does not work to return to Anaconda home screen, try launching the Jupyter notebook in Enviroment by pressing the play icon and then click on the option to open with Jupyter notebook.
3. Choose your favorite browser if prompted.
4. In the Jupyter notebook, navigate to ImcSegmentationPipeline-development/ scripts and select imc_preprocessing.ipynb.
5. Run the first two cells of script by clicking the 'Run' button two times.
6. Edit the 3rd cell of script where the font is red to identify: 1) where your input data folder is (this the path to your Data folder), 2) to call your IMC data zip file (insert zip file name here), 3) where your output folder should go (this is the path to your Data folder) 4) where your antibody panel CSV file is located (this is the path to your Data folder including the name of the CSV file at the end).
   Note: PC users will need to make their back slashes (\ ) into forward slashes (/) for this to work.
7. Run the cells up to the section titled "Next Steps" and skip the 'Optional Step' if you are using your own data.

Note: This script created subfolders within the Data folder as follows: cpout, histocat, ilastik, ometiff, tiffs, uncertainty as shown below. It filled these folders with the files needed for the next steps of the pipeline.



Prepare images for Ilastik:

1. Open CellProfiler and go to File > Import > Pipeline from File > Select: ImcSegmentationPipeline-development/ cp3_pipelines and select the file 1_prepare_ilastik.cppipe.
2. Click on View output settings and under Default input Folder, navigate to your input Data folder. Under Default Output Folder browse to your tiffs folder inside your Data folder.
3. Click on Images, then drag and drop all the files in the tiffs folder into CellProfiler.
4. Click on Analyze Images.

Note: This step prepares a series of images for ilastik training as well as random crops and puts them in your tiffs folder.

Ilastik pixel classification training:

To create a new classifier, follow the steps outlined in the Jupyter notebook or read the version provided below:

1. Open Ilastik and Create a new project for pixel classification and save the project file in the ilastik subfolder.
2. Add the .h5 random crops: Raw data > Add Separate Images > Select all .h5 crop images in the tiffs subfolder (the .h5 crops are the .h5 files with coordinates in the image name).
3. Proceed to Feature Selection.
4. Select suitable features (check everything)
5. Proceed to Training: add 3 labels:
    1. Nucleus
    2. Cytoplasm/membrane
    3. Background

   For large datasets adding the labels can take a while

   Begin drawing on the crops to label the cell compartments:
    - In the large box under 'Group Visibility' scroll down to see the last section titled 'Raw Input.' Here each channel corresponds to the targets identified for use in cell segmentation as indicated in the 'ilastik' column of the Antibody Panel metadata CSV file you placed in the Data folder. The 0 channel correspond to the sum of all channels and the remaining channels are in the order of the ilastik channels in the CSV file. This is useful for labeling the background.

- Use window leveling to change the contrast by clicking on the shaded box symbol under the section '+ Add Label.' Adjust the contrast by left clicking on image and moving mouse to the right or left. You can reset the settings by right clicking.
- Navigate Ilastik: Zoom in or out using your fingers and pinching with two fingers on the track pad, or with a mouse Ctrl+ mouse wheel scrolling (or cmd+ mouse wheel scrolling on a Mac). You can zoom in or out by using the keyboard keys + and – respectively. You can also use the arrows on the top right for zooming to fit or resetting the zoom. Panning can be achieved by holding the mouse wheel pressed + move, or dragging the mouse while pressing the shift key.
- Start training: Activate channel 0, select Background in the Training box, adjust the contrast to see the image well, and then draw on the image in the dark sections that identify the background signal. Activate the next channel that identifies Cytoplasm/membrane, adjust contrast, select Cytoplasm/membrane in Training box, zoom in then image and draw on the areas with a clear signal. Do the same for the Nucleus identification. If you see in the Nucleus channel that two Nuclei are stuck together but have a faint dip in intensity in between, label this as 2: Cytoplasm and encircle Nucleus with Cytoplasm.
- Click on 'Live Update' to see how well training is doing (this can be slow on a personal computer, possibly use a virtual machine to speed this up). You will see the current prediction in color for the different cell compartments, make sure the box next to Probability is checked for this.
- Click on the eye on the left side of Uncertainty in the 'Group Visibility' box. This shows you where the uncertain areas are in bright blue and then you can continue training on these areas.
    - Recommendations: don't train too much on one image since this skews the classifier to only work on that image, train a bit on multiple images.
    - Disable `Live Update` after use it so it's not continually updating.
    - Frequently check the `Uncertainties`: This indicates which pixels the classifier profits most if they are labeled. A well-trained classifier has low uncertainty within class regions (e.g. Nucleus) and high uncertainty at class borders (e.g. between nuclei and cytoplasm).
    - The alpha ranges help you visualize the predictions. You can adjust a cell compartment visibility to see the prediction better if needed.

Address: INO-F608A, Inselspital, CH-3010 Bern
Contact: info.imc@dbmr.unibe.ch
Tel.: +41 31 632 88 95 or +41 31 632 22 94
Website: www.imc.unibe.ch

Export Uncertainty maps: and check the level of uncertainty (the white coloring) if have many images to check.

Export probability maps: This is done once the classifier is well trained, one can save the pretrained classifier and use it on similar tissues with the same panel.

1. Under the section "Prediction Export", under "Export Settings", select the source- Probabilities and then click on "Choose Export Image Settings".
2. A window titled "Image Export Options" will open. Check the box "Convert to Data Type" and select "Unsigned 16-bit".
3. Check the box "Renormalize (min, max) from 0,00/1,00 to 0/65535
4. Under "Output File Info" select Format: tiff and under the file: navigate to the /Data/ tiffs folder and type in the file ending **/{nickname}_{result_type}.tiff**
5. Under the section "Batch Processing", click "Select Raw Data Files" and navigate to the Data/ tiffs folder and select all "_s2.h5" files. Then press "Process" all files.

Note: This step generates probablility masks that can be checked in any image viewer (ex. ImageJ FIJI) and in the pipeline the optimized probablility masks will be made into cell masks in CellProfiler.

Generate Cell masks:

1. Open CellProfiler and click on File > Import > Pipeline from File > Select: ImcSegmentationPipeline-development/ cp3_pipelines/ 2_segment_ilastik.cppipe.
2. Click on View output setting and confirm that the Input Folder is your Data folder and your Output folder is your tiffs folder.
3. Click on Images and drop all files from the tiffs folder into CellProfiler.
4. Click on Analyze Images.

Note: This step generates your single cell masks that can be imported into histoCAT.

Put masks in histoCAT folder:

1. In Jupyter, the imc_preprocessing script should still be open.
2. Run the cell at the very bottom of the script titled "Generate histoCAT folders with masks".
3. The folders in Data/ histocat can directly be loaded into histoCAT.

Address: INO-F608A, Inselspital, CH-3010 Bern
Contact: info.imc@dbmr.unibe.ch
Tel.: +41 31 632 88 95 or +41 31 632 22 94
Website: www.imc.unibe.ch

Extract single cell data for further analysis in R/Python (not needed for histoCAT)

1. Open CellProfiler and click on File > Import > Pipeline from File > Select: ImcSegmentationPipeline-development/ cp3_pipelines/ 3_measure_mask_basic.cppipe.
2. Click on *View output setting* and confirm that the *Input Folder* is your Data folder and your *Output folder* is your cpout folder.
3. Click on *Images* and drop all files from the tiffs folder.
4. Then click on *Metadata* and click *Update*.
   Note: This will automatically merge the acquisition_metadata.csv generated earlier in the script with your images.
5. Further click on *NamesAndTypes* and click *Update*.
6. Click on *Analyze Images*.

This creates 4 different CSV files in the cpout folder that can be imported into other analysis programs. Once file is called 'Cells' and has the extracted single cell data, another file is called 'Object Relationships' and has the neighborhood interactions in it. The remaining two files are less important with image parameters and cell profiler settings.

Pipeline completed!

Address: INO-F608A, Inselspital, CH-3010 Bern
Contact: info.imc@dbmr.unibe.ch
Tel.: +41 31 632 88 95 or +41 31 632 22 94
Website: www.imc.unibe.ch