NORTHWEST NAZARENE UNIVERSITY

Evaluation of Texture as a Fourth Input of Spatial Context Using Machine Learning to Map Wildland Fire Effects.

THESIS

Submitted to the Department of Mathematics and Computer Science

in partial fulfillment of the requirements

for the degree of

BACHELOR OF ARTS

Jonathan M. Branham

2018

THESIS

Submitted to the Department of Mathematics and Computer Science

in partial fulfillment of the requirements

for the degree of

BACHELOR OF ARTS

By

Jonathan Branham

2018

Evaluation of Texture as an Input of Spatial Context for Machine

Learning Mapping of Wildland Fire Effects.

Author: Jon Branham

Jonathan Branham

Approved: ____________________

Dale Hamilton, Department of Mathematics and Computer Science Faculty Advisor

Approved: ____________________

Dr. Stephen Riley, Department of Theology
Second Reader

Approved: Barry Myers

Barry L. Myers, Ph.D., Chair,
Department of Mathematics & Computer Science

ABSTRACT

Evaluation of Texture as an Input of Spatial Context for Machine Learning Mapping of Wildland Fire Effects.

BRANHAM, JONATHAN (Department of Mathematics and Computer Science), MYERS, DR. BARRY (Department of Mathematics and Computer Science), HAMILTON, DALE (Department of Mathematics and Computer Science).

A variety of machine learning algorithms have been used to map wildland fire effects, but previous attempts to map post-fire effects have been conducted using relatively low-resolution satellite imagery. Small unmanned aircraft systems (sUAS) provide opportunities to acquire imagery with much higher spatial resolution than is possible with satellites or manned aircraft. This effort investigates improvements achievable in the accuracy of post-fire effects mapping with machine learning algorithms that use hyperspatial (sub-decimeter) UAS imagery. Spatial context using a variety of texture metrics were also evaluated to determine the inclusion of spatial context as an additional input to the analytic tools along with the three-color bands. This analysis shows that the addition of texture as an additional fourth input increases classifier accuracy when mapping post-fire effects.

Acknowledgements

I would like to thank my mother for supporting me during the summer that I worked on this project. I don't think I would have made it this far without her support and encouragement throughout college. I also want to thank Dale Hamilton for the opportunity to be a coauthor on an academically published paper. Lastly would like to thank Dr. Myers for his continual feedback and willingness to teach a sometimes-floundering student. I also want to thank the other students that worked on FireMAP and their help with problems that I ran into. Also, all the other students that laid the foundation for this project, with their development of software and implantation of machine learning algorithms.

Table of Contents

Table of Figures

Overview

This project was a study about developing a texture metric as a fourth input along with the Red, Green, and Blue (RGB) bands of an image. The additional input of the most promising texture metric would help improve the classification accuracy of wildland fires, which were classified using a machine learning algorithm called an Support Vector Machine (SVM). The goal of the project was not necessarily to improve the SVM classifier itself, but to discover a texture metric, which when added as a fourth input, would improve the accuracy with which the SVM could classify an image. This would mean that the classifier would be able to understand more about a certain pixel, because it would understand the spatial context of the pixels around it. To accomplish this goal several metrics were tested, and the classification accuracy while using spatial context was examined for increased accuracy.

Background

This project was part of a summer research team at Northwest Nazarene University (NNU) called Fire Monitoring and Assessment Platform (FireMAP), whose goal is to help wildlife ecologists gather useful data on post fire effects. They do this using sUAV (Unmanned Aerial Vehicles) that acquire imagery of wildland fires with higher spatial resolution than previously used satellite imagery. The image is then input to the (the SVM) where the algorithm goes through and classifies pixels. This finally returns a classified image. The classifier needs to be trained with user labeled data, called training data. The classifications used a heirarical classification, burn extent, then biomass consumption (black/white ash).

The main point of the project was to increased accuracy in the classification of the image, which was needed to produce reliable burn extent and severity mapping products. Another FireMAP team member, Zach Garner, wrote a tool that generated the texture inputs. This program used several second order metric formulas that Haralick defines (Haralick, 1973). The metrics tested were First Order Entropy, Second Order Entropy, Homogeneity, Contrast, and Energy. Next it had to be determined which if any metric added the most value to the classifier.

Terms

The preparation for this project involved learning several terms and understand the programs to test and output useful imagery. Simply, there was a need to define the availability, usefulness, and the importance of the tools before figuring out how they worked. In this case, effectively completing this project consisted of using previously written programs, terms involved in graphic interfaces, and terms used in ecological studies and data revolving around wildland fire imagery. This resulted in study of the effects of wildland fires: what programs had been used, what programs that would make work go smoother, and a more complete understanding of how imagery data is collected and stored.

There are some terms that need to be defined to have a clear understanding of the project. The first of these terms is GLCM (gray-level co-occurrence matrix). GLCM as, "a gray-level co-occurrence matrix (GLCM) which is used to calculate how many occurrences of each combination of pixel values occurred for each pixel within the neighborhood…The texture values for each pixel are stored in a single band gray scale image." (Hamilton, 2017, 4). This is important because these GLCM's give us the spatial context of a pixel based on pixel distance. This pixel offset, and neighborhood size can be varied by the user to find the optimal parameters.

The fires burned several miles which meant that the higher resolution imagery that we collected had to be stitched together. In other words we had multiple images that we combined into one. “Acquisition of imagery for a burn area with the purpose of mapping wildland fire effects is commonly accomplished by mosaicking all the images taken during one or more sUAS flights in order to create a single georeferenced orthomosaic of the entire scene” (Hamilton, 2017, 3). These orthosmosaics contained larger amounts of data which meant, “The very large number of pixels in hyperspatial imagery requires the utmost care in selecting algorithms and metrics which extract fire effects information” (Hamilton, 2017, 3).

Another important term is texture metrics. “Haralick [7] defined 14 measures of texture for image processing from which spatial context has been measured for a variety of related image processing applications” (Hamilton, 2017,4). The metrics used are First Order Entropy, Second Order Entropy, Homogeneity, Contrast, and Energy.

$$Entropy\ (1st\ Order) = -\sum_{i,j} p(i) log_2(p(i)) \qquad (1)$$

$$Contrast = \sum_{i,j} (i-j)^2 p(i,j) \qquad (2)$$

$$Energy = \sum_{i,j} p(i,j)^2 \qquad (3)$$

$$Entropy\ (2nd\ Order) = -\sum_{i,j} p(i,j) log_2(p(i,j)) \qquad (4)$$

$$Homogeneity = \sum_{i,j} \frac{p(i,j)}{1+|i-j|} \qquad (5)$$

These five-texture metrics all yielded different results which, meant the question is what metric returns the most value to the SVM when added as a classification input along with the red, green and blue bands. All of the metrics except for First Order Entropy used GLCM to calculate the texture metrics. The most straightforward way to measure this was to calculate the information gain of each metric. Information gain is the measurement of usefulness of a certain data set in comparison to the other data sets. To determine the value, information gain was used as a measurement for each of the metrics. So, in the case of the project the data sets are the Red, Green, Blue, and texture bands. To find the information gain an Iterative Dichotomiser (ID3) was used, which is a machine learning algorithm that uses a decision tree to classify tuples (Russel, 2010, 758).

Basic knowledge of wildland fire ecology is required to understand why this project was implemented. When looking at the burned area post fire, the type of ash is considered to determine severity. Hamilton explains this, "The distinct spectral signatures between white and black ash has been shown to enable successful classification of burn severity, separating pixels with low fuel consumption (black ash) from high fuel consumption (white ash)" (Hamilton, 2017, 2). The overall goal for FireMAP is to have an accurate classifier, which can pick up white ash, black ash, and unburned parts of the terrain. This is accomplished through, "Individual pixels are classified by ash type where black ash is indicative of incomplete vegetation consumption while white ash correlates significantly to more complete vegetation consumption" (Hamilton, 2017, 2). Previously it was mentioned that there was a need for increase in accuracy. This was important because, "Increased mapping accuracy will provide actionable knowledge resulting in improved ecosystem resilience and management decisions" (Hamilton, 2017, 1).

Implementation

There were several steps to this project and the implementation required the use of several programs to be completed. The first step in this process was to determine the neighborhood size and pixel distances that we would be testing to find the optimal GLCM. To determine which neighborhood size worked best it was decided to test neighborhood sizes 3*3, 7*7, 15*15, 25*25, 35*35, 45*45, and 55*55. We generated all the GLCMs for the pixel distances of 2, 5, 10, 15, 20, 25, 30, 35 and 40. This totaled up to 63 tests for just a single second order metric. With four other second order metrics having 63 tests each, and first order entropy contributing 7, there were 259 different tests that were ran, to establish which neighborhood size and pixel distance was the most beneficial. Those numbers were just for a single fire, but in the end five fires were tested for information gain, resulting in 1295 texture files that information gain was calculated on.

Instead of classifying based off each of the neighborhood sizes and their pixel offset, which would have taken substantially longer, information gain was used to ascertain the optimal neighborhood size and pixel distance. The ID3 was used in a batch file to, "build a decision tree and report the information gain of each variable from the red, green and blue bands from the color image as well as texture" (Hamilton, 2017, 5). The information gain gave us which band was most used by the classifier to determine category. Therefore, if the information gain was higher or close to any of the other bands it meant that texture metric added value to the classification. This is because the more context that the classifier has, the more accurately it will classify.

One the input parameters for the ID3 was the training data. The training data was gathered using a web application developed by Greg Smith who was also a member of the

FireMAP research team. The application allowed the user selection of certain areas in an image and apply a label. The program would give an excel file of the pixel coordinates in the selected areas which would be read into the SVM. The SVM algorithm would then classify using the training data as a reference for what to label. The SVM algorithm also fed in the GLCM of pixel distance 10 and neighborhood size of 45 as a fourth input. After this it would classify first burned and unburned, then white ash and black ash. The image would then be run through a denoising program that would that implemented image processing morphological algorithms to remove sub-object sized clusters of pixels as well as smooth pixel cluster boundaries.

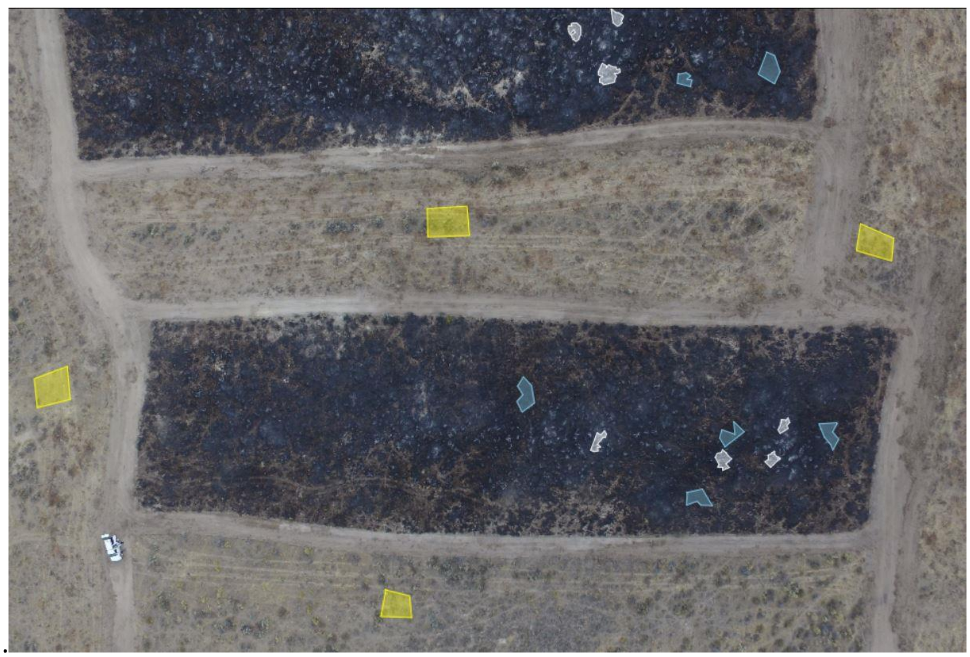
.

Figure 1 – Training Data for Reynolds Creek

A batch file was run that sent multiple files through the ID3 and the ID3 would output a text file with the tree. The information gain was the interesting part of the ID3. They are shown in Figure 1 after the G(D): Blue, Green, Red, and Entropy 1.

```
Answers: white_ash black_ash surface
I(D) = 1.5085  G(D) = 0.820706 0.988954 1.05631 0.076605
```

Figure 2 – Text File

Each fire was recorded as an orthomosaic which comprises a large amount of data. "These hyperspatial images contain a very large amount of data. For example, an orthomosaic generated from multiple flights over Northwest Nazarene University, which has a campus covering 40 hectares (100 acres) in Nampa, Idaho resulted in an image consisting of two billion pixels" (Hamilton, 2017, 3). The reason four separate tests were taken, on one fire, was to increase the reliability of the data. So, to find the mean of the information gain a MATLAB program was written, on the compilation of 20 excel spreadsheets. The program took the information gain of all the data collected and returned the mean in an excel spreadsheet. It also returned the standard deviation which was looked at for any anomalies.

Neighborhood size

Pixel Offset

0	3	7	15	25	35	45	55
2	0.18344	0.27994	0.41361	0.50953	0.5558	0.58626	0.60147
5	0.26475	0.34795	0.44309	0.52432	0.56417	0.58956	0.59915
10	0.29897	0.40769	0.4877	0.54126	0.57754	0.59691	0.60026
15	0.30473	0.41623	0.51331	0.55501	0.58325	0.59932	0.60251
20	0.30301	0.41704	0.52509	0.57065	0.59166	0.60207	0.60788
25	0.30068	0.4174	0.52748	0.5831	0.5964	0.60416	0.61204
30	0.29351	0.40794	0.52371	0.58488	0.60061	0.60686	0.61215
35	0.28505	0.40119	0.51613	0.57927	0.60207	0.60822	0.6129
40	0.27781	0.39287	0.50729	0.57209	0.59988	0.60984	0.61235

Figure 3 – Entropy 2 Mean

Figure 2 includes the table of means for Second Order Entropy. The twenty spread sheets containing information gain, were used to find the optimal neighborhood size and pixel distance. The first inclination is to pick the average with the highest information gain. This would lead to the picking of neighborhood size 55 and pixel distance 35. So instead, "The optimal neighborhood size for first order entropy was identified at the point of diminishing information gain as the neighborhood size was varied, identifying the point where the information gain gradient started to significantly reduce as neighborhood size continued to increase" (Hamilton, 2017, 6). So instead of choosing the highest point, the part where it slopes off is the sweet spot. After finding the sweet spots from all the averages of each metric, prime parameters were compared to each other. The most promising parameters are shown down in table 1. From this it was determined that Second Order Entropy yielded the highest information gain, followed by First Order Entropy, and Energy. Figure 3 shows a 3-D graph of the Entropy 2, printed from the MATLAB program, that shows the steep gains that you get until about neighborhood size of 35.

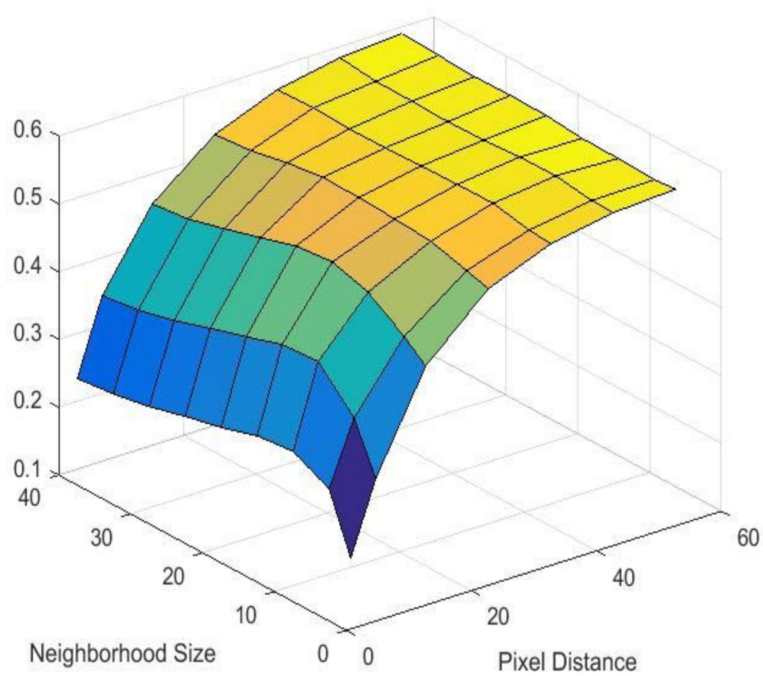

Figure 4 – Entropy 2 Average Graph

After this phase validation data is gathered, which was used to test how accurate the classifier was. “Validation data sets for each of the images were selected as regions of pixels within the image, then the pixels from each validation data set were run through the SVM, assessing the accuracy of the color bands as inputs as opposed to the inclusion of each of the texture metrics with the associated optimal parameters. Accuracy for each validation data set was calculated, determining the percentage of validation pixels the SVM classified the same as were labeled by the user. Validation data labeling was based on visual observation of the image by the user, supplemented with ground observations recorded during image acquisition flights with the sUAS ”(Hamilton, 2017, 7). Using these accuracy readings, we used the T-tests to determine over the course of several fires what the probability was that the inclusion of texture increased classification accuracy.

Results

There were two important questions that needed to be answered in this project. What the optimal neighborhood size and pixel distance of the GLCMs is, and would a texture metric, if any, increase the classification accuracy of the SVM. These two questions were answered in the tables shown below.

The optimal neighborhood size was different for each of the metrics therefore, a table was constructed to find the optimal spot for each of them. Table 1 is comprised of the different optimal values. In Table 1, First Order Entropy does not have a pixel distance, because it was generated without a pixel distance just the Neighborhood size.

Table 1: Optimal texture metric parameters.

Texture Metric	Info Gain	Nhood Size	Pixel Dist
Contrast	0.53692	35	10
Energy	0.57245	35	15
Entropy (1st Order)	0.58227	45	-
Entropy(2nd Order)	0.59691	45	10
Homogeneity	0.55731	35	15
Red	0.79268	-	-
Green	0.79583	-	-
Blue	0.64862	-	-

The accuracy of the classifier is shown in Table 2 which shows that the highest accuracy percentage is Second Order Entropy. The percentage that Second Order Entropy increases the accuracy average 2.69% across the burns evaluated. All the other metrics decrease accuracy when added to the color bands when classifying an image.

Table 2: Mean classification accuracy.

Texture Metric	Burn Extent	Ash Type
Energy	86.91	78.53
Entropy(1st Order)	84.00	75.68
Entropy(2nd Order)	94.40	83.79
Color only	91.71	77.33

These accuracy percentages were used with T-tests to find, "The statistical significance of increased accuracy across the validation sets for the burn images was established using one tailed paired T-tests. The null hypothesis is that the addition of texture as a fourth input along with the color bands does not improve accuracy. By contrast, the alternate hypothesis is that adding texture as a fourth input along with color will increase classifier accuracy. In order to apply the T-test, the accuracy of the classification was taken using just the three-color bands and then again with texture added as a fourth input. The significance required to reject the null hypothesis in favor of the alternate hypothesis is 0.05 which gives it 95 percent certainty to reject the null hypothesis in favor of the alternate hypothesis" (Hamilton, page 8). Summed up this meant that we were 95% certain that Entropy 2 as a fourth input improved the accuracy of the classifier.

Conclusion

I enjoyed working on this project because I learned a lot about my field and participated in groundbreaking research dealing with machine learning. It was very intriguing to learn about machine learning and the implications that it holds for the future. The one that impacted my work most significantly the use and understanding of the ID3. This project also allowed me the opportunity to put my name on an academically published paper.

This process showed me how much work was expected when submitting an academically published paper. The thesis gets worked over through vigorous testing that happen when proving the stated hypothesis. This project also taught me the importance of clear explanations and precise speech.

I also really enjoyed the people that I was working with. Drs. Hamilton and Myers are very inclusive when working on projects and trust you with very important parts of the project. The other students that worked with me on the project were very helpful every time I had questions. Zach Garner was the one that designed a lot of the code that I was using and was very instrumental in a lot of the tests. Ryan Pacheco was also a big help when it came to running the accuracy tests. Over all I learned a lot about team work from this project that will help me in the corporate world.

References

Hamilton, D., Myers, B., Dr., & Branham, J. (2017). *Evaluation of Texture as an Input of Spatial Context for Learning Mapping of Wildland Fire Effects* (Vol. 8). SIPIJ. October 2017

R. Haralick and K. Shanmugam. 1973. Textural features for image classification. IEEE transactions on systems, man, and cybernetics 610-621.

S. Russell and P. Norvig. 2010. Artificial intelligence: A modern approach. Prentice Hall, Upper Saddle River, NJ.