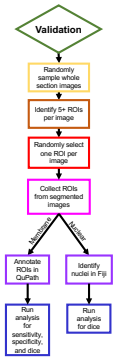
Validation

Revision: 12/13/2022 by Madi McElliott and Jeff Beamish



The following protocol outlines in detail how you may validate the accuracy of U-Net assisted analysis compared to a human reference. This assumes that both the "Pre-Processing" and "Training" protocols have been completed. This is required to validate training models for specific stains. Though, you do not need repeat this protocol for every project using U-Net assisted analysis.

BEFORE STARTING:

Set the script directory in QuPath by going to Automate→"Set Script Directory"→selected the folder where QuPath scripts are located. Should include (not all used in this protocol):

- Export Training Annotations.groovy
- Adjust BC.groovy
- Export Section Outlines.groovy
- Separate Sections.groovy

Fiji macros can be run using "Plugins" → "Macros" → "Run…" or by dragging the macro file onto the Fiji window and running from the editor (preferred in case small changes or debugging is needed).

This procedure uses tools and elements from other protocols as referenced below.

STEP ONE: Make a directory that will contain all the files for the validation process (Windows)

1. "**Original Images Full Size**" – contains the 10-15 full-sized, single channel images used for analysis
2. "**Original Images ROIs**" – contains the individual cropped images that are used for analysis
3. "**Selected ROI Locations**" – contains *.roi files that that indication the locations on the images in 2 above on the full size images in 3 above
4. "**Segmented Images Full Size**" – contains the 10-15 full-sized, segmented images that correspond to the original images in 1 above
5. "**Segmented Images ROIs**" – contains the individual cropped images extracted from the segmented full size images used for analysis
6. "**Reference Annotation QuPath Project**" – contains the QuPath project that will be used to make the reference annotations. Not used if validating a nuclear segmentation.
7. "**Reference Annotations**" – contains the reference annotations to be used for analysis. For membrane stains, cytoplasmic stains, or higher order structures, this will be a *.zip file that contains the annotations in the form of ImageJ ROIS (annotations) that can be loaded into the ROI manager. For nuclear stains, this will contain images with overlays that contain multipoints which identify each positive nucleus.
8. "**Segmented Images ROIs with Multipoints**" – contains the segmentation cropped images with an overlay of multipoints that identify positive detections. Not used for membrane stains, cytoplasmic stains, or higher order structures.
9. "**Validation Analysis**" contains the analysis and data files generated by the validation procedure.
10. Note one other directory will be generated below automatically "[Stain]-Extracted Validation Images (ALL)" which will contain all the extracted ROI samples before randomly selecting some for validation.

Commented [MM1]: This folder is empty int the mock project. I would like to delete it for the sake of keeping in concise.

STEP TWO: Generate single channel *.tif files for the project to use for validations (see Pre-Processing protocol)

*Refer to the Pre-Processing protocol to generate whole-section , single channel *.tif files. Ideally sections used for validation should be from a separate project than those used for training or at least should only utilize sections NOT used for training.*

Randomly sample whole section images

STEP THREE: Random selection of whole-section images for validations (Windows)

1. Use random number generator to select whole-section images for validation. Images should be free of large defects or artifacts (see **Pre-Processing STEP FOUR-**for quick qualitative assessment). For example if the project to use for validation has 72 original whole-section images, set the range of generator from 1-72, random numbers generated were 3,23, 28,54,67, ….
2. Copy images that correspond to randomly generated numbers into folder "**Original Images Full Size**" created in **STEP ONE – 1**.

STEP FOUR: Create ROIs for validation (Fiji)

Identify 5+ ROIs per Image

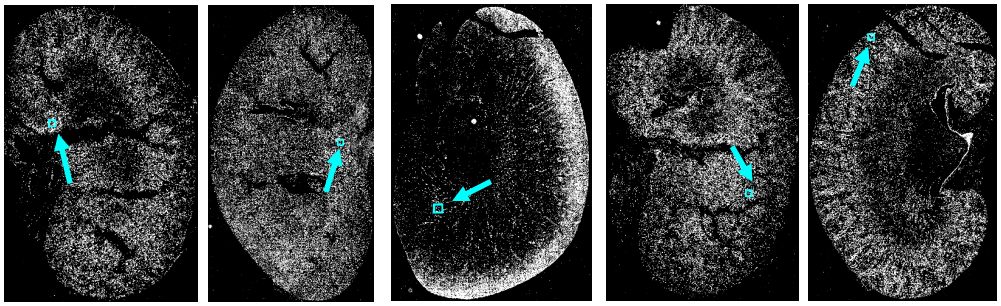
1) Install the Macro "Validation_STEP_FOUR_Make_Validation_ROIs.ijm"
    a. Go to FIJI→Plugins→Macros→Install…
    b. Note you can use set this macro as a shortcut in FIJI to allow for repeated ROIs to be made on the same image by adjusting the shortcut key in the macro name in the code (default is "[q]"). You must Install the macro for the shortcut to work. Double check that whatever short cut key you chose does not conflict with existing shortcuts.
    c. The default ROI size is 200 micron square, you may need to adjust this in the code (first executable line) if a different size is needed.
    d. This macro will create a separate folder within the "[Stain]" called "[Stain]-Extracted Validation Images (ALL)" that contains two subfolders of "ROIs (Cropped Images-ALL)" and "ROI Locations"
2) Open the first full-size image from the "**Original Images Full Size**" folder from **STEP THREE -- 2**
3) Adjust brightness and contrast to make the stain easy to see
4) Select ROIs to use for validation. These areas should be free of obvious artifacts. Exclude any ROIs that have a "positive" area that is very high (>90%) or very low (<10%), because validation metrics perform poorly at these extremes (i.e. the dice score would depend only on a few pixels). For nuclear segmentations, ROIs should have at least 5 positive nuclei (for the same reason, the dice score would depend only on few nuclei being detected).
    a. Run the macro "Validation_STEP_FOUR_Make_Validation_ROIs.ijm" using the shortcut key
    b. When prompted, first click on the general region where you want to put the ROI. For validation, the ROIs should be placed in the part of the section that will be analyzed later, for us this is generally the cortex + outer stripe of the outer medulla (OSOM). Positions in this regions should be selected at random except to avoid obvious artifacts (these would be areas excluded from analysis later). Then select OK
    c. The macro will then draw a preliminary ROI. Zoom in and reposition the ROI as needed by clicking in the center and dragging to a new position. For validation (in contrast to training), **ONLY** reposition the ROI to avoid an obvious defect. Once positioned, select OK and the macro will save the extracted images into "ROIs (Cropped Images-ALL)" and the coordinates of the cropped ROI will be in "ROI Locations".
5) Repeat a-c above to generate 5 images per full size image .
6) Close the image and load the next image from the "**Original Images Full Size**" folder from **STEP THREE -- 2**
7) Repeat steps 3-6 until all of the images have been processed.
8) For each set of ROI images from the same section (in the "[Stain]-Extracted Validation Images (ALL)/ ROIs (Cropped Images-ALL)" folder) use a random number generator to select the image to include (1 of 5). Copy these images into a folder "**Original Images ROIs**" created in **STEP ONE - 2** above.

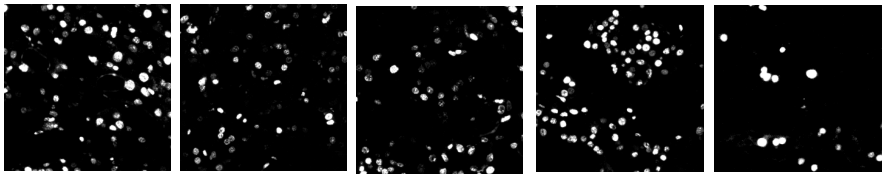
Randomly select one ROI per image

    a. For example, you would have created five ROIs for WT_GFP_M_0001_Ki67. Randomly select one of the five generated images for further analysis.

9) Copy the *.roi files for the selected images in **STEP FOUR--8** above (in the "[Stain]-Extracted Validation Images (ALL)/ ROIs Locations" folder) into a folder "**Selected ROI Locations"** created in **STEP ONE - 3** above.

Example Project Output:





STEP FIVE: Generate segmentation results for validation ("Analysis" Protocol, STEP SIX)

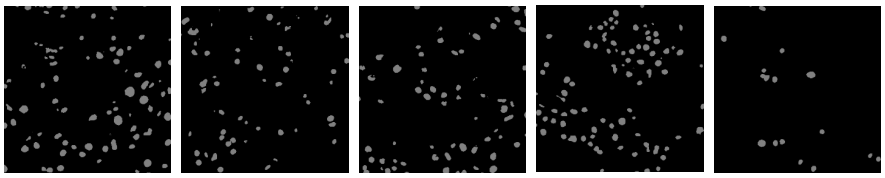
1) *Segment all the images in the "**Original Images Full Size**" folder using the U-Net model and weights you want to validate. Refer to the Analysis protocol, STEP SIX for detailed instructions for performing a batch segmentations. In analysis protocol STEP SIX – 6, specify "Original Images Full Size" folder from STEP THREE – 2 of this protocol.*
2) *For validation, we only need the results from the "Segmentation" subdirectory of the results. Copy these generated *.zip files into "**Segmented Images Full Size**" folder*

Collect ROIs from segmented images

STEP SIX: Crop segmented images used for analysis (Fiji)

1. Open Fiji
2. Open the first image in "**Segmented Images Full Size**"
3. Drag and drop the corresponding ROI coordinate file from the "**Selected ROI Locations"** folder created in S**TEP FOUR – 9**.
4. Image→Crop (or CTRL+SHIFT+X)
5. "Save As" the image with the same name and add "_ROI-#" to the end. Save it in folder named "**Segmented Images ROIs**"

Example Project Output:



STEP SEVEN: Make reference annotations and perform validation

This step is divided into two pathways depending on the type of stain being validated. Part A is for membrane stains, cytoplasmic stains, or higher order structures (for example peritubular capillaries) that requires an area to be outlined. We have found that the tools in QuPath facilitate these annotations best. Part B is for nuclear stains which are better analyzed using a point detection tool in Fiji and require a different analysis algorithm.

PATHWAY A: Membrane stains, cytoplasmic stains, and high-order structures.

Annotate ROIs in QuPath

PART I: Make reference annotations

1. Create a QuPath project using the directory made in **STEP ONE -- 6** ("**Reference Annotation QuPath Project**") and load all of the ROI image files from "**Original Images ROIs**" filled in **STEP FOUR - 8** above.
2. Within QuPath--Annotate cropped ROIs using the brush and wand tools

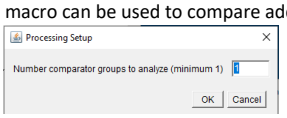
   Useful keyboard short cuts:
   - SPACEBAR-move the image without moving the annotation;
   - W-switch to the wand tool (used most of the time)
   - B-switch to the brush tool (useful when the wand tool is producing irregular results)
   - ALT-hold down while using other tools to REMOVE areas of the ROI
   - SHIFT-hold down while using other tools to ADD areas to the ROI that are not directly contiguous
   - CTRL AND SHIFT-allows you to approach other annotations without overlapping with them

3. Double check the annotations to make sure you've completed all images
4. Run the QuPath script "Export Training Annotations.groovy" using the "Automate" menu in QuPath. This script with generate *.zip files for each image that contain all annotations (which are stored as ImageJ "ROIs"). The *.zip files will be stored in the QuPath project directory.
5. Move these *.zip files into a new folder in the "**Reference Annotations**"

PART II: Perform validation analysis

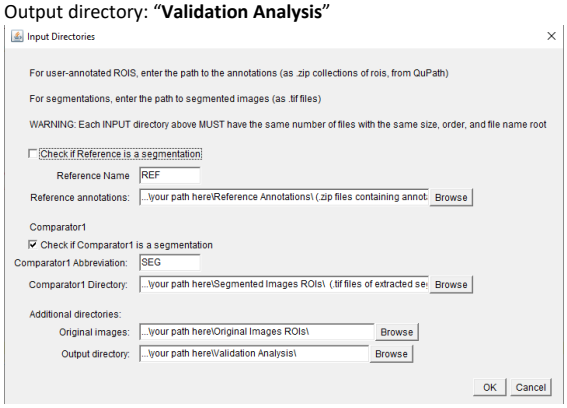
Run analysis for sensitivity, specificity, and dice

1. In Fiji, run the macro "Validation_STEP_SEVEN_A_Analyze_Validation_AREAS"
   a. In the first window, enter the number of comparators as 1 (Segmentation Only). Note: this macro can be used to compare additional segmentations as well



2. In the next window enter the paths to the appropriate folders as shown below. Each input MUST have the same number of ROIs of the same size, order, and file name root
   a. Check if Comparatory1 is a segmentation: UNCHECKED
   b. Reference name: REF
   c. Reference annotations: "**Reference Annotations**"
   d. Check if Comparator1 is a segmentation: CHECKED
   e. Comparator1 Abbreviation: SEG
   f. Comparator1 Directory: "**Segmented Images ROIs**"
   g. Original images: "**Original Images ROIs**"

h. Output directory: "**Validation Analysis**"



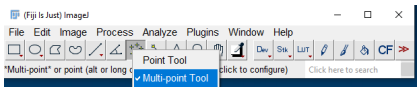
3. The macro will generate the following:
   a. A summary *.csv file which contains the relevant data
   b. For each image:
      i. A composite of the reference and Segmentation showing overlap (in green) without the original image.
      ii. An overlay on the original images showing the reference and segmentation as a semitransparent overlay.

PATHWAY B: Nuclear Stains

Identify nuclei in Fiji

PART I: Add Multipoints

1. Copy the images from "**Original Images ROIs**" filled in **STEP FOUR -- 8** above into the "**Reference Annotations**" from **STEP ONE – 7**.
2. Drag and drop one cropped ROI into Fiji
3. Annotate the images with the ImageJ multipoint tool.
   a. Select the multipoint tool to annotate each nuclei

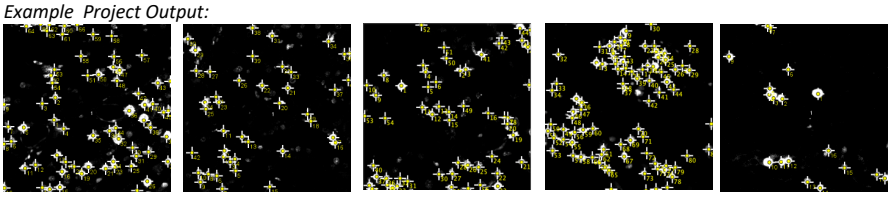


   b. Click on each positive nucleus
   c. To remove an incorrect or accidental point, click on it while holding the "ALT" key
   d. To change the size of the multipoint tool and make it more visible, double click the first multipoint created and select "ROI Properties"→"Size"→"Large"
4. Save (CTRL+S)
5. Convert the multipoint to an overlay. After placing all the multipoints, use following commands to create an overlay where the multipoints will be stored overlay.
   a. Make the current multipoints an overlay. Image→Overlay→Add Selection (or CTRL+B)
   b. Check to make sure there is only one set of multipoints (sometimes when editing, additional multipoint sets can be added and are difficulty to distinguish). First clear the ROI manager. Then move the overlay on the image to the ROI manager: Image→Overlay→To ROI manager
   c. The ROI manager should only have one entry. If not, delete any other extraneous entries listed in the ROI manager
   d. Move the overlay back to the image: Image→Overlay→From ROI Manager

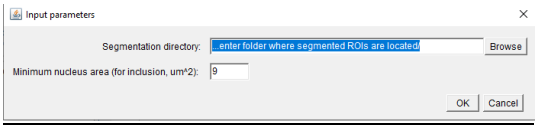
a. Save (CTRL+S)

6. *Optional:To revise multipoint rois that are already saved as overlays.
    a. Clear the ROI manager
    b. Move the overlay on the image to the ROI manager: Image→Overlay→To ROI manager
    c. Select the multipoint tool
    d. Select the entry in the ROI manager
    e. Edit and revise as in Step 4 above
    f. When done move the ROI back to the image as an overlay: Image→Overlay→From ROI Manager
    g. Save (CTRL+S)
7. Repeat for each image

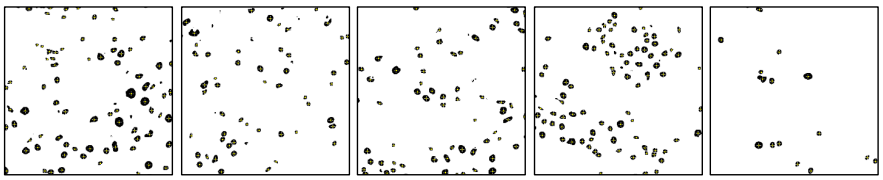
Example Project Output:



8. For nuclear stains you will also need to convert the segmentation results to points. We define a positive detection as anything > 9 micron^2 and mark the centroid.
    a. Copy the images from "Segmented Images ROIs" from **STEP SIX – 5** into the "**Segmented Images ROIs with Multipoints"** directory from **STEP ONE -- 8**
    b. Run code "Validation_STEP_SEVEN_B_Add_Multipoints.ijm". This will add multipoints to the segmentation images
        i. In "Segmentation directory:" select the path to the "**Segmented Images ROIs with Multipoints**" folder
        ii. Set the minimum nucleus size.
        iii. OK to run the macro



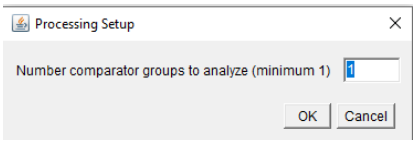
Example Project Output:



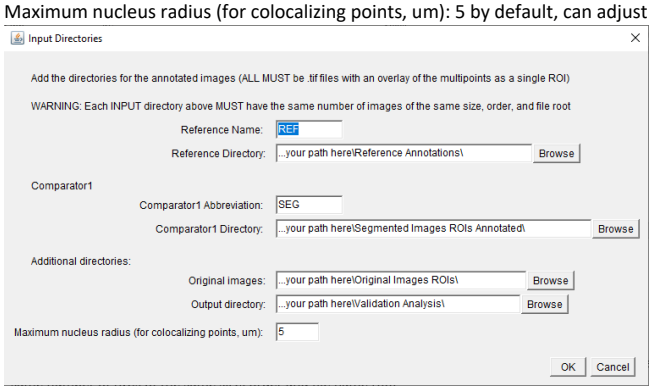
Run analysis for dice

PART II: Perform Validation Analysis

1. Run the code "Validation_STEP_SEVEN_B_Analyze_Validation_MULTIPOINTS.ijm" in Fiji
2. In the first window, enter the number of comparators as 1 (Segmentation Only). Note: this macro can be used to compare additional segmentations as well



4. In the next window enter the paths to the appropriate folders as shown below. Each input MUST have the same number of ROIs of the same size, order and file name root
    a. Reference name: REF
    b. Reference annotations: "**Reference Annotations**"
    c. Comparator1 Abbreviation: SEG
    d. Comparator1 Directory: "**Segmented Images ROIs with Multipoints**"
    e. Original images: "**Original Images ROIs**"
    f. Output directory: "**Validation Analysis**"
    g. Maximum nucleus radius (for colocalizing points, um): 5 by default, can adjust if desired



3. The macro will generate the following:
    a. A summary *.csv file which contains the relevant data
    b. A subdirectory "**Validation Analysis**" called "**Image Overlays (REF vs SEG)**" that contains images with an overlay showing the matched points.
        i. Green with yellow outline: matched point from the reference set (linked to its matched point by a green line)
        ii. Green with red outline: matched point from the segmentation set (linked to its matched point by a green line)
        iii. Yellow points: unmatched points in the reference set
        iv. Red points: unmatched points in the segmentation set
    c. A subdirectory "**Validation Analysis**" called "**Point Data Files**" which contains *.csv data files that lists each point, its coordinates, and the name of the point in the corresponding comparison data set that it was matched to, if applicable. Each comparison will have two files, one for each image.