Spring 5-2016

Using object-based segmentation to identify invasive species in Onondaga County

Jennifer Lynn Gienau

Follow this and additional works at: http://digitalcommons.esf.edu/honors

Part of the Environmental Engineering Commons, and the Weed Science Commons

Recommended Citation

Using object-based segmentation to identify invasive species in Onondaga County

by

Jennifer Lynn Gienau
Candidate for Bachelor of Science
Environmental Resources Engineering
State University of New York College of Environmental Science and Forestry
With Honors

May 2016

APPROVED

Thesis Project Advisor: ___________________________
Lindi J. Quackenbush, Ph.D.

Second Reader: ___________________________
Eric J. Greenfield, Ph.D.

Honors Director: ___________________________
William M. Shields, Ph.D.

Date: ___________________________
**Table of Contents**

List of Tables .......................................................................................................................... 1

List of Figures .......................................................................................................................... 1

Abstract .................................................................................................................................. 2

Introduction ............................................................................................................................ 3

Materials and Methods ......................................................................................................... 4

  - Study Area .......................................................................................................................... 4
  - Datasets .............................................................................................................................. 4
  - Processing .......................................................................................................................... 5

Results ..................................................................................................................................... 6

Discussion and Conclusions .................................................................................................. 8

Acknowledgements ................................................................................................................ 9

References .............................................................................................................................. 10

Appendix ................................................................................................................................ 11

**List of Tables**

Table 1: Training data for each species. .................................................................................. 11

Table 2: Error matrix for classified Japanese knotweed image. ............................................. 11

Table 3: Error matrix for classified common reed image. ....................................................... 11

**List of Figures**

Figure 1: Species locations and distribution throughout Onondaga County. ....................... 12

Figure 2: Original Japanese knotweed image. ....................................................................... 12

Figure 3: Japanese knotweed objects. .................................................................................... 13

Figure 4: Classified Japanese knotweed. ............................................................................... 13

Figure 5: Original common reed image. .................................................................................. 14

Figure 6: Common reed objects. ............................................................................................ 14

Figure 7: Classified common reed. ......................................................................................... 15
Abstract

The impact invasive species have on native habitats can be devastating as they reduce plant biodiversity, decrease wildlife habitat, and change ecological function of an area. *Lythrum salicaria* (purple loosestrife), *Fallopia japonica* (Japanese knotweed), and *Phragmites australis* (common reed) are three invasive species located in Onondaga County, New York. This study aimed to test the feasibility of identifying the presence of these invasive species using readily available oblique imagery. Imagery used in the study came from Google Earth Street View imagery matched to the coordinates of locations reported to have dense populations of common reed, Japanese knotweed, or purple loosestrife. Images were loaded into ERDAS Imagine 2015 where unsupervised object-based segmentation was performed. ArcGIS was used to determine spectral, spatial, and textural characteristics for each object within the image. MATLAB was used to classify the segmented images using a classification tree based off the spectral bands, length to area ratio of the objects, and average texture of the objects. Results show that Japanese knotweed and common reed classified by these means have an overall accuracy of about 84% and 64%, respectively. Purple loosestrife was not classified during this study due to a lack of imagery corresponding with available data points. To improve this methodology, spatial resolution and capture date of photos should be optimized. Future research should include reference photos with confirmed species identification, higher quality images, and ground validation. This research is an important step towards improving invasive species detection efforts.
Introduction

Invasive species are an ever increasing problem in our world. They alter the evolutionary pathway of native species by competitive exclusion, niche displacement, hybridization, introgression, predation, and possible extinction (Mooney & Cleland, 2001). Three non-native, invasive species found in Onondaga County in New York State are *Lythrum salicaria* (purple loosestrife), *Fallopia japonica* (Japanese knotweed), and *Phragmites australis* (common reed).

By 1829, purple loosestrife seeds were being sold in both American and Canadian nurseries, but it was also likely introduced to North America by ships dumping ballast or by raw wool or sheep imported from Europe contaminated with seeds (Lavoie, 2009). The negative impacts of invasive purple loosestrife include decrease in habitat for some wetland bird species, altering wetland structure and function, crowding out native wetland species, and reduction of wetland plant diversity (Blossey et al. 2001).

Japanese knotweed is a widespread invader of both North America and Europe, which was introduced in the 19th century because of its ornamental qualities. In places it has overrun it is known to alter nutrient cycles and reduce plant species diversity (Cottet et al. 2015).

*Phragmites* as a genera are native to the United States, but the common reed species discussed in this paper is invasive. It was brought to North America by European settlement and is of great concern in wetlands of the North American Great Lakes (Saltonstall, 2002). The biggest concern regarding the invasive phragmites is the loss of species richness followed by the potential for extirpation and loss of biodiversity of native wetland species (Pengra et al. 2007).

Invasive species, such as the three species described above, can have devastating effects on ecosystems. Due to the general fast-spreading and resilient nature of these invasive plants, scientists are tasked with creating effective methods of detecting invasive species and managing the affected areas. One way of efficiently detecting invasive plants is through the use of remotely
sensed data (Pengra et al. 2007). Through the application of remote-sensing focused methods, people can detect large areas affected by invasive species without having to travel to each individual site.

The focus of this study was to determine the utility of street-view imagery to identify the presence of the three species from several locations in Onondaga County. The source of this imagery was Google Earth Street View, which is available to any person with an internet connection.

**Materials and Methods**

**Study Area**

Onondaga County is located in Central New York and was the area of interest in this study. This study area was chosen due to its familiarity and convenient location for the researcher. Oblique imagery of the area was readily available from Google Earth Street View for use in the study. Presence data for the three invasive species study were available from the New York Invasive Species Clearinghouse (http://nyis.info/).

**Datasets**

Invasive species reference data used for the project was obtained from staff at the New York Natural Heritage Program (NYNHP). The dataset contained 2864 points collected across New York State from 2002–2015 that indicated the presence of common reed, Japanese knotweed, or purple loosestrife. It is unknown how the data was collected which may have some effect on the results. For the purpose of this project the data was narrowed to Onondaga County, which reduced the dataset to 20 observations. Each data point had associated information on the plant species, coordinates, observation date, observer, and County.
Coordinates of the invasives were input into Google Earth to download normal color (red-green-blue; RGB) Street View images. Few of the data points had a corresponding Street View image that was usable in the analysis, which further reduced the available dataset. Each data point with a usable image was downloaded and used for processing.

**Processing**

The images from Google Earth were downloaded as JPEGs and imported into ERDAS Imagine 2015 for preliminary analysis. The initial processing performed object-based segmentation for each image in the study. Object-based segmentation groups individual pixels into objects based on different characteristics (i.e. shape).

The object-based segmentation applied the unsupervised Fast Level Set (FLS) image segmentation tool in ERDAS Imagine 2015 to the three-band image from Google Earth. This tool considers the spectral content as well as the texture, size, and shape of segments in merging decisions. Default settings were used for this tool, which included a pixel:segment ratio of 500, relative weights of 0.50 for spectral, texture, size and shape, size limit minimum of 10, and size limit maximum of 100,000 pixels. The raster segmented image generated in Imagine was then brought into ArcGIS for further processing.

The segmented image was converted to a polygon feature class in ArcGIS and then joined to the initial raster image to associate the RGB data from the imagery with each object. Training data was then created for each image, shown in Table 1. The area to length ratio and texture of each object in the image were calculated. Texture calculation involved a few separate steps. First, focal statistics—standard deviation within a 3×3 window—were calculated for the raster image and the output converted to a 32-bit unsigned value. Zonal statistics were then calculated using the mean option with the texture image as the input raster and the FLS polygon
as the input feature (value = gridcode) for the area. This output was then joined to the FLS polygon in order to be used for object classification. The attribute tables for the training data and the polygon FLS image were exported to Excel and then used in MATLAB.

Having segmented the image and determined a range of characteristics for each object, the classification of the objects was performed in MATLAB, where a classification tree was generated using the two sets of training data developed in ArcGIS. Classification was based off the RGB band values, length-to-area ratio of the objects, and texture of each object within the image. The resubstitution error was calculated for each image to determine if predictions made by the classification tree generated were reliable. After the classification tree was created from the training data, it was used to classify the entire image and then those results were exported to an Excel table.

The Excel table generated from the MATLAB output was then imported into ArcGIS and used to generate an error matrix. Accuracy assessment used 100 random points generated in Excel. These 100 points were manually identified as belonging to the target species or not and this label as compared to the classified image to determine accuracy. There were not 100 points present in the error matrix generated due to the fact that some of the randomly generated points fell into the same object in ArcGIS.

**Results**

In total, there were nine Japanese knotweed, seven common reed, and four purple loosestrife locations in Onondaga County from the NYNHP data. Of these, only one Japanese knotweed and one common reed Street View image were available from Google Earth. Of the four purple loosestrife locations, none had an associated Google Earth image. The locations of the data are shown in Figure 1; locations with available imagery are highlighted in yellow in the embedded table. Overall, the observations were spread throughout Onondaga County, but
showed some clustering in certain areas (Figure 1). For example, common reed is likely clustered around wetland areas, but as mentioned previously it is unknown how the data was collected so this can’t be known for sure. Classification results aiming to separate invasive species from the surrounding area were conclusive for the Japanese knotweed image, but could be improved for the common reed image.

Figures 2–4 show the object-based segmentation process seeking to identify Japanese knotweed at the location highlighted in Figure 1. Figure 2 shows the image from Google Earth that was used in ERDAS Imagine 2015, Figure 3 shows the unsupervised FLS object-based segmentation into 1170 objects, and Figure 4 shows the classified image from the classification tree with random points overlaid. The error matrix calculated from this image is shown in Table 2 and determined that the image was classified with an overall 84% accuracy. Out of the 100 points generated for accuracy assessment, only 90 were unique for Japanese knotweed. Based on the low spectral resolution of the imagery used the accuracy levels obtained surpassed expectations.

Figures 5–7 show the object-based segmentation process seeking to identify common reed at the location highlighted in Figure 1. Figure 5 shows the image from Google Earth, Figure 6 shows the unsupervised FLS object-based segmentation into 1546 objects, and Figure 7 shows the classified image from the classification tree with random points overlaid. The error matrix calculated from this image is shown in Table 3 and determined that the image was classified with an overall accuracy of 64%. Out of the 100 points generated for accuracy assessment, only 88 were unique for common reed. The results for this image are much less desirable than those found with Japanese knotweed. This could be due to the low spectral resolution of the image or needing more variables to separate common reed from its background.
Discussion and Conclusions

Based on the classifications performed in this study, object-based image segmentation can be a reliable way to classify invasive species, but the methodology could be greatly improved. The experimentation performed in this study suggested that the same approach did not apply equally well to both common reed and purple loosestrife. Instead of having a general method to identify an invasive species, a species-specific approach should be generated. To develop a specific method would involve determining what characteristics are most important in separating a species from its surroundings. These characteristics could include different measures of texture, spectral reflectance values or reflectance ratios (i.e. green to blue), and spatial measures such as object size or shape. Exploration of additional object metrics and the significance of each metric within the classification were out of the scope of this study, but should be considered in future research.

Other ways to further improve classification techniques of invasive species is to consider image quality, shadows, seasonality of image capture, and data collection methods. Figures 2 and 5 were obtained from Google Earth and are not particularly high quality images. Low quality images with multiple shadows can negatively impact classification results by increasing confusion between classes. One way to mitigate issues related to shadowing is to use band ratios, as suggested above for species-specific approaches (Quackenbush et al. 2000).

The seasonality of capture is also important because during different times of year invasive species, like all vegetation, appear differently. Japanese knotweed, for example, blooms with white flowers in late summer and may be much less recognizable to a computer algorithm when not in bloom. Ideally, all photos used for classification should be captured when specimens are in full bloom so that classification may be easier. However, this was not the case for the images used in this study. Knowing when an image is captured is also important since if an
image is dated it may be hard to know if the invasive species is still currently present at that location.

It is important to know how the data being used for a study is collected. The reason of the clustering in this dataset is unknown. For example, the clustering of common reed in some areas could be as a result of proximity to a wetland, but alternatively the person collecting data could have just collected multiple data points in this one area. In future studies, data collection methods should be clear in order to fully understand the data and determine if clumping is a result of the surrounding area or a result of the collection methods.

Overall, this study has shown that object-based image segmentation has potential for invasive species identification, but should be further researched to improve current methods. Further research is also required to determine the feasibility of using training data from one area to classify an image from another area and conclude the presence of the invasive species. However, if a species-specific method could be developed for even one species this would be one step closer to automating invasive species identification and classification.

**Acknowledgements**

I would like to thank Dr. Lindi Quackenbush of the SUNY-ESF Department of Environmental Resources Engineering for her guidance and patience throughout this entire project. I would also like to thank Brent Kinal of the New York Natural Heritage Program for providing us with the data for the project. Lastly, I would like to thank Dr. Eric Greenfield of the US Forest Service for his valuable input.
References


Appendix

Table 1: Training data for each species.

<table>
<thead>
<tr>
<th>Training Data</th>
<th># of points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confirmed presence</td>
<td>17</td>
</tr>
<tr>
<td>Confirmed no presence</td>
<td>27</td>
</tr>
</tbody>
</table>

Table 2: Error matrix for classified Japanese knotweed image.

<table>
<thead>
<tr>
<th>Actual class</th>
<th>Predicted</th>
<th></th>
<th></th>
<th>Row total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Japanese knotweed</td>
<td>Not</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>19</td>
<td>8</td>
<td></td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>Not</td>
<td>6</td>
<td>57</td>
<td>63</td>
</tr>
<tr>
<td>Column total</td>
<td>25</td>
<td>65</td>
<td></td>
<td>90</td>
</tr>
</tbody>
</table>

Overall accuracy = 84.44%

Producer’s Accuracy (omission error)  
Japanese Knotweed = 76.00%
Not = 87.69%

User’s Accuracy (commission error)  
Japanese Knotweed = 70.37%
Not = 90.48%

Table 3: Error matrix for classified common reed image.

<table>
<thead>
<tr>
<th>Actual class</th>
<th>Predicted</th>
<th></th>
<th></th>
<th>Row total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Common reed</td>
<td>Not</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>4</td>
<td></td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>Not</td>
<td>28</td>
<td>45</td>
<td>73</td>
</tr>
<tr>
<td>Column total</td>
<td>39</td>
<td>49</td>
<td></td>
<td>88</td>
</tr>
</tbody>
</table>

Overall accuracy = 63.64%

Producer’s Accuracy (omission error)  
Common reed = 28.21%
Not = 91.84%

User’s Accuracy (commission error)  
Common reed = 73.33%
Not = 61.64%
Figure 1: Species locations and distribution throughout Onondaga County.

Figure 2: Original Japanese knotweed image.
Figure 3: Japanese knotweed objects.

Figure 4: Classified Japanese knotweed.
Figure 5: Original common reed image.

Figure 6: Common reed objects.
Figure 7: Classified common reed.