EVALUATION OF TEXTURE AS AN INPUT OF SPATIAL CONTEXT FOR MACHINE LEARNING MAPPING OF WILDLAND FIRE EFFECTS

Dale Hamilton\textsuperscript{1}, Barry Myers\textsuperscript{2}, and Jonathan Branham

\textsuperscript{1}Assistant Professor of Computer Science
Northwest Nazarene University, Nampa, Idaho, USA
\textsuperscript{2}Professor of Computer Science
Northwest Nazarene University, Nampa, Idaho USA

Abstract

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) drone imagery. Spatial context using a variety of texture metrics were also evaluated in order 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.

Keywords

Hyperspatial Imagery, Small Unmanned Aircraft Systems, Texture, Support Vector Machine, Wildland Fire

1. INTRODUCTION

This study examines the mapping of wildland fire severity from hyperspatial sUAS imagery using computer vision and machine learning. This analysis examines whether the inclusion of spatial context with color imagery can increase accuracy when mapping post-fire effects with a Support Vector Machine (SVM). Increased mapping accuracy will provide actionable knowledge resulting in improved ecosystem resilience and management decisions.

Fire consumes millions of acres of American wildlands each year, with suppression costs approaching $2 billion annually [15]. High intensity wildland fires contribute to post fire erosion, soil loss, flooding events and loss of timber resources. This results in negative impacts on wildlife habitat, ecosystem resilience, infrastructure, and recreational opportunities. In order to ameliorate these issues, remotely sensed imagery is commonly collected to assist in assessing the impact of the fire on the ecosystem.

Current methods for acquiring the imagery used for assessing fire effects rely on satellites which provide problematically low resolution. Images obtained from, Landsat, for example, have a temporal resolution of 16 days and a spatial resolution of 30 meters [14]. Manned aircraft can also be utilized for acquiring imagery, but are expensive, have a low spatial resolution, and are often unavailable, especially during fire season when manned aerial resources are devoted to suppressing active fires. Additionally, due to lack of resources, much of the body of fire history contained in fire atlases omits the spatial extent of small and moderately sized fires [13]. Accurate historical record of fire history is necessary in order to determine departure of current fire frequency from historic fire frequency, a key metric for determining ecosystem resilience.

DOI: 10.5121/sipij.2017.8501
Omission of these fires is problematic because small and moderate fires are the most ecologically diverse, adversely affecting the accuracy of fire frequency departure metrics, hindering their ability to accurately reflect ecosystem resilience [5]. Therefore, it is essential to develop novel acquisition methods which allow managers to acquire and analyze imagery with higher spatial resolution.

New advances in sUAS capabilities can enable the acquisition of imagery with a spatial resolution of centimeters and temporal resolution of minutes [11]. High-resolution imagery results in a huge increase in the quantity of data associated with a scene.

In order to fully leverage this additional data, it is necessary to develop analytic tools which identify the extent of the burned area within the image. 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 [9]. Spatial density of white ash can be considered a quantitative measure per unit area of vegetation consumption [20].

Machine learning analytics have been developed to facilitate pixel classification by ash type or vegetation structure, allowing the discrimination between black ash, white ash, crown vegetation and surface vegetation. Utilizing these classes, analytic tools can interpret the scene by relying on relationships between these classes. Classification of burned area extent has been achieved by exploiting the spectral separability between burned organic material (black & white ash) and vegetation [12]. 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) [9]. In forested biomes, low severity fires are identifiable by detecting patches of unburned vegetation within the extent of the fire. If a patch is comprised only of tree crown(s), the analysis can infer the vegetation is a tree which the fire passed under and classify the pixels as low intensity surface fire [19]. If the patch of vegetation contains herbaceous or brush species, then the patch is an unburned island within the burned area and should be classified as unburned.

2. PREVIOUS APPLICATION OF TEXTURE AS A MACHINE LEARNING INPUT FOR MAPPING POST-FIRE EFFECTS

Many projects have used a variety of machine learning algorithms mapping wildland fire extent using low-resolution satellite imagery. Some of the classifiers rely on pixel-based classification with the classifier only considering the band values for the pixel being classified. Other approaches use a variety of methods for considering the spatial context of a pixel while the classifier attempts to identify whether a particular pixel burned.

Gitas [3] segmented 1.1 kilometer (km) resolution imagery from the Advanced Very High Resolution Radiometer (AVHRR) into objects using eCognition. Fuzzy sets with membership functions were created for burned and unburned objects based both on spectral and shape information as well as relation to neighboring objects. The very low spatial resolution of greater than one km prohibits the ability to identify objects that are smaller than 100 hectares. This omission would render over 15 percent of burned area across the United States (US) undetectable due to fires under a square kilometer being sub-object in size [5].

Brewer’s [2] comparison of artificial neural network (ANN) and k-Nearest Neighbor (kNN) classifiers incorporated the spatial context of neighboring pixels in mapping burn extent. Spatial context was achieved by including the values of the 12 closest neighboring pixels around the pixel being classified. The seven Landsat bands of post-fire imagery with 30-meter resolution were compared as inputs against a combination of 11 bands from pre-fire and post-fire Landsat images. Ground-based reference data were collected from a set of reference locations both within and exterior to the burn perimeter. Image pixels corresponding to these reference points were divided between training the classifier and validating classification results. While better results
were achieved with the inclusion of pre-fire imagery, the classification only using post-fire imagery was able to map burn extent with over 80 percent accuracy. While this approach does consider spatial context, using values from seven bands for the 12 neighboring pixels in addition to the pixel being classified increases the number of inputs to either the ANN or the kNN to 91 total inputs. This high dimensional solution will increase the size and complexity of the feature space consisting of the bands for each of the considered pixels. This high dimensional data is likely to contain irrelevant or redundant inputs, causing the curse of dimensionality. Consequently, patterns are less well defined in the data, obfuscating their detection with machine learning classifiers [18].

3. CONSIDERATIONS FOR HIGH RESOLUTION sUAS IMAGERY

These approaches show a number of issues which must be addressed while developing methods, analytic tools and metrics for mapping wildland fire effects with much higher resolution than is currently available from the current generation of satellites. A sUAS with a 12 megapixel (MP) camera flown at 120 meters above ground level (AGL) will have a spatial resolution of about six centimeters resulting in imagery with such high resolution that an object like a bush or tree will be represented in an image as a set of contiguous pixels. An example image (Figure 1a) of a portion of a burn in the Owyhee Mountains in southern Idaho, USA was acquired with a sUAS at 120 meters AGL. The black rectangles in the image are burned areas, which are surrounded by fire lines dug by a bulldozer. The unburned vegetation consists primarily of grass, Wyoming Big Sage and Rabbit Brush. The scene contains two Junipers. The white object near the lower left corner of the lower burn is a Chevrolet Suburban.

In Landsat imagery which has 30 meter resolution, that same tree will be lumped together with everything else (other trees, bushes, grass, bare dirt) that is in the 30 by 30 meter square represented by a single pixel, resulting in a mixed pixel. The example image shown in Figure 1b represents with 30 meter resolution the same scene captured in Figure 1a.

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. 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. The very large number of pixels in hyperspatial imagery requires the utmost care in selecting algorithms and metrics which extract fire effects information. Special consideration must be given to what algorithms and inputs will provide the most accuracy. Additional consideration is recognizing that algorithm efficiency is a critical factor to be considered in order to ensure that derived mapping products are available to users within a reasonable amount of time. A common factor which influences the efficiency of machine learning algorithms is the dimensionality of the inputs, which is reflected by how many inputs are provided to the classifying algorithm.

When mapping burn severity from hyperspatial sUAS imagery, all the inputs are spatial in nature. Color imagery is consumed by the machine learning classifiers as three inputs: the pixel values from the red, green and blue bands. When considering a single pixel either as training data which has been labeled with a class or as unlabeled data for which the classifier needs to determine the class, the classifier only considers the band values from that particular pixel. This resulting pixel-based classification does not consider the relationship of that pixel to any of the neighboring pixels. Brewer [2] showed that improved accuracy can be achieved by providing the pixel values of neighboring pixels as an input to the classifier. Their approach was to provide the spatial context of the neighboring pixels by providing the band values of each of the neighboring pixels as separate inputs to the classifier, increasing the dimensionality of the input data from three
inputs to 39. This added dimensionality will significantly degrade the efficiency of classifiers. For example, increasing the input dimensionality of an ANN by a factor of $n$ results in an increase of the computational complexity of the backpropagation of the ANN by a factor of $n^2$. As a result, it is necessary to find a way to provide spatial context with as few additional inputs as possible in order to improve the classification accuracy without degrading temporal efficiency.

Figure 1. (a) Burn imagery acquired by a sUAS at 120 meters AGL. (b) Same scene resampled to 30 meter resolution.

4. USING TEXTURE AS A MEASURE OF SPATIAL CONTEXT

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. These texture measures have been used in a wide variety of uses ranging from vegetation structure [21] to land-use variation [8]. Texture measures have also been used as an input for image classification [19]. Of Haralick’s texture metrics, we investigated the utility of first order Entropy as well as second order Contrast, Entropy, Energy (also known as Angular Second Moment) and Homogeneity.

Each of these metrics are calculated for a pixel based on a neighborhood of a specified size from a grayscale copy of the image. Second order metrics are calculated from 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. When calculating GLCM, the distance between pixels is specified. The texture values for each pixel are stored in a single band gray scale image. The metrics defined by Haralick [1973], which were evaluated as part of this study are defined as:
\[ \text{Entropy (1st Order)} = - \sum_{i,j} p(i) \log_2(p(i)) \quad (1) \]

\[ \text{Contrast} = \sum_{i,j} (i-j)^2 p(i,j) \quad (2) \]

\[ \text{Energy} = \sum_{i,j} p(i,j)^2 \quad (3) \]

\[ \text{Entropy (2nd Order)} = - \sum_{i,j} p(i,j) \log_2(p(i,j)) \quad (4) \]

\[ \text{Homogeneity} = \sum_{i,j} \frac{p(i,j)}{1 + |i-j|} \quad (5) \]

Where \( i \) and \( j \) are possible values of pixels within the neighborhood surrounding the pixel being evaluated. The probability of finding pixels with a value of \( i \) in the pixel neighborhood is represented as \( p(i) \). The probability of pairs of pixels \( i \) and \( j \) being found in the pixel neighborhood while building the GLCM for the pixel of interest is represented as \( p(i,j) \). The resulting texture images can be used as input along with the associated image as inputs to a machine learning classifier.

5. MEASURING AND EVALUATING THE EFFECTIVENESS OF A TEXTURE METRIC AND PARAMETERS

An Iterative Dichotomiser (ID3) [17] was implemented 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. Information gain facilitated the identification of the most effective texture metric, neighborhood size and GLCM pixel distance for deriving texture for machine learning. By reporting on information gain, it was possible to observe the strength of an input’s ability to accurately split the training data based on the user designated labels as evidenced by the information content of the training data in relation to that input [6]. In order to train the ID3, training regions were designated for black ash, white ash and unburned vegetation on imagery from multiple rangeland fires. An example set of training regions associated with a burn image are shown in Figure 2.

![Figure 2. Training regions from a prescribed burn image. Training regions were used for calculating information gain to identify optimal texture measures. White regions are white ash, blue regions are black ash and yellow regions are unburned vegetation.](image-url)
5.1. Assessment of Texture Parameters using Information Gain

For the purpose of this analysis, texture files were generated for each texture metric using square neighborhoods of size 3, 7, 15, 25, 35, 45 and 55. The pixel distance for calculating the second order metric GLCM was also varied, creating a set of texture files for each metric and neighborhood size, varying the GLCM pixel offsets with values of 2, 5, 10, 15, 20, 25, 30, 35 and 40. Information gain was calculated from training sets on imagery from four burns, averaging information gain across the training sets to identify the texture metric, neighborhood size and pixel offset which has the optimal information gain for use as a fourth input to supplement color imagery as machine learning inputs. 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 as shown in Figure 3. While neighborhood size affects information gain, pixel offset is not used for calculation of first order metrics. Pixel distance is used for calculation of the GLCM, which is only used in the calculation of the second order metrics.

![Figure 3: First order entropy information gain by neighborhood size. Information gain of red, green and blue image bands are also shown.](image)

The optimal neighborhood size and pixel distance for the second order texture metrics were identified by the point of diminishing information gain as both neighborhood size and pixel distance were varied. The information gain of the optimal neighborhood size and GLCM pixel distance averaged over the training sets is shown in Table 1. For comparison, the information gain of the three color bands averaged over the training sets is also included.

<table>
<thead>
<tr>
<th>Texture Metric</th>
<th>Info Gain</th>
<th>Nhood Size</th>
<th>Pixel Dist</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contrast</td>
<td>0.53692</td>
<td>35</td>
<td>10</td>
</tr>
<tr>
<td>Energy</td>
<td>0.57245</td>
<td>35</td>
<td>15</td>
</tr>
<tr>
<td>Entropy (1st Order)</td>
<td>0.58227</td>
<td>45</td>
<td>-</td>
</tr>
<tr>
<td>Entropy(2nd Order)</td>
<td>0.59691</td>
<td>45</td>
<td>10</td>
</tr>
<tr>
<td>Homogeneity</td>
<td>0.55731</td>
<td>35</td>
<td>15</td>
</tr>
<tr>
<td>Red</td>
<td>0.79268</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Green</td>
<td>0.79583</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Blue</td>
<td>0.64862</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

The texture metric with the highest point of diminishing information gain is Second Order Entropy with an information gain of 0.59691 at a neighborhood size 45 pixels square and pixel
distance of 10. The Second Order Entropy information gain calculated for all the considered parameters is shown in Figure 4. Information gain for the remaining metrics was calculated from optimal parameters that were slightly less than Second Order Entropy.

![Figure 4: Information Gain graphed for Entropy 2, varying neighborhood size and GLCM pixel distance. Optimal neighborhood size of 45 by 45 and pixel distance of 10 is indicated where Information Gain of 0.59691.](image)

Based on the results of this analysis, each of the texture metrics would be advantageous if used as an input to a machine learning classifier. Each of the texture metrics had information gain that was nearly as high as the blue or green bands in the color image.

### 5.2. Evaluation of Optimal Texture on Machine Learning Accuracy

Further evaluation of the value of the optimal texture (metric, neighborhood size and GLCM pixel offset) as a machine learning input was accomplished by assessing the accuracy of the output classification of a Support Vector Machine (SVM) machine learning classifier. Accuracy is defined as the number of samples correctly predicted by a classifier divided by the total number of samples [6]. To assess accuracy, the SVMs were trained on the same set of images upon which texture information gain was calculated. The SVM was trained using only the three color spectra, then trained again with each of the texture metrics (with associated optimal parameters) as a fourth input in addition to the color spectra. The SVM first classified the image into burned and unburned pixels. The image region classified as burned was then hierarchically classified into white ash and black ash classes; followed by classifying the unburned regions of the image into canopy and surface vegetation. Based on the spectroscopy study by Hamilton [in press], each of these classes are spectrally separable in the visible spectra. Consequently, the SVM did not apply a kernel to convolve the data into a higher dimensional decision space. This assumption was supported by initial tests which found that running the SVM with the Radial Bias Function, Chi2 and Histogram Intersection kernels resulted in degraded image classifications. Figure 5 shows a classified output from the image in Figure 1a, recording the unburned, black ash and white ash pixels as classified by the SVM, merged into a single classified.

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. Accuracy was calculated as number of correctly predicted validation pixels divided by the total number of pixels in the validation data set multiplied by 100. In order to obtain a more complete assessment of the
accuracy of a set of classifier inputs, accuracy was evaluated based first on burn extent (ash vs unburned pixels) followed by assessment of ash type (black vs white) accuracy and vegetation structure (surface vs canopy) accuracy.

![Figure 5](image)

**Figure 5.** Classified output showing unburned, black ash and white ash pixels classified from Figure 1a. Unburned are colored black, black ash is colored grey and white ash is colored white.

Classification accuracy was averaged for each set of inputs (color versus color and texture) across multiple validation sets, then multiplied by 100. The resulting Mean Classification Accuracy are listed in Table 2 for the metrics which had the most information gain when classification accuracy was averaged across all the images included in the suite of post-fire images evaluated.

<table>
<thead>
<tr>
<th>Texture Metric</th>
<th>Burn Extent</th>
<th>Ash Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy</td>
<td>86.91</td>
<td>78.53</td>
</tr>
<tr>
<td>Entropy(1st Order)</td>
<td>84.00</td>
<td>75.68</td>
</tr>
<tr>
<td>Entropy(2nd Order)</td>
<td>94.40</td>
<td>83.79</td>
</tr>
<tr>
<td>Color only</td>
<td>91.71</td>
<td>77.33</td>
</tr>
</tbody>
</table>

Among the textures tested, Second Order Entropy had the largest increases in average accuracy with an increase of 2.69 percentage points for burn extent as well as well as an increase of 6.45 percentage points for ash type.

### 5.3. Statistical Significance of Accuracy Results

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 level that the t-test passed is 0.05 which gives it 95 percent certainty to reject the null hypothesis in favor of the alternate hypothesis.

The burn extent accuracy tests with Second Order Entropy (average increase of 2.69) rejected the null hypothesis with a P-value of .042. Likewise, the Second Order Entropy ash type accuracy
tests (average increase of 6.45) rejected the null hypothesis with a P-value of .0094. In both cases, the null hypothesis was rejected, supporting the alternate hypothesis which shows that both textures are shown to give a measurable increase in accuracy between the associated classes.

6. CONCLUSION

Each of the texture metrics were found to contain nearly as much information gain as the color image bands. Consequently, this study shows that there is value in including texture as an input to machine learning classifiers. Representing spatial context as a single input will greatly increase temporal performance of machine learning classifiers when mapping wildland fire burn extent. This will enable land managers to obtain higher accuracy burn severity maps expeditiously and economically, offering dramatic improvements over status quo imagery obtained from satellites or manned aircraft. This conclusion is supported by our assessment of classifier accuracy, which validated the accuracy of a burn classifier trained with color imagery against a burn classifier trained with the color bands as well as Second Order Entropy with neighborhood size 45 pixels square and GLCM pixel distance of 10.

7. FUTURE WORK

Additional evaluation will need to be conducted on imagery from additional ecosystem types, especially with regards to evaluating scenes with higher tree canopy cover. A collaborative relationship has been established with the Boise National Forest in southern Idaho, USA to conduct post-suppression flights over burns. To acquire additional post-fire imagery which will assist in development and calibration of our wildland fire burn severity mapping analytics. This analysis was completed using aerial imagery of burned areas using a sUAS which resulted in a spatial resolution of six centimeters. Additional analysis can be done evaluating the increase in accuracy using hyperspatial imagery as opposed to lower spatial resolution resulting from imagery acquired with other means.

The machine learning classifier used for this analysis was an implementation of a SVM. Evaluation of additional machine learning algorithms, assessing accuracy for classifying burn imagery would be very beneficial. In addition to SVM, this research effort has already done preliminary work with other supervised classification algorithms including k-Nearest Neighbor, Artificial Neural Network and the Iterative Dichotomizer 3 Decision Tree algorithms.

Additional validation of classifier accuracy can be accomplished by using post-fire plot assessments as validation data for accuracy assessment. This ground truthing data can also be used for assessing the accuracy of mapping burn severity from hyperspatial imagery as compared to burn severity mapping using Normalized Burn Ratio from Landsat imagery with 30-meter imagery.

ACKNOWLEDGMENTS

We would like to express our thanks to our mentor Gregory Donohoe for his guidance and encouragement. We acknowledge Zach Garner, Jonathan Hamilton, Nicholas Hamilton, Liam Johnston, Ryan Pacheco and Greg Smith for assisting in the development of tools as well as conducting this analysis. We would also acknowledge the National Aeronautic and Space Administration, Idaho Space Grant Consortium and the project was supported by an Institutional Development Award (IDeA) from the National Institute of General Medical Sciences of the National Institutes of Health under Grant #P20GM103408. We would also like to thank Jerry Hoagland, the Boise National Forest and the Boise District of the Bureau of Land Management for allowing sUAS image acquisition flights over burns on their property.
REFERENCES


**AUTHORS**

**Dale Hamilton** received an M.S. in Computer Science from the University of Montana in 1997 and is currently pursuing a Ph.D in Computer Science from the University of Idaho. Dale joined the faculty at Northwest Nazarene University in 2013 where he is an Assistant Professor of Computer Science. In addition to teaching a variety of Computer Science courses, Dale is the Primary Investigator on NNU’s NASA funded Fire Monitoring and Assessment Platform (FireMAP) project, using machine learning and computer vision to map wildland fire extent and severity from imagery acquired with unmanned aircraft system.

**Barry Myers** received his Ph.D. in Information Science from the University of North Texas in 2003. Dr. Myers currently serves as Professor of Computer Science and the Chair of the Mathematics and Computer Science Department at Northwest Nazarene University (NNU). Prior work experience includes systems programming at AT&T Bell Laboratories and systems analysis at Shell Oil Company. His research interests include systems development, bioinformatics, and ecoinformatics.

**Jonathan Branham** is pursuing a B.A. in Computer Science at Northwest Nazarene University, which he anticipates finishing in May 2018.