Title: Classifying Hurricane Damage Imagery with Machine Learning

Critical Resources: an internet connected computer, a zip tool, ArcGIS Pro 2.x (with Spatial Analyst and Image Analyst licenses), Lab 5 datasets.

Purpose: The purpose of this lab will be for you gain experience and knowledge with how to use and evaluate a machine learning algorithm (Support Vector Machine (SVM)) to classify imagery in a disaster response context.

Notes: Datasets for this exercise were downloaded from the NOAA Hurricane DORIAN Imagery website (https://storms.ngs.noaa.gov/storms/dorian/index.html). An area near Marsh Harbor on Abaco Island in the Bahamas was selected as a representative example of the massive damage that was incurred on Abaco Island in terms of damaged trees, blue tarp roof cover (used to cover damaged roofs), destroyed building, and intact roofs (or roof tops that have not been damaged).

A sample of imagery was extracted from a larger NOAA Hurricane DORIAN imagery tile for ease of computational processing in this lesson. The full tile where the sample imagery was extracted from was also included with the lesson files. A full complete NOAA hurricane Dorian imagery dataset can be downloaded from the website above and consists of about 10 GB of TIFF and JPEG files. These datasets are an excellent example of post-disaster imagery.

A new classification schema was created in the ArcGIS training samples manager[1] that defined analysis classes for the imagery. These schema included classes such as damaged trees, blue tarp covers, intact roofs, destroyed buildings and open roads ways. These classes are by no means representative of all the types of imagery analysis that could be conducted on the devastating effects of hurricane Dorian and they are meant simply to be a starting point that can

[1] https://pro.arcgis.com/en/pro-app/help/analysis/image-analyst/training-samples-manager.htm

be expanded upon as a learning exercise and imagery analysis for disaster response. Training sample polygons were then created from the classification schema. The training samples were saved as a polygon shapefile along with the classification schema, both of which have been included with the lesson files that can be downloaded from the book companion website as a comparison to the datasets that will be created via specific training steps in this lesson. **However, you are required to create your own assessment datasets for this exercise and cannot simply use the provided assessment datasets.**

Learning Objectives: After completing the exercise, you will know:

- How to create training samples to classify imagery with Support Vector Machine Algorithm
- How to assess the results of a machine learning algorithm by creating and modifying accuracy assessment points to "ground truth" and compare the results of a machine learning algorithm against human judgement.
- How to assess the results of a machine learning algorithm by creating and interpreting a confusion matrix that quantifies the results of comparing a machine learning algorithm classification with a human classification.

Deliverables:

Maps created via response to the instruction questions. Upload your write-up to the lab 5 drop box on myCourses.

Overview:

Hurricane Dorian, which occurred in September of 2019, was considered one of the most devastating hurricanes to hit the Bahamas since record-keeping began[2]. Abaco Island, Located in the eastern part of the Bahamas, particularly received devastating impacts from the hurricane with over 35 people reported killed because of the hurricane. Satellite imagery was used to conduct analysis of the Hurricanes devastating impacts and shows in stark detail the impacts of Dorian[3]. In this exercise, you will work with actual imagery that was collected from Abaco Island by NOAA as an exercise in learning how to use machine learning tools to classify satellite imagery in a disaster management context.

Task 1 – Download data sets

Step 1: Down the lab_5_datasets.zip from myCourses

Step 2: Unzip the datasets

Unzip the datasets to a location where you can find them easily, like the C:\temp folder.

Before proceeding with the next tasks, open ArcGIS Pro.

Task 2 – Add datasets

In this task, your will add all of the dataset that were downloaded in task 1 to your map. The software operation for adding data layers is the same for each layer so only the general steps are provided.

Step 1: Add all the exercise datasets to to your map

In this step, you will add all the exercise datasets (a .tiff) of a disaster .

ArcGIS Pro Quick Steps Reference: Map Tab > Add Data > Data > Browse for Dorian_Excerpt.tif

[2] https://www.bbc.com/news/world-latin-america-49553770
[3] https://www.washingtonpost.com/weather/2019/09/05/satellite-images-show-destruction-dorian-bahamas/

Add the Imagery basemap available in ArcGIS Pro as well as this gives a comparison of what this area looked like before Dorian struck as seen in the Dorian_Excerpt image.

Task 3 – Load Classification Schema and Create Training Samples

In this task you will load a classification scheme of that has been provided with this exercise and you will use this classification schema to start designing training sample areas within the imagery provided with this lesson.

<u>Step 1: Load Classification Schema</u>

The ArcGIS pro Imagery tools are activated by clicking on various imagery layers. To start loading the classification scheme are provided with this lesson, click on the Dorian_Excerpt layer that you added to your map is Task 2, Step 1.

Next, click on the Imagery Tab.

From the Imagery Tab, select Classification Tools > Training Samples Manager.

The Training Samples Manager will load with a default classification schema if this is your first time using this tool. For this exercise, you will load the classification schema that was provided with this lesson.

In the top half of the Training Samples Manager tool, select Classification Schema > Browse to Existing Schema and load Dorian_Imagery_Analysis.ecs provided with this lesson. The Dorian_Imagery_Analysis.ecs should now be loaded into the Training Samples Manager and you can use this schema to generate training samples.

<u>Step 2: Generate Training Samples</u>

In this next step, you will digitize training samples from the Dorian image. The idea here is to define specific polygons that are representative of different classes that will be generated by the

SVM classification algorithm, making SVM what is known as a “supervised” algorithm[4] as the training samples created by a human will help to train the algorithm on what to look for. For example you will draw polygons around damaged buildings, blue tarps, and intacts roof and any other classes defined in the classification scheme of that you just loaded into the training sample managers tool. The SVM algorithm will then use these samples to classify the entire image. The following steps outline how to digitize an area of damaged trees as a sample class, but the steps are the same for all of the classes that are defined in the classification scheme of that you loaded in the previous step.

To create a damaged tree sample class[5]:

Under the classification schema, click Environmental_Damage and expand out from underneath Damaged trees.

Click the Polygon tool.

Trace a polygon around the area located at -77.074452, 26.540376

This is a damaged tree area.

After creating the polygon, it will be displayed under the sample classes that are below classification schema.

Repeat this general process for the other classification schema classes (and remember to select the class you want to digitize first before selecting a tracing tool. Figure 1 is a loose guideline for what you are looking for.

[4] See: https://pro.arcgis.com/en/pro-app/help/analysis/image-analyst/overview-of-image-classification.htm for a discussion of supervised and unsupervised image classification in ArcGIS Pro.

[5] Full reference on these steps: https://pro.arcgis.com/en/pro-app/help/analysis/image-analyst/training-samples-manager.htm

Figure 1 – Examples of features to digitize for training samples. This image is roughly centered at -77.074413, 26.540590.

Given the relatively small size of the image, you can generate 1-2 samples for each class. The shapefile Dorian_Samples.shp provided with the exercise datasets are training classes that can be used as a comparison with the classes you create. Note that it is important that when you have completed creation of your training sample, save the training samples as a feature class with the name Abaco_Samples.

Task 4 – Run Support Vector Machine Algorithm to Classify the Image

In this task you will now run the Support Vector machine algorithm using the training samples that you created in the previous task to classify the image. To run the support Vector machine algorithm, follow these steps:

Assuming you have Dorian_Excerpt.tif highlighted in content pane, from the Imagery Tab, select Classification Tools > Classify.

The Image Classification tool will open.

At the top, it should say: Dorian_Excerpt.tif

Under classifier, select Support Vector Machine

Under training sample, select or browse for the Abaco_Samples you created in Task 3, Step 2.

Under Output Classified Dataset, name the output: Classification_1 (to indicate this is the first classification in case you to do multiple classifications based on adjustments to training samples or other factors).

You can use the default settings for the other tool parameters.

Click the Run button.

The overall process to classify the image will vary based on the number of samples and computing power you have[6]. For example, when preparing this exercise, the Training Classifier procedure took 1 minute 57 seconds using the Dorian_Samples shapefile provided with the exercise, generating the classified dataset took 22.37 seconds.

The classified data set that will be produced by the SVM algorithm will be a raster dataset that applies the colors that were defined in the classification scheme the output raster.

Create a map product of this classification result as a deliverable for this lab.

In the next task, you will learn how to assess the results of the classification in order to determine if the classification is sufficient for analytical needs or if further adjustments are needed made in terms of training samples.

Task 5 – Assess the Classification: Create Accuracy Assessment Points

The value of a machine learning algorithm is ultimately determined by how effective the

[6] See *Exercises Technical Note* at the beginning of the book for details on the computer used to create the book's exercises.

algorithm is at actually classifying items correctly. Determining the accuracy of a machine learning algorithm involves the general process of comparing samples that have been "ground truthed" by a human with the same samples that have been classified by the machine. In this step, you will generate accuracy assessment points as a first step of evaluating the accuracy of the SVM algorithm that was run in Task 4. Generating accuracy assessment points can be a somewhat labor-intensive process, so a set of accuracy assessment points have been provided with this exercise that you can use instead. Follow these steps to use the Create Accuracy Assessment Points tool[7] in ArcGIS Pro:

Analysis Tab > Tools > Geoprocessing > enter Create Accuracy Assessment Points into the Find Tools search box. Click Create Accuracy Assessment Points (Image Analyst Tools) from the search results (should be top of the list).

Enter the following into the Create Accuracy Assessment Points tool:

Input Raster or Feature Class Data Set: Classified_1 (created in Task 4)

Output Accuracy Assessment Points: dorian_assesment_points (save as a shape file or feature class in a file geodatabase)

Target Field: Classified

Number of Random Points: 100 (the more points, the more ground truthing that will need to be done)

Sampling Strategy: Stratified random

Click the Run button,

ArcGIS Pro will then generate a series of assessment points. Compare each assessment point

[7] Tool reference: https://pro.arcgis.com/en/pro-app/tool-reference/image-analyst/create-accuracy-assessment-points.htm

with the value of the classified raster that was created in Task 4. Update the GrndTruth (ground truth) field in the assessment points from -1 to the correct code value based on the numeric codes from the classification schema you first worked with in Task 3. This is where you are adding your judgement as a human. For example, if the assessment point says that a classified cell is value 11 (Damaged_Tree) and your human inspection of the Dorian_Excerpt image shows that the assessment point is in fact over an area of damaged trees, the SVM algorithm classified this pixel of the Dorian_Excerpt image correctly and thus the GrndTruth (ground truth) field in the assessment points attribute table should be updated from -1 to 11 (Damaged Tree). Repeat this verification based on human judgement procedure process for all of the features in the assessment point feature class. Use the imagery base layer provided in ArcGIS Pro to compare the Dorian disaster image with the condition on Abaco island before Dorian struck to make more informed judgements.

After you have ground-truthed each of the assessment points, you have a couple of choices as how to proceed next. For example, you could go back and redo your training samples and reclassify the image based on any errors that you found when doing your ground-truth assessment. Then, after re-running the classification a second (or more) time(s), you could then use the Update Accuracy Assessment Points tool[8] to adjust your ground-truth round points to account for changes that you made in the reclassification. However, if you are satisfied with your assessment points you can they move on to the final stage of assessing the accuracy of the machine learning algorithm and compute a confusion Matrix, a process that is described in the next step.

[8] Tool reference: https://pro.arcgis.com/en/pro-app/tool-reference/spatial-analyst/update-accuracy-assessment-points.htm

Task 6 – Assess the Classification: Compute Confusion Matrix

A confusion matrix describes the performance of a classification algorithm by comparing known data values with data values from a classifier. A confusion matrix is presented in a table format with rows representing actual values and columns corresponding to those rows with predicted values. In the case of this exercise, we just want to compare the ground truth assessment points that were created in task 5 with the classification that was generated in Tab 4. The Confusion Matrix will give us a variety of quantitative value to determine the overall utility of the SVM algorithm and inform possible modifications that need to be made to improve the accuracy. ArcGIS Pro provides a tool to compute a confusion matrix. Follow these steps to compute a confusion matrix:

Analysis Tab > Tools > Geoprocessing > enter Compute Confusion Matrix[9] into the Find Tools search box. Click Compute Confusion Matrix (Image Analyst Tools)[10] from the search results (should be top of the list).

Enter the following into the Compute Confusion Matrix tool:

Input Accuracy Assessment Points: dorian_assesment_points (created in Task 5)

Output Confusion Matrix: dorian_confusion_matrix (save as a .dbf file or standalone table in a file geodatabase)

Click the Run button.

[9] Tool reference: https://pro.arcgis.com/en/pro-app/tool-reference/image-analyst/compute-confusion-matrix.htm
[10] Also available through Spatial Analyst tools.

Table 1 shows the results of the confusion Matrix computed when developing this exercise.

Table 1 – Confusion Matrix for Dorian Image Classification

ClassValue	C_11	C_21	C_22	C_23	C_31	Total	U_Accuracy	Kappa
C_11	56	0	2	1	1	60	0.93	0
C_21	0	10	0	0	0	10	1	0
C_22	0	0	10	0	0	10	1	0
C_23	4	0	2	14	5	25	0.56	0
C_31	0	0	0	2	8	10	0.80	0
Total	60	10	14	17	14	115	0	0
P_Accuracy	0.93	1	0.71	0.82	0.57	0	0.85	0
Kappa	0	0	0	0	0	0	0	0.77

It is important to understand how to interpret the results of confusion Matrix.

Each of the classification classes are shown by their class values: C_11 (Damaged_Trees), C_21 (Blue_Tarp_Cover), C_22 (Intact_Roof), C_23 (Destroyed_Building), and C_31 (Open_Roads) Each class is compared between what the human (row) and machine (column) classified. When there is complete agreement between the machine and human, all the values except where the two agree will be 0 and the U_Accuracy (user or human accuracy) and P_Accuracy (producer or machine accuracy) will be 1 indicating complete agreement. In table 10.1, C_21 (Blue_Tarp_Cover) received U_Accuracy and P_Accuracy values of 1.

When the user/human mis-classifies a pixel, this is what is known as a false positive or Type 1 Error of error of commission where a known class is incorrectly classified and should have been classified to another known class. In Table 10.1, this can most clearly be seen with the row C_23 (Destroyed_Building). For example, looking at the row starting with C_23, note how the human classified several pixels as C_23 (Destroyed_Building) but the machine classified those same pixels differently as C_11 (Damaged_Trees, 4 values), C-22 (Intact_Roof , 2 values), and C_31 (Open_Roads, 5 values) thus decreasing the U_Accuracy down to 0.56 (or about 50% agreement). This was likely due to the fact that destroyed buildings look similar in some cases to damaged trees or open roads in terms of grey color values. The total row represents the number of points that should have been identified by the user/human per class.

Conversely, when the machine fails to detect a pixel correctly, this is what is known as a false negative or type 2 error or Error of omission. In Table 10.1 this can most clearly seen with the column C_31 (Open_Roads) where the machine mis-classified open roads (C_31) as C_11 (Damaged trees, 1 value) and C_23 (Destroyed Building, 5 values). Because of this, the P_Accuracy was down to 0.57. This was likely due to the fact that open roads in some cases look very similar to destroy buildings in terms of gray color values. Additionally, the overall Sample size of the open roads class is relatively small (10 samples). The total column represents the number of points that were identified by the machine (or producer) per class.

The Kappa value[11] is the overall level of agreement between the human/user and machine/producer. The kappa value of 0.77 would be considered "substantial agreement" according to Viera and Garrett (2005) (with values > 0.81 being "Almost perfect agreement"). Thus, the overall results indicate that the SVM algorithm performed moderately well in classifying the Dorian image.

[11] See: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3900052/ for further discussion about the Kappa statistic.

Discussion Questions

Now that you have completed this exercise, here are some discussion points:

1. What other classification classes might you have included as per the schema that was loaded In task 3 step one? For example, open grass? Also, look at the bigger image that the image used in this exercise was derived from and provided with the exercise datasets. What other types of classification classes might you find from that bigger image and how might that inform the use of classification algorithms to support a disaster response?
2. The overall accuracy of the classification based on the Kappa value discussed in Task 6 was "substantial agreement" between ground truth data and machine data. Although this was good, it was by no means perfect. How might these types of errors potentially produce problems when using the outputs of this analysis for Disaster Response? For example, if developing damage assessment maps based on machine learning algorithm outputs, what problems might occur with use of these maps in terms of inaccuracy?

Deliverable:

Submit to the lab 4 drop a single MS Word (no PDFs or individual image files) that contains:

1. Task 4 final classification map product.

2. Confusion matrix table created in Task 6 and a brief (50-75 word) discussion of your final results.

3. Answers to the two discussion questions, 50-75 words per answer.

References

VIERA, A. J. & GARRETT, J. M. 2005. Understanding interobserver agreement: the kappa statistic. *Fam Med,* 37**,** 360-363.