

Classification Tutorial 2: Collect Training Data

This tutorial is the second in a two-part series about preparing data for image classification. It requires ENVI 5.4 or later and will take approximately one hour to complete. You will learn how to interactively define regions of interest (ROIs) that represent known land-cover types. The ROIs serve as ground truth, or training, data. Training data is required for supervised classification.

Files Used in this Tutorial

Tutorial files are available from our [website](#) or on the ENVI Resource DVD in the `classification` directory. Copy these files to your local drive.


Data files were provided by the National Ecological Observatory Network, 2014. Files were accessed in July 2016, available online at <http://data.neonscience.org> from the National Ecological Observatory Network (NEON), Boulder, CO, USA.

File	Description
NEONOrthophoto.dat	Digital photograph from the NEON RGB camera at 0.25 meter resolution
NEONOrthophoto.hdr	ENVI header file for above

Open and Display an Orthophoto

In this section you will open a digital orthorectified photograph that was captured at the same time as the NEON image used in "Classification Tutorial 1: Create an Attribute Image." The two images are co-registered, so you can collect training data from the orthophoto and apply it to the multispectral image or the attribute image. You will collect training data from the orthophoto because it has a higher spatial resolution (0.25 meters) than the `NEONUrban.dat` image (0.5 meters) used in the first tutorial. The colors in the orthophoto are also more natural because it was captured with an RGB camera.

Follow these steps:

1. Start ENVI.
2. Click the **Open** button  in the toolbar.
3. Navigate to the directory where you saved the tutorial files.

4. Select `NEONOrthophoto.dat` and click **Open**. The orthophoto appears in the Image window.





Create Regions of Interest

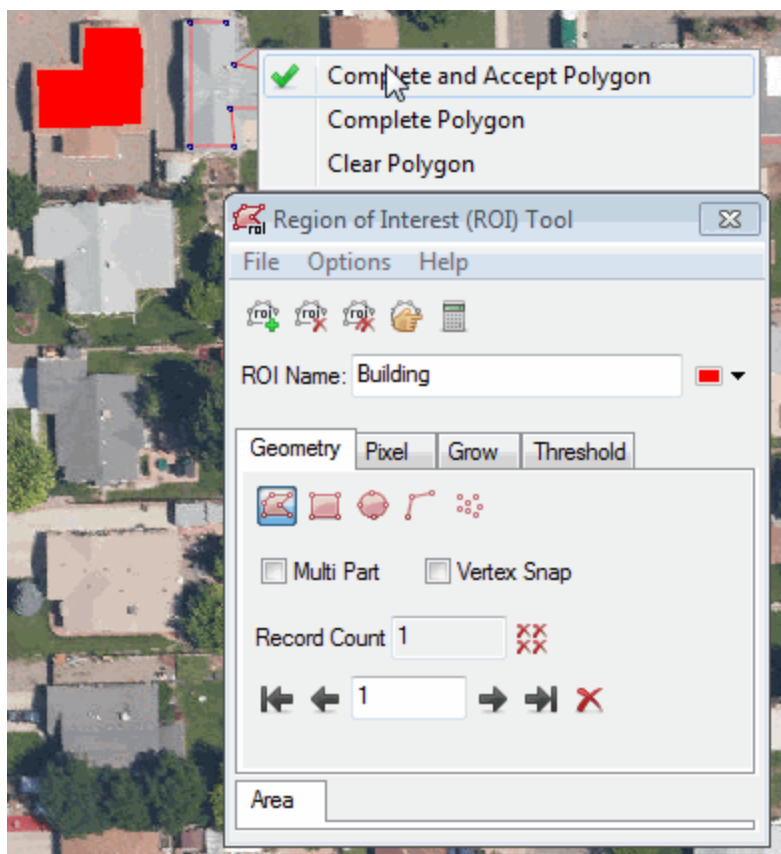
In this section you will define training data ROIs for five different land-cover types:

- Building
- Tree
- Grass
- Concrete
- Asphalt

This is a simple use case with a few classes. Real-world land-use classification typically involves many different feature types and training samples.

Collect Training Data for Buildings


1. Click the **Region of Interest (ROI) Tool** button  in the toolbar.
2. Click the **Add ROI** button  in the Region of Interest (ROI) Tool. A new Regions of Interest layer appears in the Layer Manager. The default name is "ROI #1" and its color is red.
3. Change the **ROI Name** to **Building**, and use the color picker to choose a different color if desired.
4. Zoom into the orthophoto.
5. The Polygon icon is active in the ROI Tool. Click and drag the cursor to define a polygon around some building pixels. To complete the polygon, right-click and select **Complete and Accept Polygon**. For example:



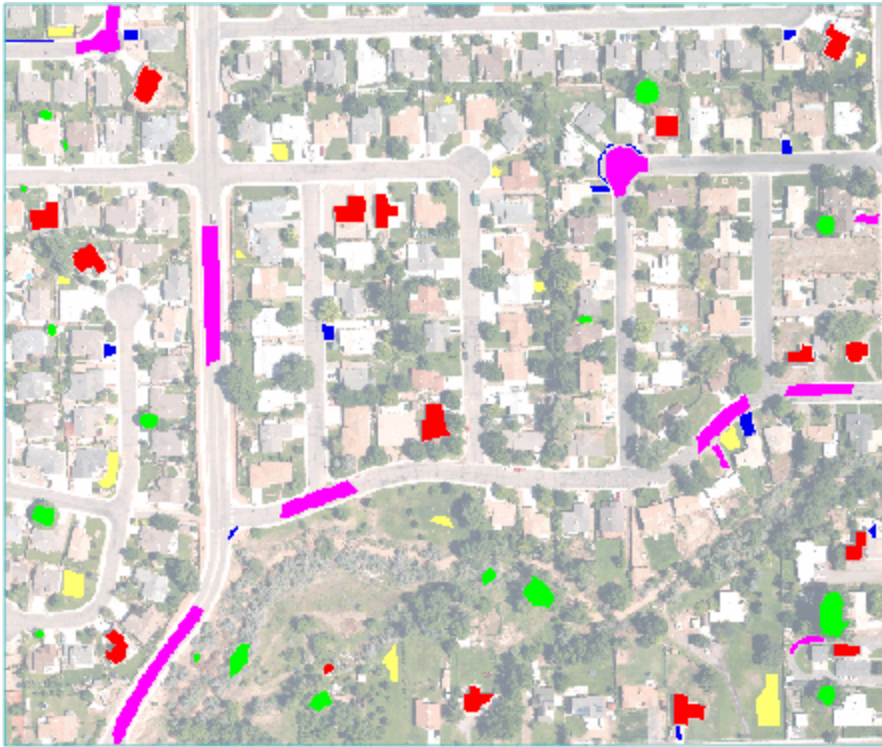
To pan around the image while creating ROIs, press the middle mouse button and move around the display. You can also click the **Pan** button in the toolbar. To resume drawing ROIs, you must (1) click the **Select** button in the toolbar and (2) click inside of the ROI Tool. The cursor will change back to a crosshair.

6. Create several more polygons around building pixels. You do not have to draw ROIs for every building. The key is to draw ROIs around buildings with different colors, intensities, and features (including solar panels and vents) so that you have a wide representation of building types.

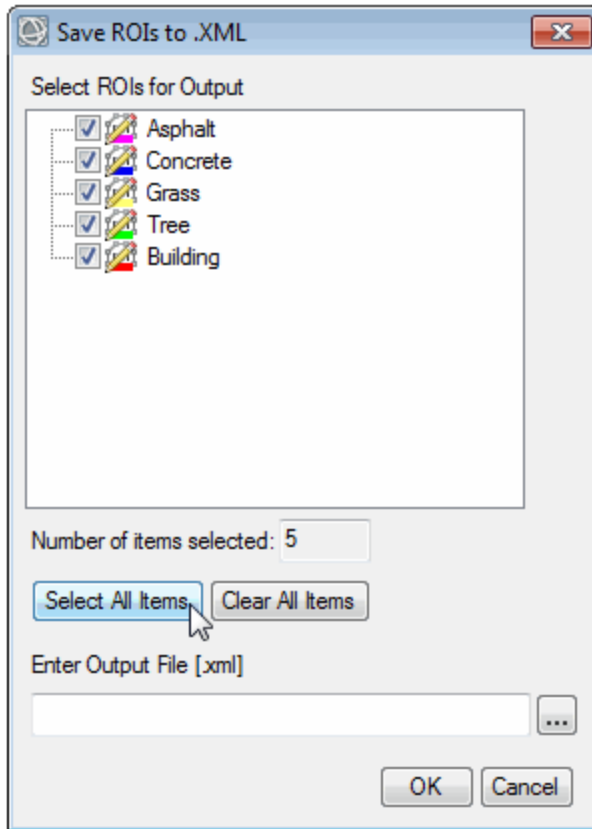
Collect Training Data for Other Land-Cover Types


1. Click the **Add ROI** button  in the Region of Interest (ROI) Tool. A new *Regions of Interest* layer appears in the Layer Manager. The default name is "ROI #1" and its color is green.
2. Change the **ROI Name** to **Tree**.
3. Define polygons that represent tree pixels. Choose different colors of trees.
4. Repeat Steps 1-3 for the other ROIs:
 - **Grass**: Select pixels that are purely grass, not a mixture of bare soil and grass.
 - **Concrete**: Select pixels from sidewalks and driveways.
 - **Asphalt**: Select pixels from roads and asphalt parking lots.

Here is an example with all five ROIs defined:



5. When you are finished creating ROIs, select **File > Save As** from the ROI Tool menu bar.
6. In the Save ROIs to .XML dialog, click the **Select All Items** button to select all five ROIs.



7. Click the  button next to the **Enter Output File** field.
8. Select an output folder, and name the output file `TrainingDataROIs.xml`.
9. Click **OK** in the Save ROIs to .XML dialog.
10. Exit ENVI.

Next Steps

Now you can use the attribute image (from "Classification Tutorial 1: Create an Attribute Image") and training data ROIs from this tutorial for classification. Here are some different options:

- See the following code examples in ENVI Help for instructions on running an end-to-end classification workflow using the attribute image and training ROIs:

- Code Example: Support Vector Machine Classification Using API Objects
- Code Example: Softmax Regression Classification Using API Objects
- Or, use the ENVI Classification workflow by selecting **Classification > Classification Workflow** from the ENVI Toolbox. When the workflow starts, select an input **Raster File** and click **Next**. Select the **Use Training Data** option, then select `AttributeImage.dat` for the source image and `TrainingDataROIs.xml` for the training data.