A CONTEXTUAL CLASSIFICATION APPROACH FOR FOREST LAND COVER MAPPING USING HIGH SPATIAL RESOLUTION MULTISPECTRAL SATELLITE IMAGERY – A CASE STUDY IN LAKE TAHOE, CALIFORNIA

BY

YI ZOU

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Geography
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2015

Urbana, Illinois

Adviser:

Professor Jonathan A. Greenberg

ABSTRACT

Maps of classified surface features are a key output from remote sensing. Conventional methods of pixel-based classification label each pixel independently by considering only a pixel's spectral properties. While these purely spectral-based techniques may be applicable to many medium and coarse-scale remote sensing analyses, they may become less accurate when applied to high spatial resolution imagery in which the pixels are smaller than the objects to be classified. At this scale, there is a higher intra-class spectral heterogeneity. Detailed forest and vegetation classification is extremely challenging at this scale with both high intra-class spectral heterogeneity and inter-class spectral homogeneity. A solution to these issues is to take into account not only a pixel's spectral characteristics but also its spatial characteristics into classification. In this study, we develop a generalizable contextualized classification approach for high spatial resolution image classification. We apply the proposed approach to map vegetation growth forms such as trees, shrubs, and herbs in a forested ecosystem in the Sierra Nevada Mountains.

Keywords: high spatial resolution imagery; contextual classification; forest classification; remote sensing, lifeforms

To Mother and Father

ACKNOWLEDGMENT

I would like to express my great gratitude to my academic advisor Dr. Jonathan A. Greenberg for his time and numerous efforts guiding me through the completion of my master's studies in the past two years. I would like to thank him for giving me challenging topics but at the same time always being available when I needed help. I am grateful to be his first round graduate student. I would like to thank him from the very beginning: I greatly appreciate his efforts taking me around the campus and introducing me to his faculty colleagues when I first met him and visited the campus two years ago. During the past two years, I was very thankful to have the experience working so closely with Dr. Greenberg, which I believe not every graduate student has such experience with his/her academic advisor. Within the past two years working with Dr. Greenberg, I was amazed by his positive attitude to challenges, his amazing problem solving skills and experience, his vision and knowledge about the field of remote sensing, his passion about research, and his personality. I feel really honored to be chosen by Dr. Greenberg to study with him and to work in his lab. The past two years graduate studies with Dr. Greenberg has become a great memory in my life. Besides, I would like to thank Dr. Greenberg for his research funding and support. I would like to thank him for giving me research assistantship and taking me to high quality conference in the field and allowing me to get access to the powerful computer infrastructures, high quality computer hardware and software, very cool field equipment, and the imagery and data archive in his lab. Finally, I would like to express my gratitude again, without my advisor Dr. Greenberg it is impossible for me to work solely to complete my masters study.

I also would like to thank my thesis committee members, Professor James Dalling and Professor Shaowen Wang for taking the time and efforts reviewing my thesis and giving me extremely valuable feedback and advice from both technical and application aspects of this thesis work. I greatly appreciate their time serving as my committee members. Professor James Dalling is the head of the department of plant biology and professor Shaowen Wang consistently travels out of the state to conferences and meetings. I greatly appreciate they taking time out from their busy schedule to serve as my committee members.

I also would like to thank the GGIS department for giving me a year of fellowship to allow me to focus on my study. I would like to thank the previous department DGS Professor Flint Colin for accepting my application and bringing the good new to me. I would like to thank our department staff Susan and Matt for their support. I would also like to thank my TA instructor Dr. Jonathan Tomkin and the school of SESE staff Lara. It was a great experience being a teaching assistant with Dr. Tomkin in the course of ESE200.

I would like to thank my parents for their support and understanding for allowing their only child to pursue high education abroad. I would like to thank my boyfriend Dr. Jia for his mental support and constantly traveling to Urbana-Champaign bearing the horrible Amtrak train service to visit me. In addition, thanks Dr. Jia for his advice on the mathematical notations in my thesis writing. I must say that without Dr. Jia's support, it is impossible for me to complete my master study.

Finally, I would like to thank everybody I met here in UIUC. Thanks Evan Bowling in Computer Vision (CS543) course for being an awesome group partner for our course semester long project. Thanks Dr. Jeffrey Wayne Matthews for his comprehensive high quality course

NRES419 - Environment and Plant Ecosystem, which I greatly appreciate all the knowledge I gained from the course. Thanks all my dear graduate student colleagues in the department, it was very nice to have you all in my past two years. In addition, I would like to thank all my friends from the University of Minnesota for their mental support. Finally, special thanks to some of my close friends Danyang Chen, Hui Li, and Xinyi Tu for their accompany.

Figures

Tables

Chapter 1 Introduction

1.1. Research Background and Motivation

1.1.1. Research Background

One of the most important outcomes of remote sensing analyses is to use imagery to classify and delineate different objects and land cover types on the earth's surface. Historically, remote sensing has been used to perform land use/land cover ("LU/LC") mapping, in which complex mixtures of vegetation and non-vegetated surfaces are collapsed into a relatively small number of discrete classes labeled according to, typically, their relative fractional cover. For instance, the NLCD 2011 Classification System (Jin et al., 2011) defines a "forest" as an area of land with greater than 20% tree cover. While medium and coarse scale remote sensing (e.g. Landsat and MODIS imagery) with their relatively high temporal frequency are extremely valuable for various long-term research and management objectives such as long term sustainable forest management, biodiversity monitoring, carbon accounting, habitat protection, and sustainable timber production, these sensors lack the spatial detail to resolve fine landscape features such as individuals trees and shrubs, and are thus unable to produce many of the inventory products that are necessary for a full understanding of ecosystems processes (Falkowski et al., 2009; Pu, 2002). The impact of the focus on LU/LC mapping can be significant; the integration of these discrete LU/LC classes into ecosystem modeling versus using continuous fractional covers of plant functional types leads to profound differences in modeled energy and water flows (Bonan et al. 2002). Clearly, there is a need for remote sensing analyses to go beyond producing discrete LU/LC classes, and move towards an inventory-based approach to monitoring ecosystem characteristics, analogous to what field inventories can produce.

"Hyperspatial" remote sensing, defined as image data with pixels smaller than an object of interest (typically ≤ 1 meter ground sample distance, "GSD"), has been employed in monitoring and obtaining forest inventories data at an individual plant scale, including fractional plant functional type cover, crown size, species, and above ground biomass (Gougeon, 2003; Greenberg et al. 2005; Greenberg et al. 2006; Key, 2001). Historically, aerial imagery comprised the bulk of this type of image data, and the analysis techniques relied on significant field assessments and manual interpretation of the imagery. With the increasingly availability of hyperspatial satellite imagery and the advance of computing infrastructure and computing power, there is a general trend of obtaining and updating long-term detailed forest data via semi-automatic or automatic processing of high spatial resolution satellite imagery (Gougeon, 2003; Domain, 2005).

Although hyperspatial satellite imagery (e.g. IKONOS, WorldView 1, 2 and 3, and Quickbird) outcompetes medium and coarse resolution satellite imagery for the level of details that can be observed and extracted from the earth's surface, new challenges have emerged for detailed forest and vegetation classification using high spatial resolution imagery. Some of the challenges reported in the literature include how to deal with shadows caused by trees, high intra-class spectral variations, and high inter-class spectral homogeneity (Lu and Weng, 2007). Conventional classification techniques used with medium and coarse spatial resolution imagery focus on purely spectral-based classifications, but these techniques do not appear to work in many circumstances when applied to hyperspatial imagery. One likely explanation for this is that at medium to coarse scale resolution (>1m GSD), each discrete LU/LC class tends to have a separable spectral reflectance signature through the electromagnetic spectrum that is captured by different bands in the imagery. At this scale, the intra-class spectral variations often cluster around the ideal spectral reflectance signature distribution, and can be described by a probability distribution to model the classification based purely on the spectral characteristics. At high spatial resolution, however, the pixel sizes are often smaller than classes of interest, and there exists high spectral variations within a given class as well as among classes. The weak separability of spectral characteristics often makes it impossible to use solely spectral information to achieve satisfactory classification results. Using conventional spectral-based classification techniques on hyperspatial imagery results in pixels with identical spectral responses belonging to different classes being misclassified, and classification results often have a "salt-and-pepper" effect, where pixels falling within a single object (e.g. a tree) may be classified into multiple different classes. The limitations and decreasing of classification accuracy when applying conventional pixel-based approaches in high spatial resolution image classification have been stated in several studies (Woodcock and Strahler, 1987; Marceau et al., 1990; Greenberg et al., 2006; Yu et al., 2006; Lu and Weng, 2007).

1.1.2. Research Motivation and Objectives

This study was motivated need to improve the precision and accuracy of use/land cover classification applied to hyperspatial satellite imagery. The objective of this study was to overcome several challenges of applying conventional classification techniques to hyperspatial imagery classification, especially in forestry and vegetation classification. Specifically, we are asking:

- Is there a way to incorporate spatial information with spectral information together to perform classification in hyperspatial multispectral satellite imagery for forest classification/applications?
- Can contextual (neighborhood pixels) information be modeled to label the class of a given pixel?
- Can the contextual information be modeled with an advanced non-parametric machine-learning algorithm?
- Can this contextual classification solve some challenges such as shadow and high intra & inter-class spectral variations in hyperspatial imagery classification, especially in vegetation classification?

In this study, a general algorithm and approach is developed to overcome some of the challenges in hyperspatial imagery classification. The proposed approach takes advantages of fine contextual details in hyperspatial imagery and uses the contextual and spectral information together to improve classification capability.

1.1.3. Direction & Usage of the Proposed Approach

Different from many existing studies and approaches that can only be used to perform a certain level of classification (Franklin, 2001), the proposed approach is able to map and perform classifications in different scales to suit different classification needs. Individual species composition in lifeform class level can be classified first and then an upper level classification (e.g. forest cover type) can be generated based on the lifeform level classification. To better use the proposed approach, it is critical to understand the idea and theoretical framework discussed in the methods section in chapter 2. Generally, the proposed approach can be tried in high spatial resolution classification tasks if 1) one's classification classes share similar spectral properties, 2) there exists a high spectral variations within class, 3), there is no obvious geometric properties of the classification classes, 4) individual or groups of plants composition is desired (e.g. capture open canopy crowns as well as shrubs in the open space), 5) similar context of plant species composition patterns exist across large landscape (e.g. large area herb land can be found growing together with hardwood species, which both tend to live in areas with a relatively high water table – the valleys), 6) shadow is a problem (one important and exciting contribution of the proposed approach is that shadow is not a problem anymore, on the contrary, shadow can contribute important information to aid in classification.)

1.2. Related Works

1.2.1. High Spatial Resolution Classification Techniques in Detailed Vegetation Classification – A Brief Review

Many techniques have been developed to improve classification performance in hyperspatial imagery. The approaches generally fall into four groups: 1) texture analysis, 2) object-oriented approaches, and 3) multisource approaches, and 4) contextual approaches. Each approach has its unique advantages and works well in certain application settings. We discuss below each of the four approaches in detail together with referenced literatures and application examples.

Texture Analysis

Textural features extract information of local spatial statistics in a grey-scale image using fixed-size local windows surrounding a given pixel. Texture features can capture and evaluate whether a single layer grey-scale image has characteristics of homogeneity, linear structure (grey-tone linear dependencies), contrast, edges, complexity, etc. Natural texture features of an object surface such as the grain of a wood and the ripple of a wave can be represented by specific mathematical transformation and can be calculated by computer in a robust way. Haralick (1973) first developed a set of generalizable procedure for extracting texture properties by computing local spatial statistical variations of tone values of an image data and attracted numerous attentions from the remote sensing community. By setting a searching distance of a moving window, different angular nearest-neighbor gray-tone spatial-dependence matrices were computed to represent different texture features. These calculated layers are then added as additional bands, and the classification is then conducted using conventional pixel-based spectral classification classifier. Adding texture features is the mostly widely used approach in pixel-based high spatial resolution imagery classification among the remote sensing community (Lu and Weng, 2007).

Different texture features capture different properties of an image. Whether or not the computed texture features actually are useful in discriminating different classes in a classification system needs to be tested and investigated. This requires time and research efforts to identify suitable texture features with well-tuned parameters that work well with one's classification needs and the landscape of the study area. In addition, texture features are computed using a moving window. Determining a proper window size and shape to capture the texture features of the targeting classification classes is critical. Adding additional texture feature layers can increase or decrease classification accuracy (Mather and Tso, 2009).

Among the vegetation remote sensing community, feature extraction methods proposed by Haralick have been previously applied (Beguet et al., 2014; Kim 2009; Lu and Weng 2007; Pu 2012; Ruiz, 2004; Wulder 2004; Yu et al., 2006). Since the level of classification, the composition of landscape (vegetation) in the study area, and texture features selected and parameters used varied widely, below I summarize and cite two works that related to detailed forest cover type and species composition classification using texture features with hyperspatial imagery data.

Ruiz et al. (2004) tested four texture feature extractions: 1) statistical methods based on grey level co-occurrence matrices, 2) energy filters and edgeness factors, 3) Gabor filters, and 4) wavelet transform based methods to classify seven forest landscape types: high-density forest, mid-density forest, areas combining forest-shrub, shrubs, scattered trees, scattered shrubs and olive trees. They explored a Mediterranean forest where there was a wide variety of structural and vegetation diversity. A QuickBird panchromatic image resampled to 2.4 m GSD was used to compute various texture features. They reported that the texture approach is efficient in separating landscape types that have spectral heterogeneity where unique intensity variations can be found in grey-scale (black/white) image; while landscape types that have unique spectral response but are very homogeneous in grey-scale image are not suitable to use texture features to separate the classification objects/classes, such as pasture land and cereal crops or fallow. In addition, they reported the main limitation of using texture features was the border effect. Since each texture feature was computed using a local moving window of neighborhood pixels, error values were more cumulated at the transition area and they suggested further studies to reduce the border effect. Finally, they stated that there was no benchmark for the use of texture extraction method for classification, every single classification study has to test and find out the best texture features and parameters to suit the classification area and the purpose of the study, and apriori knowledge of the study area is required.

Franklin et al. (2000) tested texture features at different classification hierarchies at 26 forest plots in mixed-wood and boreal forests in Canada using airborne hyperspatial multispectral imagery (0.3m GSD) for species composition classification (hierarchical level 1: species name with percentage coverage; hierarchical level 2: conifer or hardwood with percentage coverage; and hierarchical level 3: three classes of mixed-wood, deciduous, and coniferous). They reported that adding texture features generally improved classification accuracy by 10-15%, however in some of their 26 plots, adding texture features did reduce the accuracy. In some of their plots, species composition appeared spectral distinctions for classification purpose, adding texture features in contrast confused this distinctiveness and reduced accuracy. They discussed the uncertainty in the incorporating the texture features in high spatial resolution imagery classification in forest applications and provided a checklist for better choosing texture features in various classification needs with tracing references in their discussion session.

Object-based approach

Object-based image classification approach also has been employed to overcome the problem of high spectral variability within class in hyperspatial imagery (Hamada, 2011; Kim, 2009; Lu and Weng, 2007; Myint, 2011; Qian and Gong, 2006; Yu et al., 2006). Object-based approaches first use image segmentation techniques to group relatively homogeneous pixels into polygon objects, and then performs classification on the segmented objects. By grouping objects in grey-scale layer, the confusion of high spectral variation in a certain pixel position through all bands is avoided. The quality and accuracy of the segmentation in the first step is critical to the final classification results.

Hamada et al. (2011) applied object-based approach with various spectral band combination indices as features to classify life-form cover types in California sage scrub communities using different spatial resolution remote sensing imagery (SPOT at 10m GSD, QuickBird at 0.6m GSD, and infrared aerial photography at 0.15m GSD). True shrub, subshrub, herb, and bare ground were their classes. The object-based approach was successfully applied to capture individual shrub canopies and patches. They reported true shrub and bare ground were relatively easy to obtain reliably classification results, but subshrub and herb remain challenging. The object-based approach worked well with QuickBird imagery (0.6m GSD) where compared to the 0.15m GSD aerial photography, as the 0.6 m GSD did not have significant spectral variability within a class and at the same time maintains spectral separability between classes. The high spectral variability within a class at a 0.15m spatial resolution confused the segmentation algorithm to segment relatively homogeneous gray tone objects in the first step of the object-based approach, which significantly influenced the final classification results.

Kim et al. (2009) applied object-based approaches combined with texture features to map the three forest types using hyperspatial IKONOS imagery in a close canopy (dense) national forest in North Carolina. Their classes were deciduous broad-leaved forest, evergreen needle-leaved forest, and mixed evergreen deciduous forest, approximated the upper Level 3 Formation Hierarchy level in the Federal Geographic Data Committee (FGDC) National Vegetation Classification Standard Version 2 (FGDC, 2007). One of their research questions was whether segmentation quality greatly influences the classification results and they found the answer to be yes. In addition, for forest application in their study site and classification level, they found that adding texture features did not increase classification accuracy primarily because the texture properties among their classes were very similar.

Yu et al. (2006) applied object-based approaches to perform detailed vegetation mapping of 48 classes using hyperspatial airborne remote sensing. Their classification scheme followed the California Vegetation Classification System (Vegetation Classification and Mapping Program, September 2003 edition), which was more detailed than the Level 3 in the USGS land-use, and Land-cover Classification System (Jin et al., 2011). They performed the object-based approach with different features such as texture features, ancillary topographic data and geometric features and achieved average overall accuracy of 51 percent to 58 percent, where 11 out of 43 of their vegetation alliances achieved results of 60 percent and higher (Yu et al., 2006).

Among the remote sensing literature, the suitability of using object-based approaches appears to be limited to when the grey-level image has relatively clear edges to delineate and segment objects that relate to the targeted classification objects/classes. Urban applications in these circumstances were likely to achieve better classification results using object-oriented approach. For example, Sun et al. (2005) compared pixel-based and object based approach in classifying urban environment and found object-oriented outcompeted pixel-based approach. This makes sense that classes with similar spectral responses like roads, path, and parking lots that are not distinguishable by the pixel-based approach can be segmented out as different objects in object-based approach (Soe 2011; Sun et al., 2005).

In vegetation applications, object-based approaches have been found to be successful in the level of land cover type classification using SPOT imagery by Duro (2011). They reported that object-based approaches did a better job than pixel-based approaches in classifying some of their classes such as wetland, riparian and mixed grassland. However, in general, the object-based approach and the pixel-based approach had similar overall classification accuracy in their classification scheme and study area.

Multisource approach

The multisource approach of image classification uses multiple gridded data sources, geometrically co-registered and rescaled to a consistent resolution, fused together into a single dataset then used with a classifier. The input datasets often contain remote sensing imagery from multiple sensors, as well as ancillary spatial data such as topography and climate. Solberg (1996) applied various imagery fusion techniques to incorporate multisource remotely sensed data in forest classifications. Solberg (1996) co-registered 0.5m GSD multispectral aerial photography with 0.5m GSD rasterized laser altimeter data and a reconstructed forest map to classify seven classes: shadow, shrub, dry grassland, deciduous forest, rock, peat, and conifer. Local spatial-spectral features were added to perform the classification. Using an iterative contextual classification algorithm, the iterated conditional modes (ICM) (Besag, 1986) were used to update a given pixel based on all available information in which the class with the maximum conditional probability was labeled. This technique is based on Bayesian probability theory, a detailed description of which can be found in Tso and Mather (2009). From the experiment results, Solberg found that adding the reconstructed forest map contributed the most improvement of accuracy in their study area, a coniferous forest in southeast Norway. Solberg also tested adding SAR data and an 81-band airborne spectrometer dataset collected at a spatial resolution of 5.6m x 5.6m to map six classes of different tree species with variations in tree height: Oak 1 (13-30m), Oak 2 (13-30m), Oak 3 (>30m), Beech 1 (<30m), Beech 2 (>30m) and Pine in the Fontainebleau forest in France. Solberg achieved a 98.7% accuracy by including all the multisource data and using the contextual classification technique. Another finding of Solberg's (1996) study was that choosing a more advanced statistical model resulted in a better improvement to classification performance than adding additional remotely sensed data. For multisource approaches in image classification, advanced techniques and reviews on data fusion can be found in Solberg (1999) and Abidi and Gonzalez (1992).

Contextual Approach

The basic philosophy of the contextual approach to image classification is to label the class of a pixel based on its neighborhood pixels, but in a more generalizable way than texture or object-based approaches. The application of contextual approaches to image classification have been primarily confined to the pattern recognition and computer vision communities, and is rare amongst the remote sensing community (Fröhlich, 2013). The increasing availability of hyperspatial imagery now makes it possible to apply advanced image processing techniques developed within the pattern recognition and computer vision communities to remote sensing applications.

We will first review the conventional pixel-based, purely spectral classification approaches from a statistical perspective to help better understand the idea of contextual approach of classification: each material on the earth's surface has specific reflectance signatures through the electromagnetic spectrum. Passive remote sensing imagery usually has multiple bands capturing electromagnetic radiation (EMR) across the visible and infrared regions. In an ideal situation, two pixels belonging to the same classification class should have the same reflectance signatures throughout the electromagnetic spectrum (feature space). In realistic situations, there is noise from the sensor and from the variations within the same class. Thus, the spectral response of pixels in the same class tends to cluster around the ideal response in the feature space. This cluster of the responses usually can be described with a suitable probability distribution. By sampling enough pixels responses in a class, the responses of pixels through feature space can be learned by supervised learning algorithms to make predictions on unlabeled pixels based on their responses through feature space. However, not all classes have separable reflectance signatures in the measured parts of the electromagnetic spectrum. In addition, with the increase in pixel resolution with hyperspatial satellite imagery, the spectral variations of pixels within the same class increase significantly. Thus, the spectral response of pixels within the same class might not have clustered distribution around the ideal spectral response of the class thus it is hard to use statistical learning algorithms to predict based sole on spectral information.

Contextual techniques are developed to tackle this problem. One approach is based on probabilistic relaxation (Dizenzo et al., 1987; Faugeras and Berthod, 1981; Peleg. 1980; Richards et al., 1982). The probability of the neighborhood pixels of a given pixel is used to update the label of the given pixel iteratively. The specification of neighborhood pixels used is based on different statistical measurement and techniques. For example, Press (1996) proposed a directional neighborhoods approach in which Press proposed a set of 21 neighborhood cliques. For example, in a 5 x 5 window, clique #1 contains pixel positions 4,5,8,9,10,13 which compasses the pixel direction of northeast; clique #2 contains pixel positions 9,10,13,14,18,19 which compasses the pixel direction of east, and etc. The homogeneity was computed for each clique and the clique with the highest value of homogeneity is chosen as neighborhood pixels to label and update the center pixel. Another approach can be traced back to Toussaint (1978) based on the theme of sequential compound decision theory. Generally speaking, this approach attempts to find a path to search the entire image for the determination of a given pixel (Kartikeyan, 1994). Markov Random Field (Kindermann, 1980) is a tool that characterizes contextual information from the nearest neighborhood pixels (NN) to model prior probability density functions (p.d.f.), in turn, with the available of prior information, the classification can be expressed in terms of maximum a posterior (MAP) where Bayesian theory can be applied (Tso and Mather, 2009). Even though there are numerous research and techniques available in contextualized classification in the pattern recognition, computer vision, and statistics community, there are very few applications in the remote sensing literature, especially in vegetation remote sensing community.

Chapter 2 Manuscript

Abstract

Maps of classified surface features are a key output from remote sensing. Conventional methods of pixel-based classification label each pixel independently by considering only a pixel's spectral properties. While these purely spectral-based techniques may be applicable to many medium and coarse-scale remote sensing analyses, they may become less appropriate when applied to high spatial resolution imagery in which the pixels are smaller than the objects to be classified. At this scale, there is often higher intra-class spectral heterogeneity than inter-class spectral heterogeneity, leading to difficulties in using purely spectral-based classifications. A solution to these issues is to use not only a pixel's spectral characteristics but also its spatial characteristics. In this study, we develop a generalizable contextualized classification approach for high spatial resolution image classification. We apply the proposed approach to map vegetation growth forms such as trees, shrubs, and herbs in a forested ecosystem in the Sierra Nevada Mountains.

Keywords: high spatial resolution imagery; contextual classification; forest classification; remote sensing, lifeform

2.1. INTRODUCTION

Remote sensing has long been used to assist in the monitoring and managing of forestlands. Since the 1940s, aerial photography has been used for manually interpreting forest attributes (Gougeon 2003). Around the 1970s, the availability of coarse and medium spatial resolution satellite imagery started being used to map and classify forestlands at larger scales (Colwell, 1960; Heller, 1975; Hoffer and Staff, 1975; Lachowski et al., 2000; Jensen, 1979). However, initial attempts of mapping forest land cover types with medium to coarse scale satellite imagery were limited by their relatively low spatial resolution. Mapping tree canopy and individual tree characteristics within stands was found to be difficult or impossible using medium to coarse scale satellite imagery (Brockhaus and Khorram, 1992; Carreiras et al., 2006; Congalton et al., 1991; Franklin, 1994; Katoh, 1988). With the increasing availability of high spatial resolution satellite imagery ("hyperspatial", <1m ground sample distance, "GSD") in the late 1990s (Ehlers et al., 2003; Ehlers, 2004) combined with the progression of more advanced image processing algorithms, more detailed forest data can be obtained from these hyperspatial imagery (Falkowski, 2009). Satellite imagery (such as IKONOS, Quickbird, GeoEye, and WorldView 1, 2 and 3) outcompetes aerial photography in assisting forest monitoring and management with their relatively large spatial coverage, relatively high temporal frequency, and low within-scene distortion. There is a general trend of obtaining and updating long-term detailed forest data, moving away from manually interpretating imagery and moving towards semi-automatic or automatic processing of the hyperspatial satellite imagery (Domain, 2005; Gougeon, 2003).

Conventional classification techniques developed back in the 1970s used with medium or coarse scale resolution satellite imagery continue to be applied today. Conventional pixel-based classification techniques typically use a pixel's spectral reflectance value to distinguish different classification classes. This method works well with coarse and medium resolution as each pixel covers relatively large areas (>10m GSD), at which land cover/land use classes typically have unique spectral characteristics. At higher spatial resolutions (<1m GSD), a pixel may be smaller than an object of interest (e.g. a tree), potentially resulting in high intra-class spectral variance. It is easy to have two pixels with similar spectral reflectance response in the electromagnetic spectrum belonging to two different classes in hyperspatial imagery, making purely spectral pixel-based classification a challenging task. Depending on the scale of classification, for example, for a "tree" cover class, different tree species may contribute to a wide range of spectral reflectance variance (Roberts et. al., 2004). In addition, at high spatial resolution, individual tree canopies are visible, containing pixels both on the sunlit side as well as on the shaded side of the canopy which leads to high variations of spectral reflectances within a class. For vegetation classification, there is a high chance that pixels in two different cover classes share similar or identical spectral reflectance responses resulting in misclassification when using solely spectral-based classifications. The limitations of applying conventional pixel-based classification approaches to hyperspatial imagery have been stated in many studies (Greenberg et al., 2006; Lu and Weng, 2007; Marceau et al., 1990; Woodcock and Strahler, 1987; Yu et al., 2006).

Many approaches and techniques have been developed to improve classification accuracy in hyperspatial imagery (Haralick, 1973; Lu and Weng; 2007; Mather and Tso, 2009; Ouma et. al, 2006; Puissant et. al, 2005; Wulder et al., 2004). Haralick (1973) first developed a set of generalizable procedure for extracting texture properties by computing local spatial statistical variations of tone values from single layer grayscale image and attracted numerous attentions from the remote sensing community. Among the vegetation and forestry remote sensing communities, methods such as object-based classification, multi-image fusion techniques, as well as incorporating ancillary data into the classification schemes have been applied to improve the classification performance in high spatial resolution imagery (Hamada, 2011; Kim, 2009; Lu and Weng, 2007; Myint, 2011; Qian and Gong, 2006; Solberg, 1996; Yu et al., 2006). Perhaps the most widely used approach in pixel-based high spatial resolution imagery classification involves adding different local-window texture features into a classification scheme (Beguet et al., 2014; Kim 2009; Lu and Weng 2007; Pu 2012; Ruiz, 2004; Wulder 2004; Yu et al., 2006).

However, whether or not the computed texture features actually represent discriminations of different classes in one's classification system needs to be tested and investigated. This requires time and research efforts to identify suitable texture features with well-tuned parameters for one's classification needs and the properties of the landscape of the study area. In addition, texture features are computed using a fixed-size moving window, determining a proper window size is a critical step in using these features in a classifier. Adding texture feature layers have been found to both increase and, at times, decrease classification accuracy (Franklin et al., 2000; Mather and Tso, 2009; Ruiz, 2004).

Incorporating texture features into a classifier has several reported issues: 1) extracting and computing local texture features with square or rectangular window shapes create artifacts at the spatial edges between different classes - the border effect problem (Moser, 2009; Ruiz, 2004); 2) testing combinations of parameters in each feature extraction and computation is time consuming and the choice to use what types of features highly depends on specific classes; 3) object-oriented approaches using texture features work well in urban environment where buildings and roads are structurally distinguishable (Duro, 2011; Soe 2011; Sun et al., 2005), whereas might not work well in forested landscapes where different classes might not have structural and spectral homogeneous units (Ruiz, 2004). In summary, there is no single technique available to meet the challenges in hyperspatial image classification across all common target classes and across the wide diversity of different landscapes.

In this study, we developed and tested a new method and framework for incorporating spectral and spatial information into a classifier applied to hyperspatial imagery. The basic operation of the proposed contextual classification approach is to label the class of a pixel based on its neighborhood pixels across multiple electromagnetic spectrum channels (bands) in a general, flexible way.

The paper is organized as follows. A detailed description of the proposed method is presented in Section II. A case study of the proposed methods for realizing lifeform classification applied to Worldview-2 imagery of the Lake Tahoe Basin in the Sierra Nevada Mountains

(California/Nevada) is described in Section III. Results are discussed in Section IV. Conclusion and discussion are presented in Section V.

2.2. METHOD

General Idea and Framework

Let N be a three-dimensional image array with row and column numbers N_1 x N_2 and n layers (Figure 1). $X_{i,j}$ is a single cell location in N with a set of n layer values. Let ω be the classified two-dimensional array of N. The number of rows and columns in ω is the same as the number of rows and columns in N. The values of ω are chosen from a user defined classification category. The conventional method of obtaining $\omega_{i,j}$ is to use values in $X_{i,j}$, in other words, $\omega_{i,j}$ is only dependent on $X_{i,j}$.

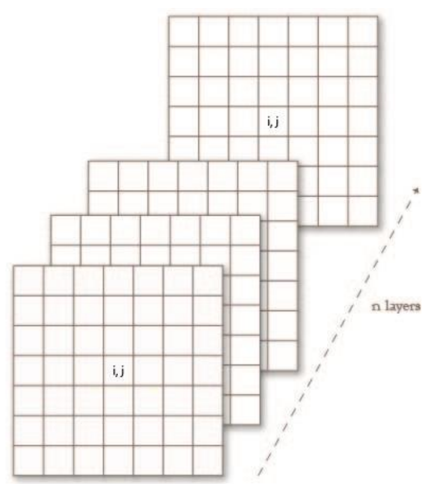

Figure 1. Conceptual diagram of contextual classification

In this study, we propose to use not only $X_{i,j}$ but also its nearby cells ranging from $X_{i-v,j-v}$ to $X_{i+v,j+v}$ to determine the class of $\omega_{i,j}$. Here we propose a simple way to control the number of neighborhood cells: the moving window approach. Let ρ be an odd number. The classification of $X_{i,j}$ is dependent on a set $S_{i,j}$ of pixels, where

$$S_{i,j} = \{X_{k,l},\ i-\frac{\rho-1}{2} \leq k \leq i+\frac{\rho-1}{2},\ j-\frac{\rho-1}{2} \leq l \leq j+\frac{\rho-1}{2}\}.$$

For example, we have an 8 band multispectral image and we want to classify a single pixel using contextual information from a 7 by 7 window around that pixel. In the feature extraction stage, in this case, n equals 8, ρ equals 7, the number of features being extracted would be from $X_{i-3,j-3}$ to $X_{i+3,j+3}$, which is 7 x 7 x 8 = 392 features, in other words, the class of $\omega_{i,j}$ is dependent not only on $X_{i,j}$ (8 values) but 392 values for all bands and all cells in the neighboring 7x7 window. Since we will be using a statistical learning approach for image classification, each feature will become a predictor variable that will be used to build a statistical model classification.

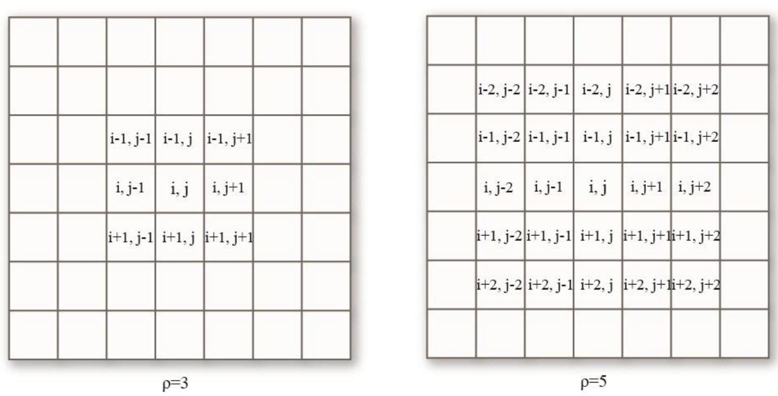

Figure 2. Conceptual diagram of contextual classification feature extraction

Random Forests

As ρ becomes larger and larger, the number of features being extracted increases exponentially. For example, if we consider contextual information of 15 by 15 neighborhood, the features available to be used to determine the class of a single pixel $\omega_{i,j}$ would be 15 x 15 x n layer. This large number of input features likely includes useless or redundant information and also greatly increases the computational needs to train and implement the classifier. Therefore, we need to utilize a classification algorithm that can used with overpredictive datasets without overfitting the classes, as well as potentially allowing the key features to be identified so an optimized model can be produced. While many machine learning classification algorithms can be leveraged, we chose to use Random Forests by Breiman (2001).

Random Forests is an ensemble decision tree based classifier, where each tree is trained using a bootstrap sample of m from the original training samples M and each split occurs at a variable chosen from a random subset of h from the original variables H (Breiman, 2011). For classification, the final result is based on a majority vote over all of the trees (the "Forest", Figure 3). By using a newly sampled random subset of variables for splits in each tree, the correlations between each tree in the ensemble are reduced and computational time is saved. Since the subset of variables chosen to be used in each tree is different and variables being sampled to a set can be in a new set next time, this helps stabilize the classification accuracy and acquire generalizable classification results where a small change in the nature of the training samples will not dramatically alter the classification accuracy (Breiman, 2001).

While Random Forests is generally insensitive to highly overpredictive training datasets, two issues arise when using Random Forests with a contextual features dataset. First, balanced training datasets are needed to properly represent classes that may have different sample sizes (Chao et. al, 2004). If the training dataset is not balanced, far more trees are needed to properly sample from the minority classes (Breiman, 2001). Second, models trained from the full feature set take a significant amount of computational time to apply to image datasets, so optimizing the features used to train the final model by removing useless or redundant features can significant speed up the prediction.

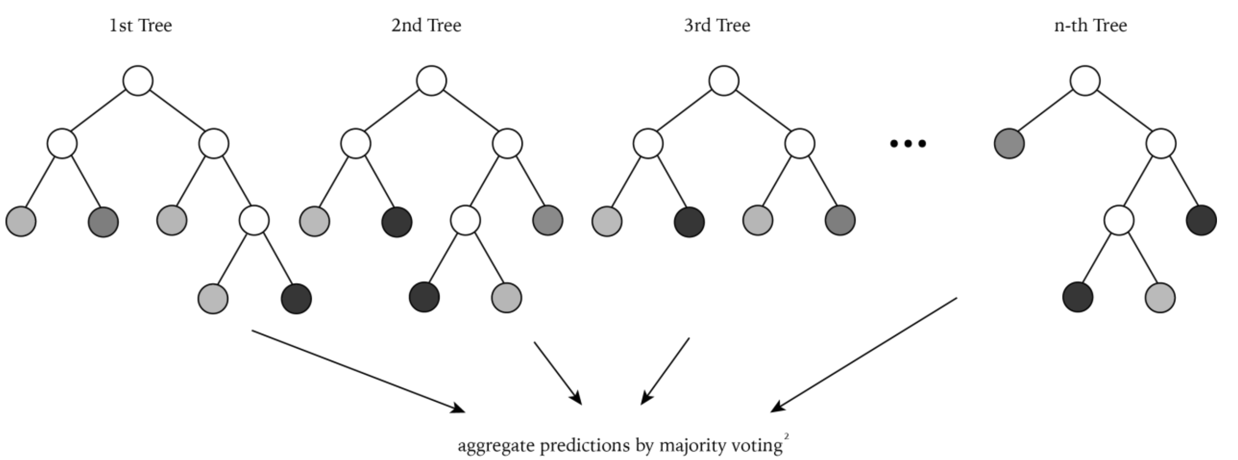

Figure 3. Conceptual diagram of Random Forests model– n represents n number of binary decision tree. Each tree grows independently and the final prediction obtained is through aggregation by majority voting. Diagram is adapted from Fröhlich et al (2013)

Balanced Training Data

Several resampling techniques can be applied to balance the training dataset before using the Random Forest classifier. Downsampling techniques can be employed to balance the majority classes to have the same frequency as the minority classes. Upsampling techniques can be used to sample the minority classes with replacement to make all classes distributed equal as the majority class (Kuhn, 2014). SMOTE sampling technique can be employed to re-sample the data to a user-defined numbers of samples by up-sampling the minority class and down-sampling the majority classes (Torgo, 2010).

Variable Selection

Within the use of Random Forests, important predictor variables can be calculated and ranked by the variables' predictiveness. Several techniques have been developed to calculate the variable importance (VIMP) (Breiman, 2001; Diaz-Uriarte and Alvares 2006; Ishwaran et al., 2010). We employed the Maximal Subtree method proposed and implemented by Ishwaran et al (2010). Basically, a Maximal Subtree of a variable is defined as the terminal node where a given variable encounters its first split. Then the minimal depth is calculated to measure the distance from the root node to a variable's maximal subtree. The shorter the minimal depth of a given variable, the greater the impact of that variable has on prediction (Ishwaran et al., 2010).

2.3. Application of Contextual Classification to Lifeform Classification

Study Area

The study area is the northeastern side of Lake Tahoe Basin, located in the Sierra Nevada Mountains along the border of California and Nevada (Figure 4). Elevation ranges from 1,900m to 3,050m above sea level (a.s.l.). The climate of the Lake Tahoe Basin follows a Mediterranean pattern with long, cool wet winters and short, warm dry summers. Precipitation usually occurs between October and May as snow. The topographic complexity leads to high variations in temperature, precipitation, and solar radiation and has resulted in a rich diversity of vegetation types. The area is dominated by a variety of conifer species including White Fir, Jeffrey Pine, Red Fir, Lodgepole Pine, and Incense Cedar. Broadleaf tree and tall shrub species are also present, and include Quaking Aspen, Mountain Alder and several species of Willows. In addition to trees and tall shrubs, many species of shrubs and herbs are also present (Barbour et al., 2007).

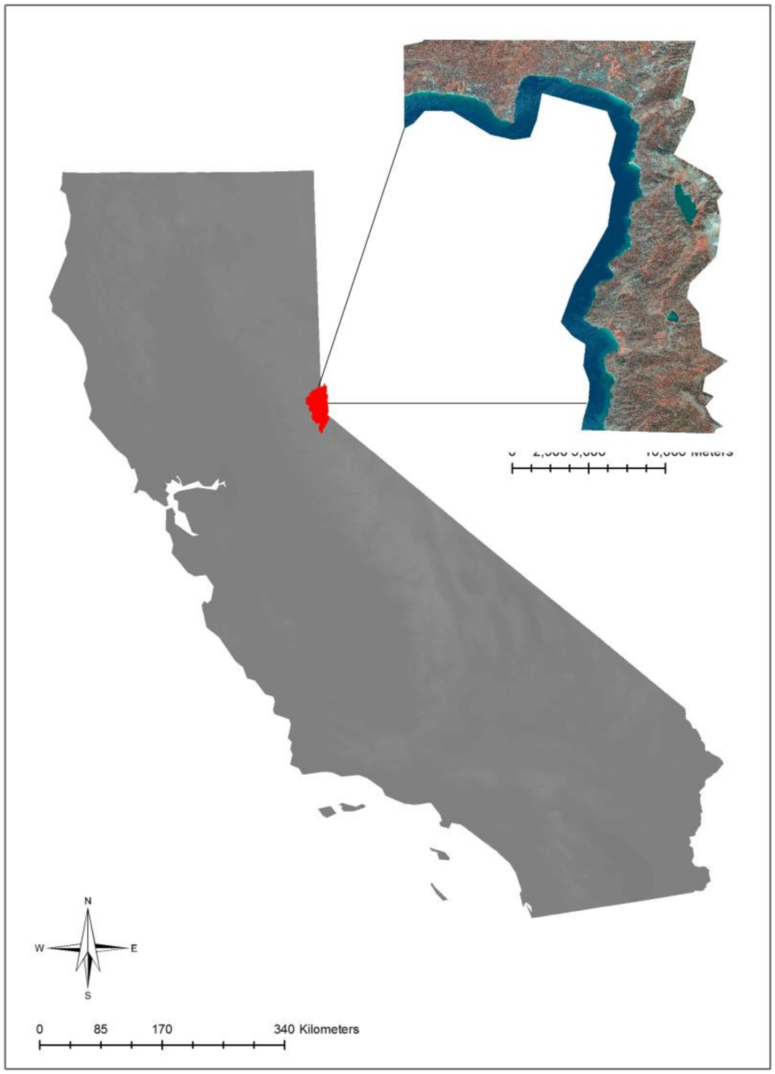

Figure 4. Study Area. Lake Tahoe is labeled in red. The imagery employed in this study covers the northeast of Lake Tahoe.

Remote Sensing Imagery

We used WorldView-2 ("WV2") imagery in this analysis. Worldview-2 is an 8-band, multispectral image that collects data at 0.45 m across a wide, panchromatic band, and at 1.85m for the 8 visible/near-infrared spectral bands that range between 400 to 1040nm. The acquisition was acquired at an off nadir angle of 8.39° on September 10^{th}, 2010. The azimuth was 137° and solar elevation was 54.22°. The imagery covers a total area of 422.48 km^2 (Figure 4). The imagery was atmospherically corrected and orthorectified by the USDA Forest Service Region 5 Remote Sensing Laboratory. The multispectral imagery was pansharpened to 0.45 m.

Training Data

We prepared training data by photointerpreting eight lifeform classes from the pansharpened WV2 imagery. More specifically, training polygons were digitized at the individual plant level for trees or at the level of a small, homogenous patch for shrub and herbs. Based on the plant species present in the study area, we grouped our training data into eight lifeform/growthform classes (Table 1).

Table 1. Eight lifeform/growthform classes with representing species

Lifeform/Growthform Classes	Dominant Species/Materials
Conifer Tree	White Fir, Jeffrey Pine, Red Fir, Lodgepole Pine, and Incense Cedar
Hardwood Tree	Quaking Aspen
Tall Shrubs	Mountain Alder, Willows
Shrubs	Arctostaphylos nevadensis, Ceanothus cordulatus, Ceanothus prostrates, Chrysolepis sempervirens, Quercus vaciniifolia, Spiraea densiflora ssp. splendens
Herb	Dry/wet grasses, Meadows
Soil	Barren soil, sand and dune
Barren and Impermeable	Barren, cliffs, and bedrock
Water	Water

We constructed a stratified random sampling design based on the Tahoe Basin Exiting Vegetation Map (TBEVM, Greenberg et al. 2006) vegetation subclass, which was based on the FGDC mapping standard (Table 2).

Table 2. Federal Geographic Data Committee Vegetation Classification Standard Subclass Categories

SUBCLASS
Deciduous closed tree canopy
Evergreen closed tree canopy
Mixed evergreen-deciduous closed tree canopy
Evergreen open tree canopy
Deciduous open tree canopy
Mixed evergreen-deciduous open tree canopy
Evergreen shrubland
Deciduous shrubland
Mixed evergreen-deciduous shrubland
Herbaceous Vegetation
Sparse Vegetation

Polygons representing each subclass were randomly chosen (one polygon/subclass), photointerpretation was then conducted on each chosen polygon to identify the target lifeform classes. The photointerpretation was guided by field data collected by USDA Forest Service and by the University of California at Davis in 2011, Google Earth historical imagery, Google street view, and vegetation index layers such as NDVI. We updated the training data iteratively by using the first round of training data to train a model, using the model to predict on randomly chosen image subsets stratified by FGDC subclass, correcting prediction errors and including new training sites into the training data. The iterative process of updating training data helps to capture contextual information into the training and learning process. For example, we did not include the long tail of conifer tree shadows over a flat surface of shrubs in the early training data. However, we noticed that the predictions of those long tails were classified incorrectly in the early rounds of image subset predictions. We then added new training data of these shadowed shrubs to the training data. For subsequent prediction subsets, this contextual information was learned and predicted correctly by the machine-learning algorithm. The iterative process of updating training data is illustrated in figure 5 and the final spatial distribution of training data is shown in figure 6.

Validation data was prepared using USDA Forest Service field data as well as additional independent photointerpreted data. Table 3 summarized the validation data for each class.

Table 3. Validation Dataset.

Class	# of pixels
Conifer Tree	55
Hardwood Tree	19
Herb	25
Shrub	27
Tall Shrub	15
Soil	20
Barren and Impermeable	37
Water	21

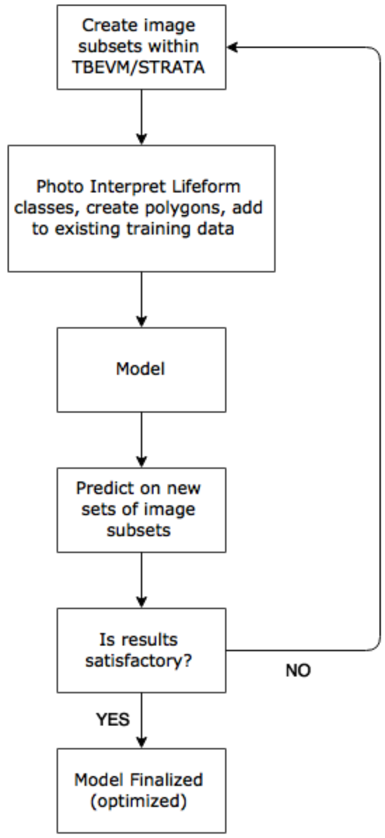

Figure 5. Flowchart of iterative process of updating training data.

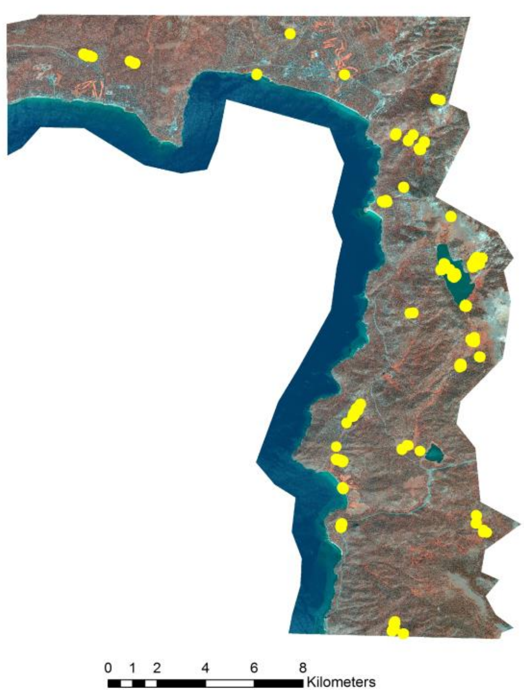

Figure 6. Spatial distribution of the final training data. Yellow dots represent the location of training samples.

A total of 237,649 training samples were extracted from training polygons using 15 by 15 window followed the method in Section II (Table 4).

Table 4. Training polygons and number of pixels extracted using 15 x 15 window

Class	# of polygons	# of pixels
Conifer Tree	101	9,832
Hardwood Tree	14	47,438
Herb	20	41,472
Shrub	18	3,478
Tall Shrub	13	13,139
Soil	21	11,649
Barren and Impermeable	11	22,562
Water	6	98,079

Model Construction and Optimization

We utilized the RandomForestSRC implementation of Random Forest (Ishwaran, 2015) within the R Statistical Computing framework (R version 3.1.3, 2015) to train a classification model relating the eight (8) lifeform classes to the contextual information extracted from the neighboring pixels. To begin, we implemented a statistical downsampling based on the minority class (N = 4,378) to balance the training samples.

RandomForestSRC provides a method to choose a reduced set of variables important to the model (Ishwaran et al. 2010). We used the Maximal Subtree method to calculate the variable importance (VIMP) (Ishwaran et al., 2010). This allows us to first determine the correct window size, with the goal of confirming that the window size used was not too small for the analysis in question. As such, we ran the RandomForestSRC model on several versions of the training data, in which only features from within a given distance from the center pixel were used, ranging from 1 to 15 pixels in radius. At each window size, the number of total available variables and selected variables were recorded, and percentage of important variables chosen from all available variables was calculated. Figure 7 shows that by extracting contextual information with window sizes ranging from 1 to 7 pixels, all the features within this radius were selected as important to construct the model, which means that these windows are likely too small (there may be additional benefit from searching the feature space at a greater distance). Beyond 7 pixels, the percentage of variables chosen decreased to a minimum of at a 13 x 13 window size and then stabilized at 15 x 15, which confirmed that 15x15 was sufficiently large to capture the important contextual information.

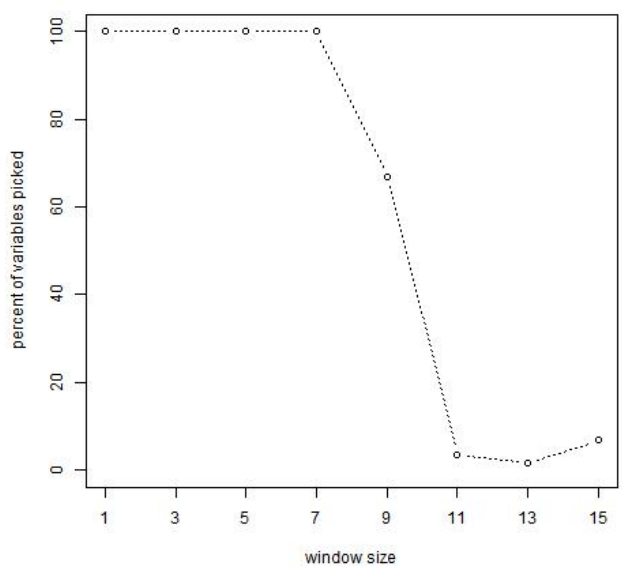

Figure 7. Optimal window size test plot.

Once the window size was confirmed, the next step was to optimize the variables needed to run the model without a loss of accuracy. This step repeated the previous step with the 15x15x8 window of features (a total of 1800 available features) and using the VIMP to select which of the 1800 variables were chosen as important for the model.

As a final optimization, we reduced the total number of trees needed to construct the final model by examining the change in error as additional trees were added, and stopping the model when the error rate improved by less than 0.1% by adding an additional tree. From a starting point of 1000 trees, we found that our model required only 899 trees.

Once the optimizations had been performed, a final model was constructed which was used with the validation data and to predict on images. To evaluate the model itself, we summarized the histograms of the radius, azimuth, and bands chosen for use with this model.

2.4. Results

Of the 1800 features found within a 15x15 window, with 8 spectral bands, 122 variables were selected for use in the final model. Of these 122 variables, five of the eight WV2 spectral bands were found to be important: blue (400-450nm), red (630-690nm), "red edge" (705-745nm), and both near infrared bands (770-895nm, and 860-1040nm) (Figure 8). Variables were used across all radii, although there was a notable decrease in the number of variables chosen at distances between 2 and 5 pixels, with the most important radii being close to the center, and at the maximum distance from the center. Figure 9 shows a histogram of variables chosen for a given range of radii vs. the total possible of variables possible at that radius. Figure 10 shows a histogram of the variables used as a function of azimuth. The distribution of the selected important variables in azimuth showed that most variables were picked up in the northeast and southwest directions. Figure 11 visualizes the spatial and spectral distribution of the selected variables used in the model.

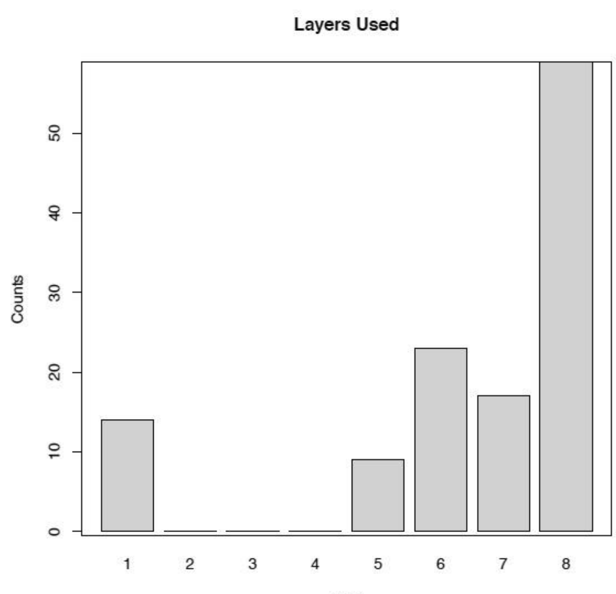

Figure 8. Visualized spatial distribution of the selected variables through different bands in the imagery.

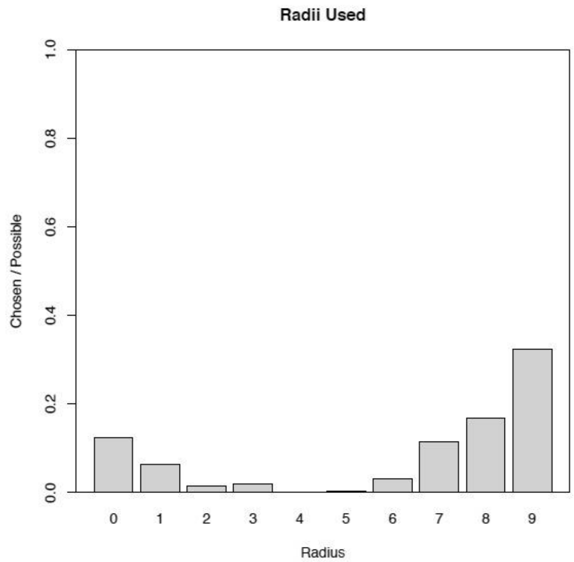

Figure 9. Percentage of variable selected from all available in a given range of search radius.

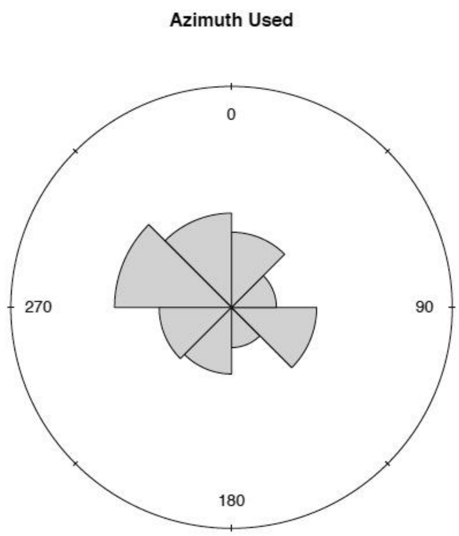

Figure 10. Azimuth plot. All the important variables selected in direction to the center pixel.

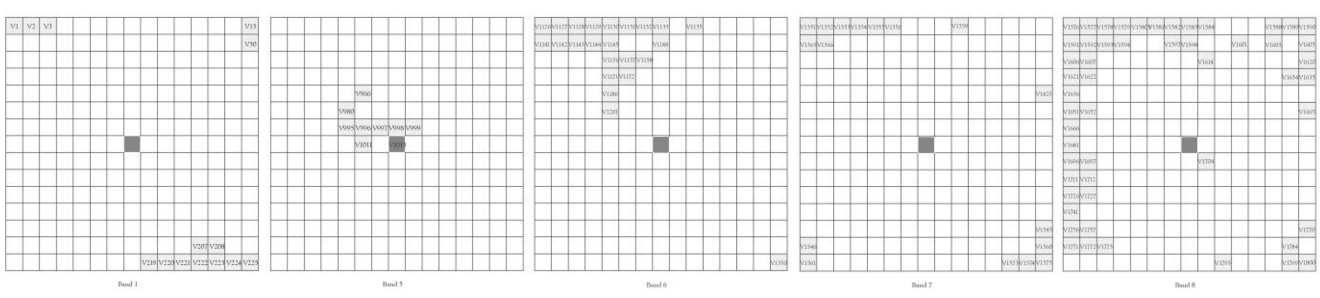

Figure 11. Visualized the spatial and spectral distribution of the selected variables used in the model.

Accuracy Assessment

We used an independent set of validation data to evaluate the results. These data were obtained by the USDA Forest Service as well as additional independent photointerpreted data points. A total of 219 validation points were used. A confusion matrix was produced, producer's and user's accuracy were computed (Table 5 and 6). Compared to the validation data, the overall classification accuracy of the contextual classification model was 76% with Kappa of 0.72.

Table 5. Confusion Matrix of All 8 Classes

	Barren and Impermeable	Conifer Tree	Herb	Hardwood Tree	Soil	Shrub	Tall Shrub	Water	# of Classified Pixels
Barren and Impermeable	**28**	0	0	0	4	0	0	0	32
Conifer Tree	1	**52**	2	0	5	3	0	0	63
Herb	1	0	**13**	1	0	3	0	0	18
Hardwood Tree	0	1	1	**15**	1	2	1	0	21
Soil	2	0	0	0	**16**	0	0	0	18
Shrub	5	2	1	2	8	**18**	0	0	36
Tall Shrub	0	0	8	1	0	1	**14**	0	24
Water	0	0	0	0	0	0	0	**21**	21
# of Ground Truth Pixels	37	55	25	19	34	27	15	21	233

Table 6. User and Producer Accuracy

	Producer's Accuracy	User's Accuracy
Barren and Impermeable	75.7%	87.5%
Conifer Tree	94.5%	82.5%
Herb	52.0%	72.2%
Hardwood Tree	78.9%	71.4%
Soil	47.1%	88.9%
Shrub	66.7%	50.0%
Tall Shrub	93.3%	58.3%
Water	100%	100%

The class of water achieved the highest 100% accuracy because of its distinguishable spectral signatures as well as unique contextual pixel neighborhood, as a water pixel is much more likely to be next to another water pixel. For vegetation classes, conifer trees stood out as the most accurate class. This could be partially because that the distinct crown shadow contributing unique contextual patterns for distinguishing conifer trees from other vegetation classes.

In addition to the standard pixel-level accuracy results, we applied the final model to several randomly chosen scene subsets to provide additional qualitative visual assessments of the accuracy of the model (Figure 12, 13, and 14).

Samples of Classified Image Subsets

Two image subsets of relatively open conifer canopy area were classified as shown in figure 12. The shrubs and soils in the open canopy areas were captured by the model and classified properly.

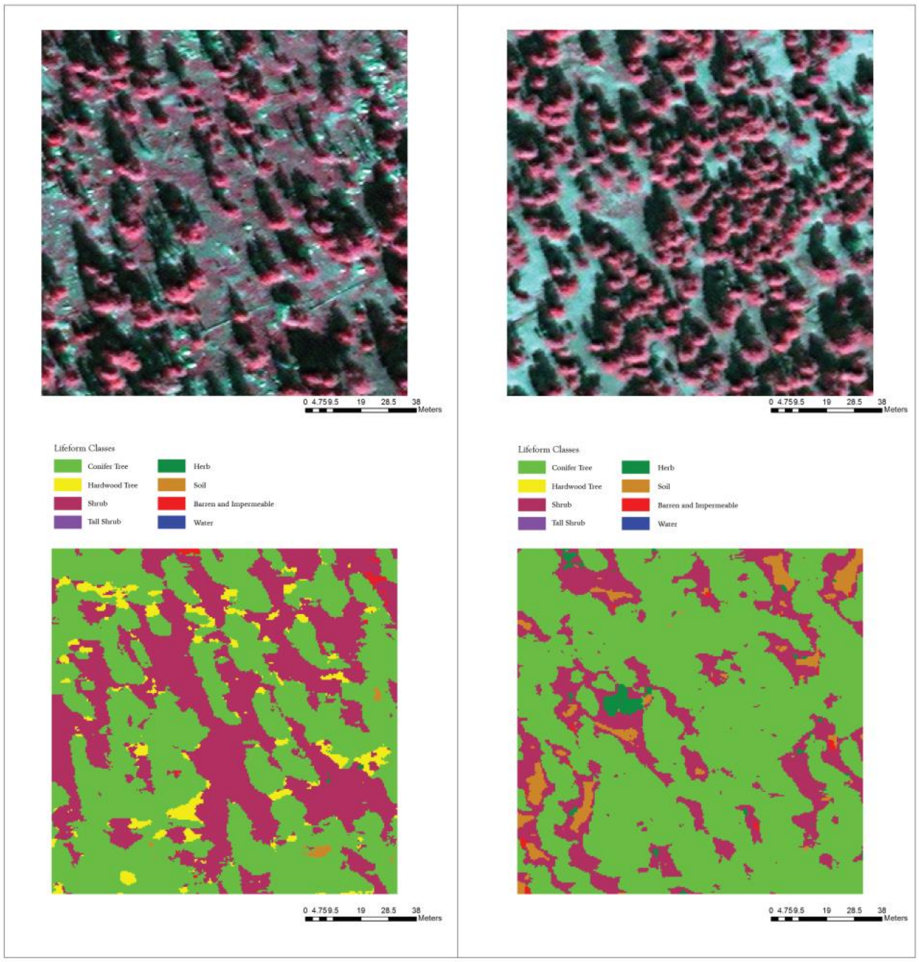

Figure 12. Lifeform classification result on 251x251 image subsets.

Two image subsets of 1) mixed open conifer and deciduous trees and 2) open conifer with shrubs were classified as shown in figure 13. The model was able to distinguish hardwood trees, tall shrubs, shrubs, and conifers especially shown in the left classification result in figure 13.

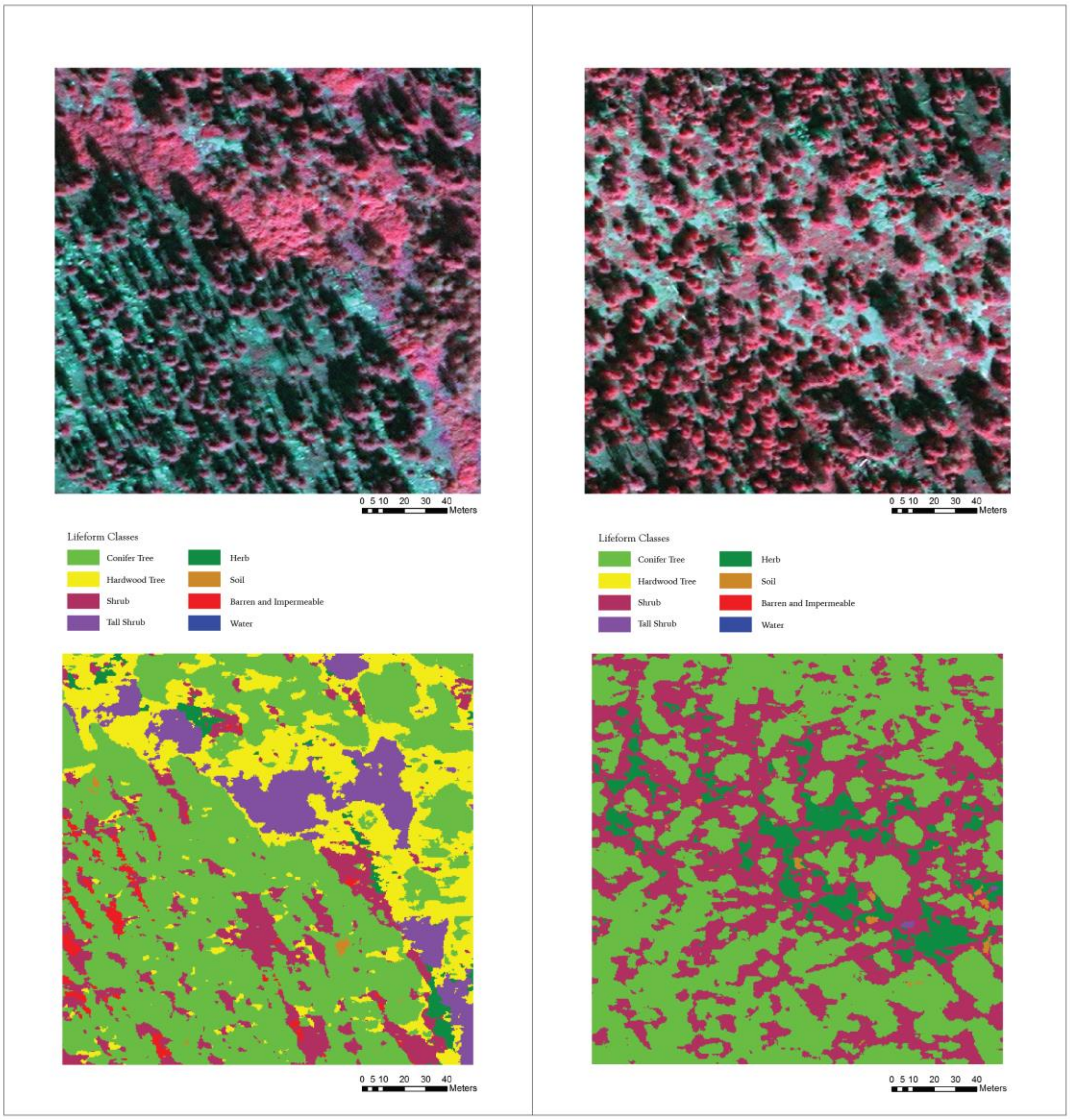

Figure 13. Lifeform classification result on 401x401 image subsets.

The large image subset shows the classification result of a mix conifer, deciduous and shrub land near the lakeshore (Figure 14).

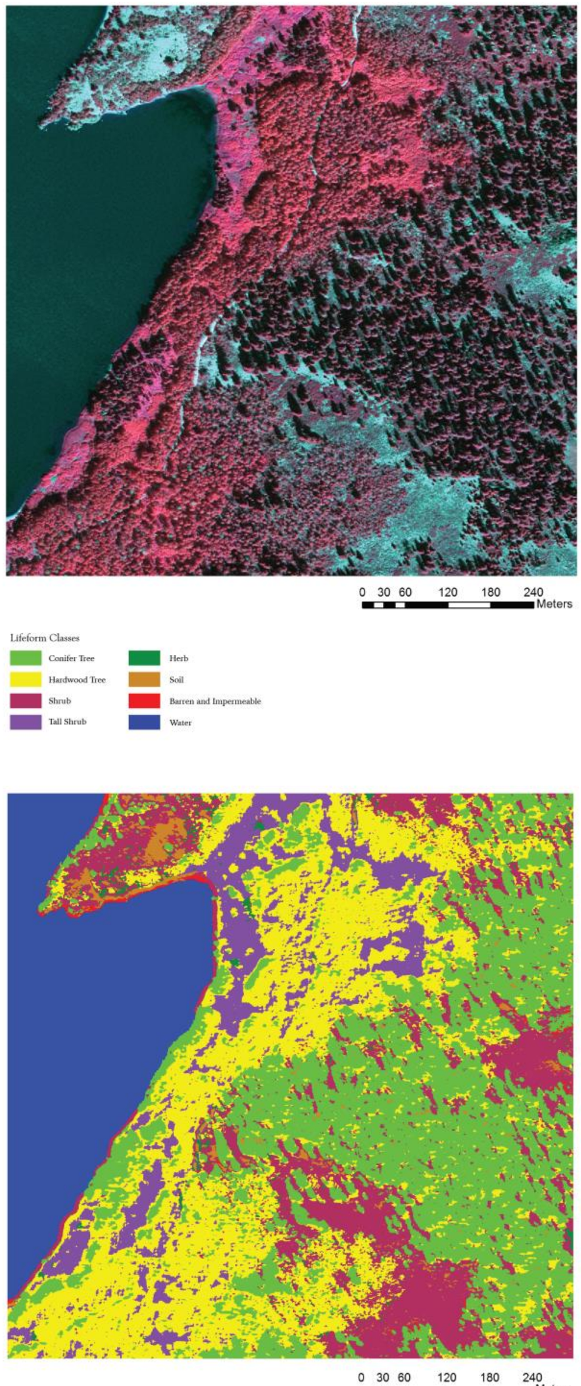

Figure 14. Lifeform classification result on 1600 x 1600 image subset.

2.5. Conclusion and Discussion

In this study, we developed a generalizable contextual classification approach to incorporate spatial information combined with spectral information together into a machine learning classifier.

The results of the variable analysis pointed towards a sensible choice of important variables. The distribution of the selected azimuths showed that most variables were picked up at northwest and southeast directions, which was along the solar azimuth at the time of the image was acquired. This makes sense, in that the model picked up the contextual information such as shadows as a clue for the determination of a class. Tall conifer and hardwood trees have significant within-canopy shadows, whereas shorter shrubs and herbs do not.

The radii chosen also made sense in light of the characteristics of the trees used in the training dataset. Figure 15 shows that the conifer tree crown size in meters from the training data, which indicated that conifer tree crown radii varied from 1 to 5 meters in our study area. Most tree crowns were at about 2.0 to 2.5 meters in radius. A 15 x 15 window yields 9.9 pixel search radius which is equivalent to about 4.5 meters in ground. This was twice the distance from the center of a tree crown pixel to the edge of the crown. Generally speaking, the optimal search window size was greater then the size of the classification objects/classes.

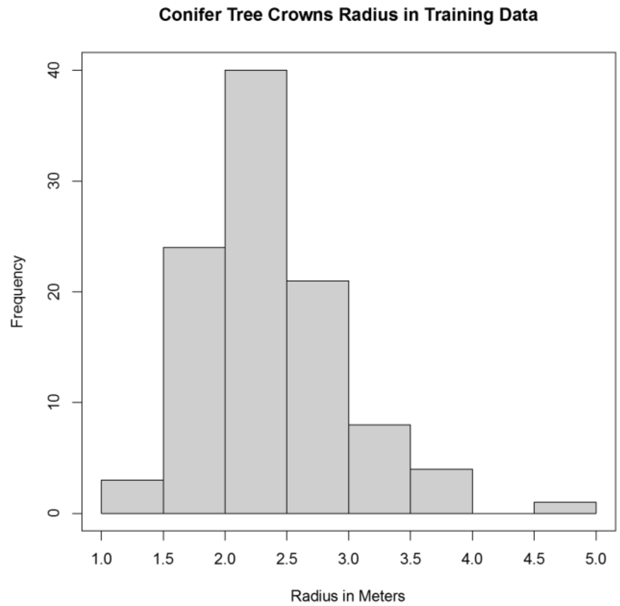

Figure 15. Conifer tree crown size from the training data.

The spectral profile of the central pixel for each class is graphed in figure 16 where it is clear that the conifer tree, hardwood tree, shrub, herb, and tall shrub classes shared very similar spectral responses through all the 8 spectral bands in the WV2 imagery. In addition, there existed a very high inter-class as well as intra-class spectral variation (Figure 16 &17). A conventional pixel-based, purely spectral classification, would be unlikely to separate these five classes.

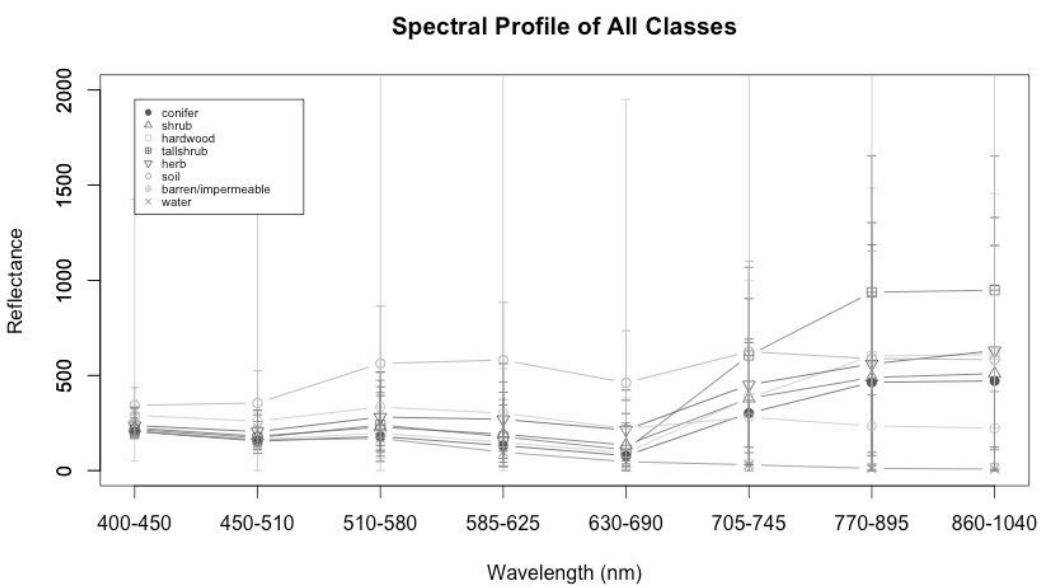

Figure 16. Spectral profile of all classes.

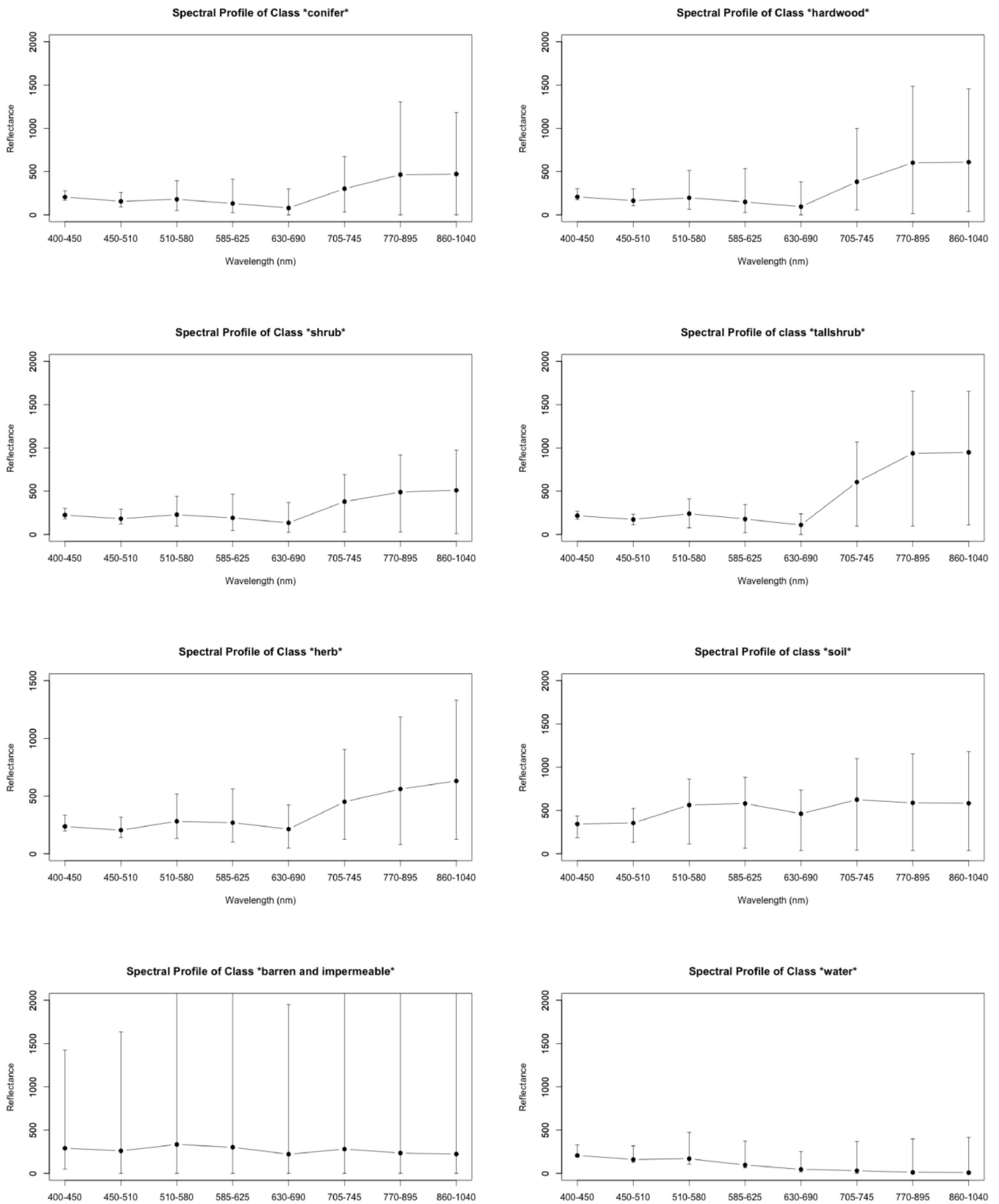

Figure 17. Plots of spectral profile for each classification classes. For each class, the training samples mean were plotted as black dot. Maximum and Minimum samples were plotted as error bars to show the high intra-class spectral variation.

The contextual classification approach overcame several challenges in classifying high spatial resolution imagery. By taking spatial information into the classification scheme, the spectral value is not the only parameter in determining a class of a pixel, which overcame the high intra- and inter-class spectral variation problem. Shadows reported by previous studies as noises and problems in classifying high spatial resolution imagery not only did not cause errors in the analysis, but were leveraged with the contextual feature extraction method, allowing the tree shadows to be valuable clues for the contextual classification (Lu and Weng, 2007). Problems of the border effects when using texture layers computed by box-shape moving windows did not occur in this proposed method even though the box-shape moving window was employed in extracting features.

Contextual classification methods can be easily adapted to different landscape types and class types. The quality of training data was the most crucial part of the proposed approach. We found that it was important to capture a wide variety of spatial contexts for each of the classes (e.g. trees in open and closed canopies, different species of trees, etc.), and that multiple iterations with training data and visualizing results, refining and correcting errors were helpful in producing an accurate model. Finally, our results point towards a potential difficulty in porting a model from one image scene to another, particularly if the images are captured with different solar azimuths. The development of azimuth-agnostic training data may be necessary to allow the production of a model from one scene's data to be applied to another scene.

Future Research & Experiment Directions

There are several aspects that could be used to improve classification performance. First, in our moving window approach, different shapes of the moving window such as circles or lines may be more appropriate than a rectangle which biases towards diagonal distance. Second, while Random Forest was used in this study, the proposed framework should be somewhat agnostic towards the particular machine learning algorithm uses, so other machine learning algorithms such as Support Vector Machine (SVM) or Neural Networks may lead to improved classification accuracies or computational efficiency. Third, we would like to conduct quantitative analyses on the training sample to see how different sampling design methods (e.g. spatial or class-specific stratification for stratification; proportional to area or equal allocation for sample allocation) might influence the classification results (Jin, 2014). Finally, we suspect that incorporating spectral (e.g. NDVI) and spatial indices (distant pixel reflectance divided by center pixel reflectance) may improve model accuracy, so we intend to incorporate these into future models.

In conclusion, we demonstrated that the contextual classification method could be leveraged to provide a generalizable and flexible approach to mapping vegetation in diverse ecosystems such as those found in the Lake Tahoe Basin. The output classification maps are able to begin to fill the information gap between large-scale Land Use/Land Cover type maps and detailed ground-based field inventories. We believe these approaches can provide improved and more detailed classification maps than are possible to derive from medium resolution remotely sensed imagery. As the spatial resolution of imagery from newer earth-observing sensors continues to increase, many advanced techniques developed from fields like computer vision become more relevant to the remote sensing community. There is tremendous potential for interdisciplinary research projects to collaborate researchers in both fields to explore exciting techniques and raise interesting research questions and applications.

References

Barbour, Michael G., Todd Keeler-Wolf, and Allan A. Schoenherr, eds.*Terrestrial vegetation of California*. Univ of California Press, 2007.

Bechtold, William A., and Paul L. Patterson. "The enhanced forest inventory and analysis program: national sampling design and estimation procedures." (2005).

Beguet, B., et al. "Classification of forest structure using very high resolution Pleiades image texture." *Geoscience and Remote Sensing Symposium (IGARSS), 2014 IEEE International*. IEEE, 2014.

Besag, Julian. "On the statistical analysis of dirty pictures." *Journal of the Royal Statistical Society. Series B (Methodological)* (1986): 259-302.

Bonan, G. B., S. Levis, L. Kergoat, and K. W. Oleson, Landscapes as patches of plant functional types: An integrating concept for climate and ecosystem models, Global Biogeochem. Cycles, 16(2), doi:10.1029/2000GB001360, 2002.

Boyd, D. S., and F. M. Danson. "Satellite remote sensing of forest resources: three decades of research development." *Progress in Physical Geography* 29.1 (2005): 1-26.

California Department of Fish and Game, List of California Terrestrial Natural Communities Recognized by the California Natural Diversity Database, Wildlife and Habitat Data Analysis Branch, Vegetation and Classification and Mapping Program, September 2003 Edition.

Chen, Chao, Andy Liaw, and Leo Breiman. "Using random forest to learn imbalanced data." *University of California, Berkeley* (2004).

Duro, Dennis C., Steven E. Franklin, and Monique G. Dubé. "A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery." *Remote Sensing of Environment* 118 (2012): 259-272.

Elsevier Science, and Gonzalez Rafael C Abidi Mongi a. *Data Fusion in Robotics & Machine Intelligence*. academic press, 1992.

Falkowski, Michael J., et al. "Supporting large-area, sample-based forest inventories with very high spatial resolution satellite imagery." *Progress in Physical Geography* 33.3 (2009): 403-423.

Franklin, S. E., et al. "Incorporating texture into classification of forest species composition from airborne multispectral images." *International Journal of Remote Sensing* 21.1 (2000): 61-79.

Franklin, Steven E. *Remote sensing for sustainable forest management*. CRC Press, 2001.

Fröhlich, Björn, et al. "Land cover classification of satellite images using contextual information'." *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 3 (2013): W1.

Greenberg, Jonathan A., et al. "A bottom-up approach to vegetation mapping of the Lake Tahoe Basin using hyperspatial image analysis." *Photogrammetric Engineering & Remote Sensing* 72.5 (2006): 581-589.

Greenberg, Jonathan Asher, Solomon Z. Dobrowski, and Susan L. Ustin. "Shadow allometry: Estimating tree structural parameters using hyperspatial image analysis." *Remote Sensing of Environment* 97.1 (2005): 15-25.

Gougeon, François A., and Donald G. Leckie. "The individual tree crown approach applied to Ikonos images of a coniferous plantation area."*Photogrammetric Engineering & Remote Sensing* 72.11 (2006): 1287-1297.

Hamada, Yuki, Douglas A. Stow, and Dar A. Roberts. "Estimating life-form cover fractions in California sage scrub communities using multispectral remote sensing." *Remote Sensing of Environment* 115.12 (2011): 3056-3068.

Haralick, Robert M., Karthikeyan Shanmugam, and Its' Hak Dinstein. "Textural features for image classification." *Systems, Man and Cybernetics, IEEE Transactions on* 6 (1973): 610-621.

Ishwaran H. and Kogalur U.B. (2014). Random Forests for Survival, Regression and Classification (RF-SRC), R package version 1.6.0.

Ishwaran, Hemant, et al. "High-dimensional variable selection for survival data."*Journal of the American Statistical Association* 105.489 (2010): 205-217.

Jin, Suming, et al. "A comprehensive change detection method for updating the National Land Cover Database to circa 2011." *Remote Sensing of Environment*132 (2013): 159-175.

Jin, Huiran, Stephen V. Stehman, and Giorgos Mountrakis. "Assessing the impact of training sample selection on accuracy of an urban classification: a case study in Denver, Colorado." *International Journal of Remote Sensing* 35.6 (2014): 2067-2081.

Key, Thomas, et al. "A comparison of multispectral and multitemporal information in high spatial resolution imagery for classification of individual tree species in a temperate hardwood forest." *Remote Sensing of Environment* 75.1 (2001): 100-112.

Kim, Minho, Marguerite Madden, and Timothy A. Warner. "Forest Type Mapping using Object-specific Texture Measures from Multispectral Ikonos Imagery." *Photogrammetric Engineering & Remote Sensing* 75.7 (2009): 819-829.

Kindermann, Ross, and James Laurie Snell. *Markov random fields and their applications*. Vol. 1. Providence, RI: American Mathematical Society, 1980.

Lu, Dengsheng, and Qihao Weng. "A survey of image classification methods and techniques for improving classification performance." *International journal of Remote sensing* 28.5 (2007): 823-870.

Mather, Paul, and Brandt Tso. *Classification methods for remotely sensed data*. CRC press, 2009.

Marceau, Danielle J., et al. "Evaluation of the grey-level co-occurrence matrix method for land-cover classification using SPOT imagery." *IEEE Transactions on Geoscience and Remote Sensing* 28.4 (1990): 513-519.

Max Kuhn. Contributions from Jed Wing, Steve Weston, Andre Williams, Chris Keefer, Allan Engelhardt, Tony Cooper, Zachary Mayer and the R Core Team (2014). caret: Classification and Regression Training. R package version 6.0-24. http://CRAN.R-project.org/package=caret

Moser, Gabriele, and Sebastiano B. Serpico. "Contextual high-resolution image classification by Markovian data fusion, adaptive texture extraction, and multiscale segmentation." *Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International*. IEEE, 2011.

Myint, Soe W., et al. "Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery." *Remote sensing of environment* 115.5 (2011): 1145-1161.

Ouma, Yashon O., T. G. Ngigi, and R. Tateishi. "On the optimization and selection of wavelet texture for feature extraction from high‐resolution satellite imagery with application towards urban‐tree delineation." *International Journal of Remote Sensing* 27.1 (2006): 73-104.

Peleg, Shmuel. "A new probabilistic relaxation scheme." *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 4 (1980): 362-369.

Pu, Ruiliang, and Shawn Landry. "A comparative analysis of high spatial resolution IKONOS and WorldView-2 imagery for mapping urban tree species."*Remote Sensing of Environment* 124 (2012): 516-533.

Puissant, Anne, Jacky Hirsch, and Christiane Weber. "The utility of texture analysis to improve per‐pixel classification for high to very high spatial resolution imagery." *International Journal of Remote Sensing* 26.4 (2005): 733-745.

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Richards, J. A., D. A. Landgrebe, and P. H. Swain. "Pixel labeling by supervised probabilistic relaxation." *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 2 (1981): 188-191.

Roberts, Dar A., et al. "Spectral and structural measures of northwest forest vegetation at leaf to landscape scales." *Ecosystems* 7.5 (2004): 545-562.

Ruiz, L. A., A. Fdez-Sarría, and J. A. Recio. "Texture feature extraction for classification of remote sensing data using wavelet decomposition: a comparative study." *20th ISPRS Congress*. 2004.

Torgo, L. (2010). Data Mining with R, learning with case studies. Chapman and Hall/CRC. URL: http://www.dcc.fc.up.pt/~ltorgo/DataMiningWithR

Toussaint, Godfried T. "The use of context in pattern recognition." *Pattern Recognition* 10.3 (1978): 189-204.

V. Subcommittee, F. Geographic, and D. Committee, "National Vegetation Classification," *Environ. Manage.*, vol. 2008, no. February, pp. 703–15, 2008.

Weng, Qihao, ed. *Advances in environmental remote sensing: sensors, algorithms, and applications*. CRC Press, 2011.

Woodcock, Curtis E., and Alan H. Strahler. "The factor of scale in remote sensing." *Remote sensing of Environment* 21.3 (1987): 311-332.

Wulder, Mike, K. Olaf Niemann, and David G. Goodenough. "Local maximum filtering for the extraction of tree locations and basal area from high spatial resolution imagery." *Remote Sensing of environment* 73.1 (2000): 103-114.

Wulder, Michael A., Werner A. Kurz, and Mark Gillis. "National level forest monitoring and modeling in Canada." *Progress in Planning* 61.4 (2004): 365-381.

Wulder, Michael A., et al. "High spatial resolution remotely sensed data for ecosystem characterization." *BioScience* 54.6 (2004): 511-521.

Xiaoxia, Sun, Zhang Jixian, and Liu Zhengjun. "A Comparison Of Object-Oriented And Pixel-Based Classification Approachs Using Quickbird Imagery." (2005).

Yu, Qian, et al. "Object-based detailed vegetation classification with airborne high spatial resolution remote sensing imagery." *Photogrammetric Engineering & Remote Sensing* 72.7 (2006): 799-811.

Zenzo, Silvano Di, et al. "Gaussian maximum likelihood and contextual classification algorithms for multicrop classification." *Geoscience and Remote Sensing, IEEE Transactions on* 6 (1987): 805-814.

Zortea, Maciel, Michaela De Martino, and Sebastiano Serpico. "A SVM ensemble approach for spectral-contextual classification of optical high spatial resolution imagery." *Geoscience and Remote Sensing Symposium, 2007. IGARSS 2007. IEEE International*. IEEE, 2007.