

Assignment 8: Supervised Classification

(60 Points Total)

Data available under [Resources>Supervised Classification](#).

You have been provided with a Sentinel-2 image of Vancouver, British Columbia ([sentinel_vancouver.img](#)) that was provided by the European Space Agency (ESA) and downloaded from the United States Geological Survey's EarthExplorer. These data have been provided at a 10 m spatial resolution with the following band designations:

Layer_1: Blue (Band 2)

Layer_2: Green (Band 3)

Layer_3: Red (Band 4)

Layer_4: NIR (Band 8)

Layer_5: SWIR (Band 11)

Layer_6: SWIR (Band 12)

For reference, these are the Sentinel-2 band designations.

Sentinel-2 Bands	Center Wavelength (μm)	Spatial Resolution (m)
Band 1: Coastal Aerosol	0.443	60
Band 2: Blue	0.490	10
Band 3: Green	0.560	10
Band 4: Red	0.665	10
Band 5: Red Edge	0.705	20
Band 6: Red Edge	0.740	20
Band 7: Red Edge	0.783	20
Band 8: Near Infrared (NIR)	0.842	10
Band 8A: Narrow Near Infrared (NIR)	0.865	20
Band 9: Water Vapor	0.945	60
Band 10: Shortwave Infrared Cirrus	1.375	60
Band 11: Shortwave Infrared (SWIR)	1.610	20
Band 12: Shortwave Infrared (SWIR)	2.190	29

You have also been provided with a set of validation points ([validation_points.shp](#)). The codes in the "class" field represent the following classes:

0 = Water

1 = Developed

2 = Barren

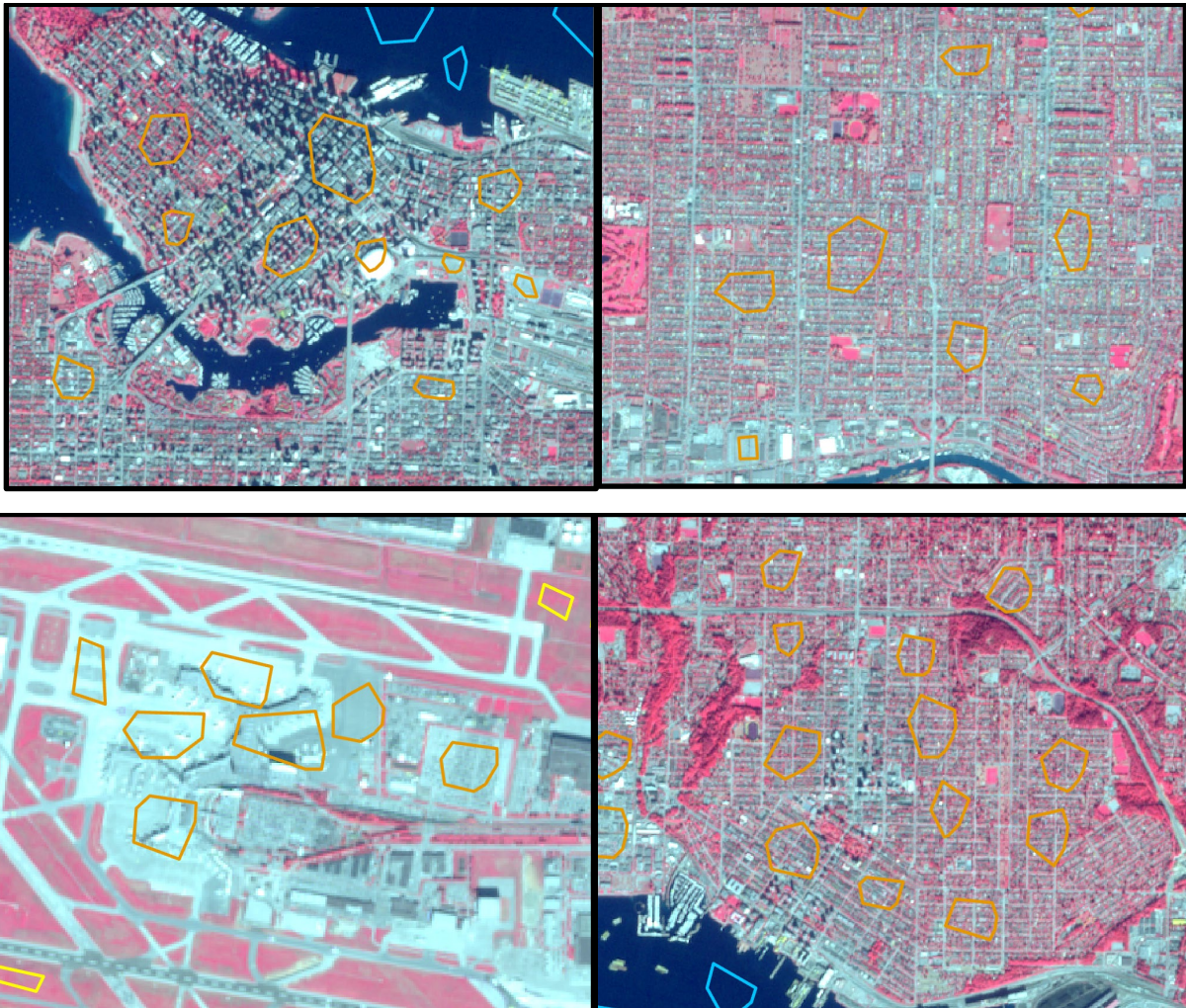
3 = Forest

4 = Herbaceous

The examples below describe the classes of interest.

Developed

This class will include all commercial, urban, and residential areas. Common features within developed areas include buildings, roads, yards, and parking lots.



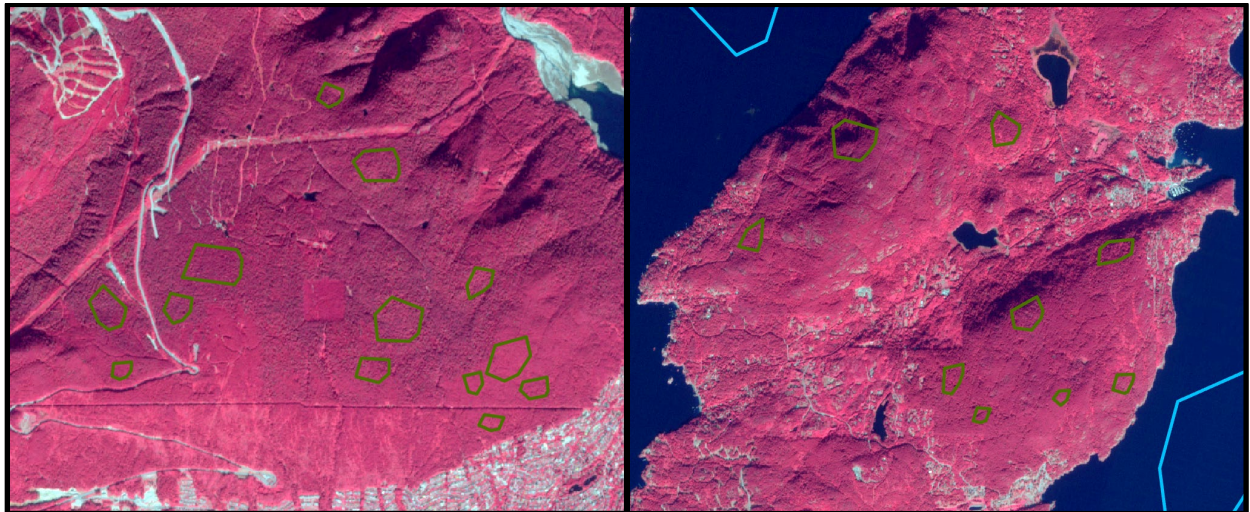
Barren

This class will include non-vegetated areas not associated with development, such as bare rock or soil. This is not a common cover type in this image. However, there are some bare rock surfaces in the mountainous area in the top part of the image north of the city.



Forest

This class will include all forested areas in the image. Do not include small stands of trees in residential areas.



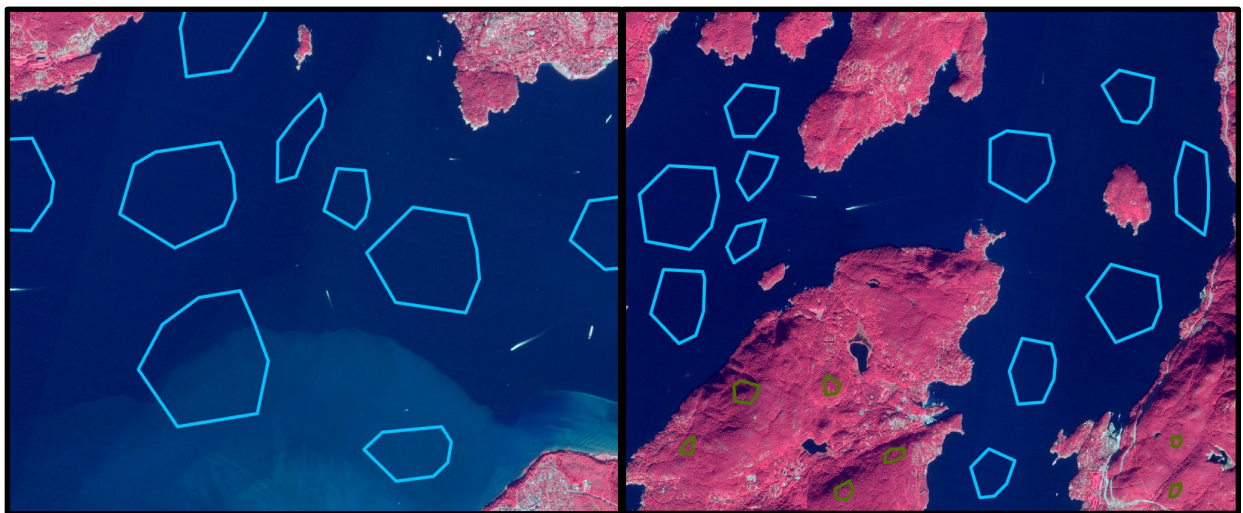
Herbaceous

This class will include vegetated areas that are not dominated by trees, such as fields, pastureland, and cropland.



Water

This class will include all water features, such as the ocean, ponds, lakes, and rivers.



Description of Problem

Produce two land cover classifications that differentiate these five classes and compare them using confusion matrices and the provided validation data. Use the Orfeo Toolbox within QGIS.

You will need to do the following:

1. Create a shapefile in which to digitize your training data. Make sure to include a “class” field as an integer type. Use the codes for each class as defined in the validation data.
2. Digitize a variety of examples of each type. Make sure to capture the full spectral variability and variety within each class. Make sure to assign the correct code to each sample. You can create training data as points or polygons.
3. Compute statistics for the [sentinel_vancouver.img image](#) using the ComputeConfusionMatrix tool from the Orfeo Toolbox. Save it as an XML file.
4. Train two classifiers using the TrainImageClassifier tool from the Orfeo Toolbox and your training samples. You do not need to optimize the algorithms.
 - a. SVM with RBF kernel
 - b. Random Forests
5. Classify the image using both models with the ImageClassifier tool from the Orfeo Toolbox.
6. Calculate confusion matrices for each classification using the ComputeConfusionMatrix tool from the Orfeo Toolbox. You will need to change the value for NoDATA pixels to 5 since 0 is assigned to a class.

Deliverables

- Create a single layout showing the input image, two classification results, and training data over the input image. You should have a total of four map frames. Also include a title and text labels for each map frame. Include the error matrices for both classifications. They should be well formatted and include the overall accuracy and class user’s and producer’s accuracies. These can be formatted outside of QGIS and then inserted as images. Include a legend that explains the colors assigned to each class. On the layout, write a short paragraph to summarize the results of the algorithm comparison and the sources of confusion between the classes. Render your results as a PDF.
 - The four map frames are well organized on the page. **(6 Points)**
 - Each map frame is labelled with text. **(3 Points)**
 - The same colors are used for each class in both classifications and to show the training data. **(6 Points)**
 - A descriptive main title and citation for the European Space Agency are included. **(3 Points)**
 - A north arrow and scale bar should be included for just one of the maps. **(6 Points)**
 - The error matrices are well formatted, use the class names for the rows and columns as opposed to the class codes, and include the overall accuracy and class user’s and producer’s accuracies. The counts in the table should be rounded to whole numbers while the overall accuracy and class user’s and producer’s accuracy should be rounded to three decimal places. **(12 Points)**

- The paragraph clearly summarizes the results by comparing the RF and SVM output in regard to overall accuracy. The paragraph should also include a description of the sources of error or confusion between the classes. **(12 Points)**
- The map should be overall neat, well organized, and use the space well. **(12 Points)**