Lab 3.
Segmentation Parameterization:
Mean-Shift Segmentation

Due Date: February 16, 2022
Form for submission: [Link]

Learning Objectives

1. Understand the impact of various parameters have on the image segmentation process
2. Implement segmentation using your choice of tool

A popular clustering algorithm used for image segmentation is the Mean-Shift segmentation. This process is a region growing method and the general steps are:

1) Seed point and local neighborhood is defined. The smaller the region the fewer points can affect the density distribution. A global neighborhood flattens out the data and estimates a single mean.
2) Compute local density (pdf in non-parametric method)
    a) Kernel function used to ave over the neighboring points. The closer the point is in space to the seed point the most impact it has
    b) Kernel Density Estimation
        i) Use smoothing kernel
        ii) Typical choice uses a Gaussian Kernel
3) Recenter the region of interest to the new center of mass using the shift vector. (Think of walking towards the areas where density increases)
4) Repeat 2 & 3 until convergence

Instructions

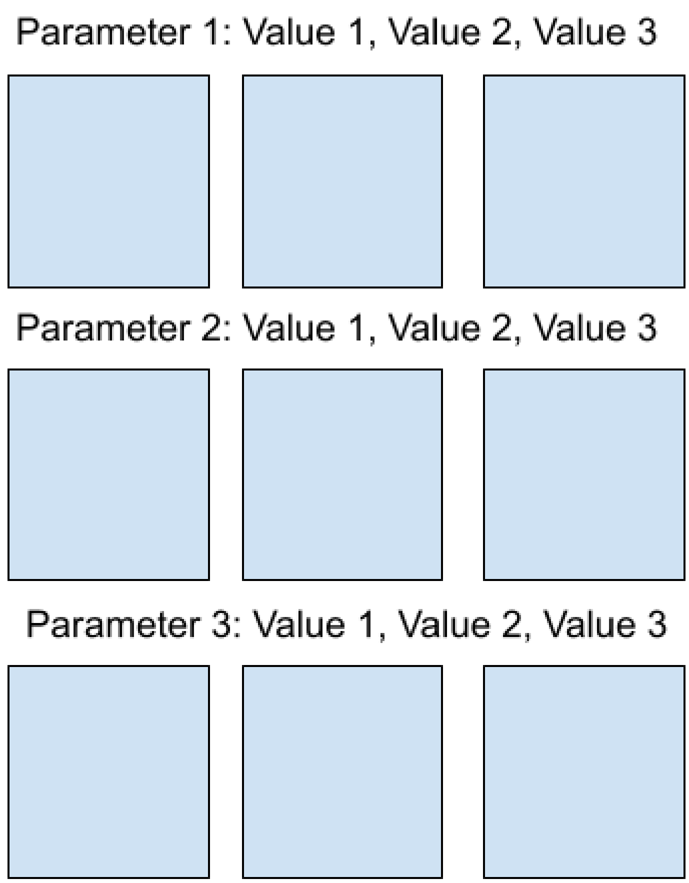
For this assignment, you will perform the mean-shift segmentation algorithm on two images (one rural and one urban). I have already clipped the images to a relatively small size so as not to overburden the computer. You can use whichever tool you are most comfortable with to perform the segmentation- ArcPro, Python, or QGIS. The set up for the three different tools is a bit different and I will provide the parameters for each below.

For each image

1. Select one parameter at a time to change (for example bandwidth). For the first implementation hold all parameters at their default values.
2. Change the parameter you selected in Step 1 by some increment. Hold all other parameters at their defaults.

3. Repeat the process, changing that original parameter again, preferably by the same value (so if you were changing bandwidth from 4 to 6 in step 2, change from 6 to 8 in this step).
4. Repeat one more time so that you have completed three segmentations with varying that parameter (so for example, three segmentations of bandwidth 4, 6, 8).
5. Return the parameter you just changed to the default value, select a second parameter to change and repeat steps 1-2 for this parameter for three different values.

By the end of this process you should have multiple segmented images that you can evaluate. As the second part of your lab, you should write a short paragraph or two about what changes you see across the images as you change the parameters. An optimal layout for your final images submission is illustrated below. Feel free to incorporate the written part either separately or around the images.



Mean-Shift Algorithm Parameters

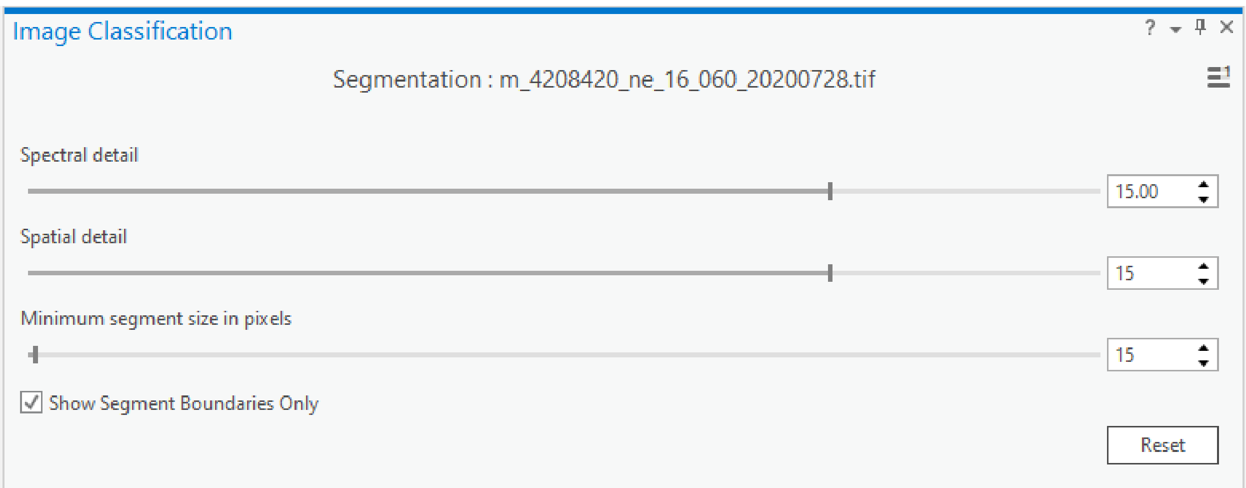
Each of the tools use different names for the same basic parameters. In order to clarify them for you, I have made a list of them for each of the tools below.

ArcPro

Spatial Detail (range 0-20): level of importance given to the proximity between features. A smaller value for this property will result in smoothed outputs.

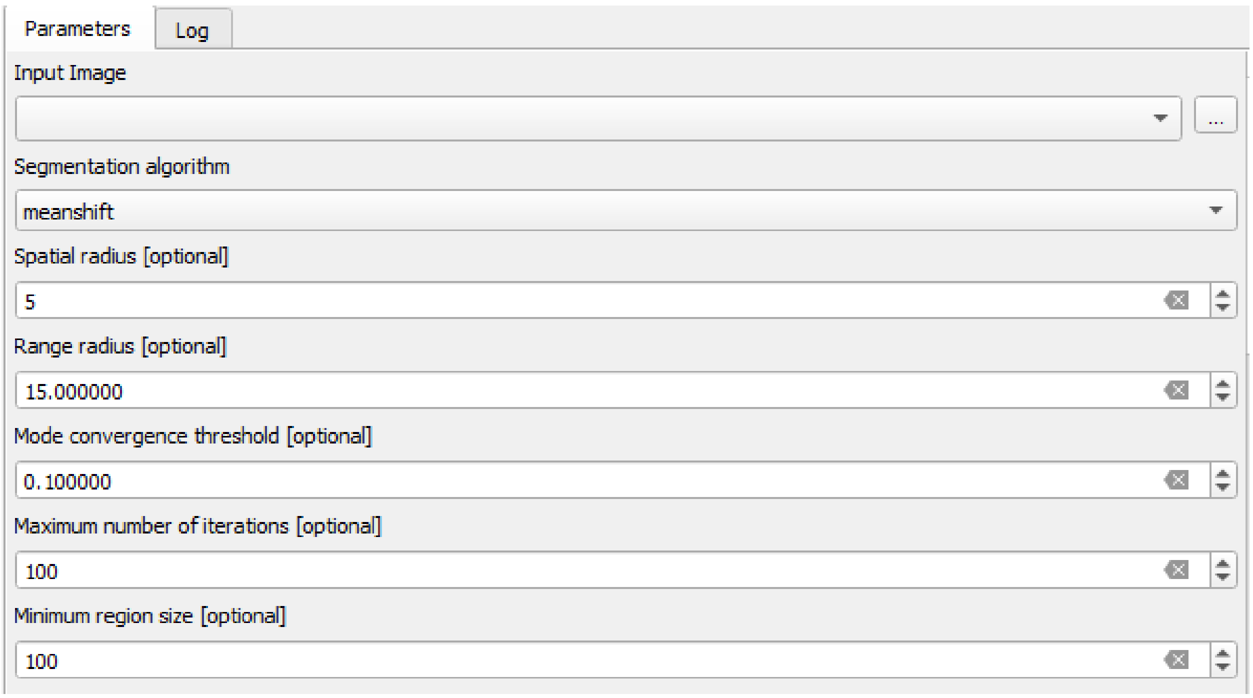
Spectral Detail (range 0-20): level of importance given to the spectral differences of features. A higher value will emphasize smaller differences between features.

Minimum Segment Size (unit in pixels) Segments smaller than this size are merged with their best fitting neighbor object.



QGIS

Spatial Radius: neighborhood size definition
Range: the radius (expressed in radiometry unit) in the multispectral space.
Mode Convergence iterations end when the mean-shift vector is below this threshold or if iteration number reached maximum number of iterations
Minimum Region Size minimum size of a region (in pixel unit) in segmentation



SKLEARN

**Note that this package is parameterized very differently than the other two, so you only really have control over bandwidth, which can be set algorithmically with estimate_bandwidth, or manually by setting the value in the Mean Shift function. If you choose to do this option, evaluate this parameter along with the quantile parameter- specifically, how does changing the quantile parameter affect the bandwidth estimate?

```
sklearn.cluster.estimate_bandwidth(X, quantile=X, n_samples=X,
random_state=0)


class sklearn.cluster.MeanShift(*, bandwidth=None, seeds=None,
bin_seeding=False, min_bin_freq=1, cluster_all=True, n_jobs=None,
max_iter=300)
```

Quantile (range 0-1, median = 0.5): determines the number of pairwise distances used

N_samples: number of samples to use in estimation
Random state: random sample generator
N_jobs: number of parallel jobs, -1 means all processors.
Bandwidth: bandwidth used in the RBF kernel

For advanced programming students: If you would like to compare the output of the SKLearn mean-shift algorithm with another mean-shift algorithm from a different package, that would be really interesting.