

Supervised classification of civil air patrol (CAP)

Elena Sava¹  · Laura Clemente-Harding^{1,2} · Guido Cervone^{1,3}

Received: 10 June 2016 / Accepted: 26 November 2016
© Springer Science+Business Media Dordrecht 2016

Abstract The mitigation and response to floods rely on accurate and timely flood assessment. Remote sensing technologies have become the de facto approach for observing the Earth and its environment. However, satellite remote sensing data are not always available, and it is crucial to develop new techniques to complement them with additional sources. This research proposes a new methodology based on machine learning algorithms to automatically identify water pixels in Civil Air Patrol (CAP) aerial imagery. Specifically, a wavelet transformation is paired with multiple classifiers to build models that discriminate water and non-water pixels. The learned classification models are first tested against a set of control cases and then used to automatically classify each image separately. Lastly, for each pixel in an image, a measure of uncertainty is computed as a proportion of the number of models that classify the pixel as water. The proposed methodology is tested on imagery collected during the 2013 Colorado flood.

Keywords Supervised classification · Remote sensing · Damage assessment · Natural hazards · Flood · Civil air patrol

1 Introduction

The study of environmental hazards and mitigation of their effects are essential for the development and sustainment of our society. Over the past decade, thousands of events have caused significant damage to people, property, and the environment. Some of these events occurred in statistically rare areas and with unprecedented strength.

✉ Elena Sava
esava@psu.edu

¹ Geoinformatics and Earth Observation Laboratory, Department of Geography and Institute for CyberScience, The Pennsylvania State University, State College, PA, USA

² Geospatial Research Laboratory, Engineer Research and Development Center, Alexandria, VA, USA

³ Research Application Laboratory, National Center for Atmospheric Research, Boulder, CO, USA

Remote sensing is the de facto technique for observing the Earth and its environment as it makes it possible to acquire information about Earth from space (Joyce et al. 2009). In particular, the coupling of remote sensing and geographic information systems (GIS) have greatly facilitated the operation of flood mapping and risk assessment (Nagasawa 2014; Schumann et al. 2009; Vijayaraghavan et al. 2012). This is due to the availability and high spatial and spectral resolution of remote sensing data and the visualization of flooding using GIS. When combined, these two aspects create the potential to further analyze and estimate probable damage due to flooding (Sanyal and Lu 2004; Mason et al. 2007; Sakamoto et al. 2007).

The idea of extracting features from imagery using statistical and structural approaches can be traced back to the early 1970s (Haralick 1979). Researchers first started using 80 m Landsat data to identify flooded areas in Iowa and in the Mississippi River Basin. In particular, the near-infrared band available on many remote sensing platforms has been found to be particularly suitable for identifying water because of its strong absorption of energy in the 700 nm–1.3 mm range of the electromagnetic (EM) spectrum (Smith 1997; Rango and Anderson 1974; Rango and Salomonson 1974; Moody et al. 2014).

Employing a similar understanding of water's low spectral response in the optical range (visible NIR), higher spatial resolution platforms have been used to improve the capability of water identification and flood delineation. Platforms, such as Landsat Thematic Mapper (TM), coupled with maximum-likelihood classification have improved the identification and extraction of water bodies and increased the precision with an overall accuracy of 97% (Frazier et al. 2000). Landsat TM is unable to penetrate through thick clouds and dense canopy resulting in an underestimation of flooded areas. The fusion of satellite imagery with digital elevation model (DEM) data has led to the improvement of flood extent detection under densely vegetated regions. (Wang et al. 2002). Furthermore, the Advanced Very High Resolution Radiometer (AVHRR) imagery has been paired with DEM, drainage networks, population density, and other GIS data to generate flood hazard maps. Additionally, these hazard maps have been used to plan new infrastructures with consideration for potential flood damage (Islam and Sado 2002). This allows for better identification and understanding of water depth and comparison of damage between multiple regions.

While these techniques work well when proper data are available, satellite data, as previously mentioned, are not always present or may only provide partial or incomplete coverage of disaster areas. Thus, despite the massive quantities of data available, gaps in coverage are often present due to satellite revisit time limitations, cloud cover, vegetation canopy, and other factors that hinder the observation of flooded areas. For these reasons, it is crucial to develop new techniques in order to produce timely flood assessments before, during, and after an event.

Accurate and actionable damage assessment requires integrated information from many different sources. A major concern for emergency management strategies is the availability of reliable and accurate hazard maps. The presence of accurate mapping can greatly assist a respondent's decision-making process regarding the direction of aid to areas with the greatest need. Several integrated approaches of remote sensing with GIS have led to better management strategies, more accurate flood hazard mapping, and flood risk assessment (Hadjimitsis 2007; Lamovec et al. 2013; Schnebele et al. 2014a; Cervone et al. 2016).

The integration of information extracted through non-authoritative and remote sensing datasets coupled with other datasets provides tremendous potential for identification, monitoring, and assessment of disasters. For instance, crowdsourced photographs and volunteered geographic data were fused using geo-statistical interpolation to create an

estimation of flood damage in New York City after hurricane Sandy on October 2012 (Schnebele and Cervone 2013; Schnebele et al. 2014b). During the large fire series that took place in Southern California from July 2007 to May 2009, local residents shared observations using social media outlets such as Flickr or Twitter, local news Web sites, and Google maps as a way to localize the information they contributed (Goodchild and Glennon 2010).

Civil Air Patrol (CAP) imagery is one data source available to augment the use of satellite remote sensing. The CAP is a congressionally funded, nonprofit corporation functioning as an auxiliary to the United States Air Force. The CAP conducts a variety of missions in support of federal, state, local, and non-governmental entities, including search and rescue, disaster relief support, and aerial reconnaissance for homeland security.¹ However, working with such imagery poses challenges because the data are usually collected using conventional digital cameras that only capture information in visible light. Therefore, each image contains only three spectral channels in red-green-blue (RGB) color space. This coarse sampling of the visible spectrum limits the identification and classification of pixels performed on RGB imagery making it difficult to extract features.

This is particularly true in the CAP data used in this study because water has very different signatures in the images. These variations are due to different collection times, orientation, body of water observed (e.g. river vs lake vs flooded urban area). The differences occur both in the spectral characteristics of the water pixels and the geometrical signature of the water areas. On the contrary, advanced multispectral sensors available through various remote sensing platforms are structured to collect data in different parts of the electromagnetic (EM) spectrum usually ranging from ultraviolet (UV) to emissive infrared (IR). CAP images range from oblique to rooftop images, and therefore, the methodology can be applicable to both cases as shown by Cervone et al. (2016).

This work addresses this technical challenge with a new computational methodology based on the use of artificial intelligence and machine learning classification algorithms to automatically identify water pixels in CAP imagery associated with the 2013 Boulder, CO floods. More specifically, it proposes a multi-strategy approach based on maximum likelihood and supervised machine learning classification coupled with wavelet transformations. The novelty of this approach lies in the use of freely available data documenting flooding and damage in real time in order to draw an effective, integrated strategy for the use of CAP imagery and satellite remote sensing in case of emergencies. An additional benefit of the current research is that the source code and executable to perform the water extraction are released free of charge. While several computer vision algorithms have been published to perform related feature extraction, their implementations are sometimes not readily available, or might need to be modified to work with CAP imagery.

1.1 Related work in image classification

The synoptic, repetitive nature of satellite remote sensing data coupled with GIS data has allowed response teams and specialists to monitor the progression of a flood event in near real time (Vijayaraghavan et al. 2012). However, satellite remote sensing technologies do not always provide clear and complete data coverage of an area. In the absence of satellite imagery and NIR information, water identification and extraction can be derived from standard RGB images taken by airplanes, or by using the emerging availability of Unmanned Aerial Vehicle (UAVs). Feature identification in images is an active subject in

¹ <http://www.gocivilairpatrol.com>.

the computer vision community. Two main approaches are being investigated: pixel based and object based. In pixel based analysis, the characteristics of the individual pixels are used, while in object based the image is transformed into a series of semantic concepts.

1.2 Pixel based image analysis

Several methods have been proposed to semantically segment RGB images using multiple scales (Mallat 1989; Csurka and Perronnin 2011; Pantofaru et al. 2008; Shotton et al. 2006). Most multiscale segmentation methods combine two principles: first, a recognition process that assigns each pixel to a class and second, a regularization process that groups pixels together which are contained in the same feature (Salamati et al. 2014).

RGB imagery can be decomposed into a family of derived images. Due to the flexibility to explore fine details from high- and medium-frequency subdivisions of an image, wavelet decomposition has gained popular applications in the field of signal and image processing and classification (Baaziz et al. 2010; Li et al. 2014; Huang and Aviyente 2008). Wavelet sub-bands of texture for automatic clustering of image classification and segmentation have also been used (Arivazhagan and Ganesan 2003; Aujol et al. 2003). By using a wavelet transformation algorithm and analyzing the textures characterized by the energy of wavelet coefficients based on each sub-band decomposition, a supervised classification model can then be applied to classify the images. The algorithm produces segmentations of an entire image according to prior known texture parameters by using level sets created according to wavelet coefficients and the interaction between each other (Aujol et al. 2003). In more recent research, wavelet transforms have been investigated for multiscale time–frequency signal analysis. For example, the integration of wavelet-based feature with pixel-wise-based classification has been used to improve the classification performance in noise-corrupted environments for hyperspectral data (Li et al. 2014). Additionally, a wavelet transformation and a k-means algorithm are used as an approach for image retrieval by combining color and texture information of features (Bai et al. 2012).

Texture is an important feature for the analysis of many types of images and may be seen in all images whether obtained from an aircraft or satellite platform. Various approaches have been suggested for texture retrieval and mainly vary by the method used for extracting textural features. Texture extraction methods may be defined into four categories: (1) statistical, (2) structural, (3) model based, and (4) transformation based (Materka and Strzelecki 1998; Bharati et al. 2004). Transformation methods of texture analysis such as Fourier, Gabor, and wavelet transforms represent an image in a space whose coordinate system has an interpretation that is closely related to the characteristics of a texture (such as frequency or size) (Materka and Strzelecki 1998). Methods based on the Fourier transform perform poorly in practice due to its lack of spatial localization (Skarbnik 2009). Gabor filters provide means for better spatial localization. However, the use is limited in practice because there is usually no single filter resolution at which one can localize a spatial structure in natural textures. This is known as the filter bank design, which involves the selection of a suitable number of filters at different orientations and frequencies (Bianconi and Fernández 2007). Compared to the Gabor transform, the wavelet transforms presents several advantages: varying the spatial resolution allows it to represent textures at the most suitable scale, there is a wide range of choices for the mother wavelet function, each providing advantages for texture analysis in a specific application (Materka and Strzelecki 1998).

Additional methods using filters have been applied for feature extraction and classification. Gabor filters are used as a texture extraction method, and classification is performed

with maximum likelihood for the classification of aerial and satellite digital images (Recio et al. 2005). A wavelet-based texture retrieval using generalized Gaussian density and Kullback Leibler distance has also been presented (Do and Vetterli 2002). A color texture image classification based on hue–saturation–value (HSV) color space wavelet transform and motif co-occurrence matrix for feature extraction has also been used (Chang et al. 2010). A support vector machine (SVM) is then applied to learn and classify feature classes on texture images based on extracted information. Another technique uses wavelet coefficients in low pass bands with a structure-based neural network to develop a supervised classification based on the distribution of histograms from wavelet coefficients (Zou et al. 2008).

1.3 Object based image analysis (OBIA)

Object based image analysis (OBIA) is a semi-automated method that uses groups of pixels that represent concepts (e.g. buildings, trees) to classify objects from high-resolution remotely sensed data. In an OBIA, the outcome of segmentation is directly related to several adjustable criteria such as scale, color, and shape as defined by the user. In particular, the scale and shape parameters are critical as a large scale leads to bigger objects being classified and a small scale may not be representative. Adjusting the shape parameter also affects the overall fusion value as it is computed based on both spectral heterogeneity and shape heterogeneity. The result of OBIA is also dependent on additional metrics such as measures of texture, length, and spatial relationship to other near and far neighboring objects.

OBIA software fuses the disparate datasets, usually Light Detection and Ranging (LiDAR) and high-resolution multispectral data, and then classifies objects based on spectral and geometric characteristics. Mueller et al. (2004) employed an object based segmentation with special focus on shape information to extract large man-made objects, mainly agriculture fields in high-resolution panchromatic data. It is important to note that there is no standardized or widely accepted method available to determine the optimal scale for all types of applications, areas with different environmental and biophysical conditions, and different kinds of remotely sensed images (Myint et al. 2011).

OBIA algorithms make extensive use of image segmentation to split an unclassified image into separate regions (Im et al. 2008; Lee and Warner 2006; Stow et al. 2008). From an algorithmic perspective, it is generally divided into four categories: (1) point based, (2) edge based, (3) region based, and (4) combined (Schiewe 2002). A series of studies have exhibited methods for comparing different techniques of sample segment objects against corresponding reference objects as well as for assessing segmentation accuracy (Weidner 2008; Winter 2000). Using trial-and-error tests along with an index called the ‘comparison index,’ Möller et al. (2007) developed an approach to perform a segmentation scale that is close to optimal. The approach used a segmentation procedure called fractal net evolution approach, implemented in the eCognition software, to generate objects at different scales (Batz and Schäpe 2000).

There are several advantages and disadvantages of various segmentation approaches as discussed by Muñoz et al. (2003), there has yet to be a segmentation algorithm that integrates regions and boundary information and thus research is still needed in the computer vision community and its applications. OBIA algorithms work best with multispectral imagery because the algorithm relies heavily on the spectral signature of the objects to be classified. Furthermore, OBIA relies heavily on distinctive geometrical characteristics of the objects (e.g. homes tend to be rectangular and of comparable size). In

our case, the spectral information is very limited, and water areas do not have predictable geometric patterns.

2 Case study: 2013 boulder flood

This work analyzes data for the 2013 flood that occurred in Colorado, particularly affecting Boulder and the neighboring counties. Boulder is located near the foothills of the Rocky Mountains approximately 50 km northwest from Denver. This unique location has made Boulder the number one flash flood community in the state of Colorado as major flash floods have developed in recent years (Pittman 2014). In the last 15 years alone, the City of Boulder has spent more than \$45 million dollars on infrastructure resilient to floods and sustainability projects (Pittman 2014). However, the 2013 flood was not the flash flood that officials had been preparing for. It was an unprecedented event where Colorado received a year's worth of precipitation in only eight days, exceeding 45 cm of rain.

The 2013 Colorado flood propagated quickly and with little warning. A catastrophic storm system trapped moist monsoonal air over the Colorado Rocky Mountain Front Range producing heavy rainfall over a one week period. From September 11, 2013, to September 15, 2013, a record-breaking precipitation occurred along the Colorado Front Range impacting 18 counties and resulted in severe floods throughout many parts of Colorado from Larimer and Weld Counties southward to El Paso County. Boulder County was worst hit with 25 cm of rain on September 12 and up to 43 cm by September 15, 2013. With precipitation exceeding 50 cm in parts of Boulder County, the area experienced catastrophic flooding, property destruction totaling over \$2 billion in damage, and a tragic loss of life. The effects of the storm were not only felt in Colorado. Heavy precipitation and flooding occurred in several neighboring states such as Nebraska, New Mexico, and Wyoming (Gochis et al. 2014).

2.1 CAP data

Civil Air Patrol data were downloaded from the US Geological Survey Hazards Data Distribution System Explorer (USGS HDDS system²). The HDDS is a unique Web-based interface that enables users to search for satellite, aerial imagery, and other geospatial datasets available in near real time. The Web portal is designed to assist in disaster response by providing imagery and documents acquired before and after an event.

A series of Unix Bash and Python commands were used to automatically select and download a total of 6444 images. Images covered the flooding period of September 14–22, 2013. Each image contains metadata providing information on the camera type used to capture the image and several spatiotemporal characteristics (e.g. date/time of acquisition, lat/long coordinates). Based on this metadata, Table 1 shows the number of CAP images collected per day over Boulder County.

The image classification was performed using the RStudio³ parallel package using six cores on a MacPro. Because each image can be analyzed independently, the proposed solution is 'embarrassingly parallelizable' and thus it is exceptionally well suited for multi-core/multi-node architectures.

² <http://hdds.usgs.gov>.

³ <http://www.rstudio.com>.

Table 1 Total number of CAP aerial images acquired between September 14, 2013, and September 22, 2013

Saturday 9/14	Sunday 9/15	Monday 9/16	Tuesday 9/17	Wednesday 9/18	Thursday 9/19	Friday 9/20	Saturday 9/21	Sunday 9/22	Total
1788	52	1852	1020	82	705	110	0	834	6444

3 Methodology

The goal is to classify each CAP image in the dataset automatically with minimal expert supervision and input. The methodology consists of three main steps:

- a. Manual: Defining a very small, selected number of training regions as examples and counterexamples of flooded areas.
- b. Automatic: Multiple classifiers are run in parallel to build a model describing water and non-water regions. Each classifier is applied either on the original images or on a transformation and classifies all pixels in each image into either water or non-water pixels; thus, no pixels remain unclassified. This extracts relevant information such as texture or homogeneity. Among the possible transformations, wavelet decomposition proved to be a particularly effective method. Wavelets are applied to the grayscale representation of the original RGB images, and the computed wavelet scale coefficients are used as training regions.
- c. Automatic: Each pixel classification in the images is performed by combining the results of the models. A measure of confidence is computed by ranking how many of the models classified a pixel as water.
- d. After each image is classified, a despeckling filter is applied to remove small defects that may be present in an image. This is commonly applied to RGB images to remove color channel noise (Fig. 1).

This approach utilizes freely available CAP data to draw an effective integrated strategy for near real-time damage assessment using only RGB information.

3.1 Wavelet transformation

Wavelet transformation is a mathematical function and signal processing technique that helps localize data by decomposing a signal into different frequency sub-bands.

For any given real-valued function ϕ with zero average $\int_{-\infty}^{\infty} \phi(t)dt = 0$, let

$$Wf(u, s) = \int \frac{f(t)}{\sqrt{(s)}} \phi\left(\frac{t - u}{s}\right) dt \tag{1}$$

be the real continuous wavelet transform of a function f . Since ϕ has zero mean, the previous integral measures the variation of f in a neighborhood of time (position) u of size proportional to the so-called scale factor $s > 0$. There are several other wavelet functions ϕ that can be used for the transformation. The Haar wavelet was chosen because of its simplicity and computational efficiency (Amolins et al. 2007; Hein and Feklistova 2011).

The wavelet transformation technique can be applied to a variety of fundamental signal processing tasks such as enhancing images and sounds recordings by removing noise or for

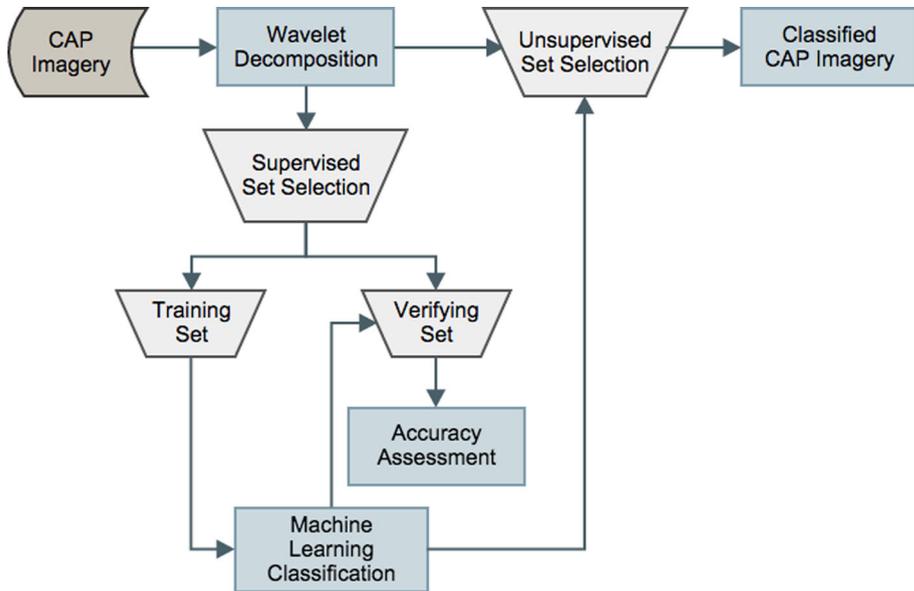


Fig. 1 Methodology flowchart: notice that wavelet transformation is applied to all the images regardless of their use in the main classification as training or testing

data reduction by compression (e.g. jpeg compression) (Baaziz et al. 2010). When applying a wavelet transformation, data are compressed and transformed based on multi-resolution analysis (Huang and Aviyente 2008). This consists of the decomposition of an image into subimages containing different fractions of signal value known as wavelet coefficients (Fig. 2).

This preserves the relative distance between objects at different levels and allows natural clusters in the data to be easily identified. In practice, the transformation divides the data equally at each iteration (Han et al. 2006). First, it applies a data smoothing technique such as sum or weighted average, and then, it performs a weighted difference, which allows the details of features to be identified. The output results in a smoothed or low-frequency version of the original data by removing the high-frequency information. This procedure occurs recursively by applying the transformation on the image resulting from the previous iteration and continues until a termination condition is met (Arivazhagan and Ganesan 2003; Amolins et al. 2007).

3.2 Supervised image classification

Supervised classification is an operation that uses an inductive reasoning process where a set of rules are learned based on labeled pixels in order to classify new unlabeled pixels (Dougherty 2012). This method consists of two steps: first, the user manually selects a series of examples and counterexamples to form the training sets. The user tunes the model on the training data using a machine learning algorithm to perform a correlation task. As shown in Fig. 3, these regions represent areas for each class of interest to be classified (e.g. water vs not water).

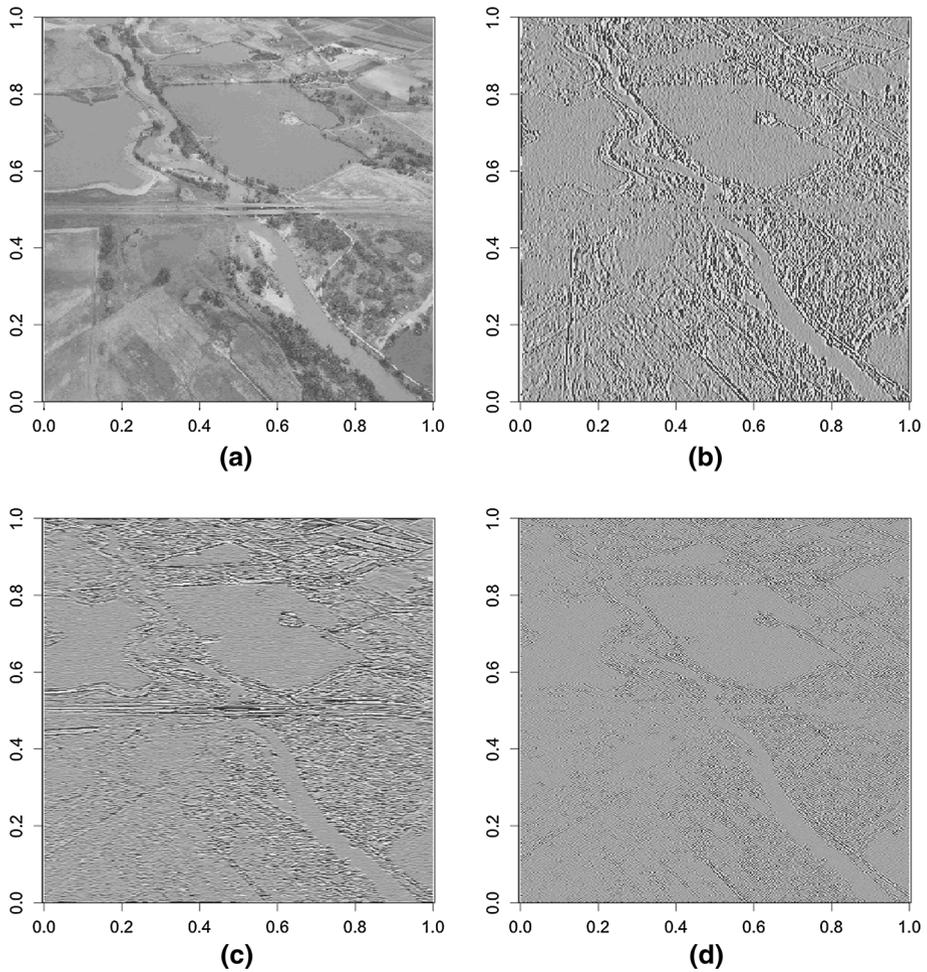


Fig. 2 Illustration of the wavelet transformation process. The original image is displayed in the upper left quadrant (a). The other three quadrants represent the decomposition of the wavelet transformation. The first wavelet coefficient is represented by figure (b), the second wavelet coefficient is shown in figure (c), and third wavelet coefficient in figure (d)

Second, the classification algorithms use the positive training samples to create a generalization of spectral and wavelet coefficient properties for each of the classes. This generalization contains discriminant characteristics of the features of the input data, and they are used to classify new pixels the model has never seen before into one of the classification classes.

Supervised classification has the advantage of automatically generating models used for classifying thousand of images from a very limited number of training regions. This allows human intervention to be kept to a minimum (e.g. identification of 40 regions in a set of ten images). However, the classification of water pixel is not trivial. In fact, color information alone is not sufficient to discriminate between classes, as water pixels vary in RGB color as



Fig. 3 Sample image used to select examples and counterexamples. *Rectangles in black* represent water (positives) while the *red rectangles* represent regions without water (negative)

shades of blue, green, and brown. Furthermore, soil can have similar color characteristics, which would lead to misclassification if only regression algorithms were used.

Thus, because supervised algorithms learn from features present in all images (e.g. color or wavelet coefficient), the learned models can be applied to images that were not part of the original training set.

- *Decision Trees*

Decision trees use a divide and conquer paradigm to generate a hierarchical set of attribute-value decisions to discriminate between pixel classes defined in a set of training examples and counterexamples (Breiman et al. 1984; Quinlan 1993; Steinberg and Colla 1995). This training set is composed of ‘labeled’ data, where each element represents a pixel which has been previously classified into one of the output class (e.g. water or non-water). Each element has two parts: first, the output class and second, a vector of values representing the characteristics for the pixel, such as RGB values or wavelet coefficients.

The learned decision tree is then applied to all the unseen (unlabeled) data, which are automatically classified into one of the output classes. Starting from the root, a consecutive set of comparisons are made between a pixel and the learned decision tree to determine which pixel belongs in each class.

The main advantages of decision trees over other classification algorithms include the capability for highly dimensional data, resilience to noise, and computational efficiency to learn the tree and apply it to classify new examples. They also have very few input parameters which make them easier to optimize for a given problem. This work uses the regression tree library for R statistical language 1.0.36 package.

- *Maximum Likelihood*

Maximum likelihood classification (MLC) is one of the most commonly used supervised classification techniques. MLC builds a statistical model for each class summarizing the multi-dimensional mean and variance for each of the attributes. Pixel classification is made by assigning individual pixels to the class that minimizes the Euclidean distance between the values of the pixel and the MLC models. The computation of the models is based on the assumption that both the training data and classes display multivariate normal frequency distribution (Campbell and Wynne 2011).

3.3 Classification of pixels using learned models

Supervised algorithms, as described in Sect. 3.2, utilize the training set images in order to learn the models. The learned models are then applied to all 6424 images that are not part of the original training set. This task is fully automated and requires no additional human supervision. The classification of the entire dataset is geographically represented with a map showing areas most likely affected.

4 Results

4.1 Model learning

In order to test and validate the hypothesis that CAP imagery can provide actionable knowledge during flood emergencies, the proposed methodology is applied to all the available data described in Sect. 2.1.

The first step consists in learning a model that employs both decision tree and maximum likelihood classifiers. A small set of regions (44 for water examples and 40 as non-water counterexamples) is manually chosen from 20 randomly selected CAP images. Selecting representative training samples is a key step for supervised classification. The identified regions are homogeneous, rectangular in shape, and vary in size.

4.2 Validation

The learned model is validated on a small sample of manually selected pixels to assess prediction accuracy. In this work, both classifiers are weighted equally. Thus, an unseen pixel is classified as 100% water if both classifiers match, 50% if only one classifier matches, and 0% when neither of the classifiers match. Figure 4 depicts the results using ternary plots. A ternary plot shows the relationship between three variables on three separate axes, under the assumption that the sum of the variables is equal to 1 for all examples.

Figure 4 shows two ternary plots displaying the results of the tests performed for areas labeled as containing water (Fig. 4a) or not containing water (Fig. 4b). Each point represents one of the areas used for the tests. The axis shows the percentage of pixels in the areas classified as no water NW (left vertex), water by a single classifier C1 (top vertex), or C2 (right vertex) water by both classifiers. The three vertices of the ternary plots correspond to C1, C2, and NW, and each represents 100% accuracy in the classification of pixels. Each of the points represents the training sets.

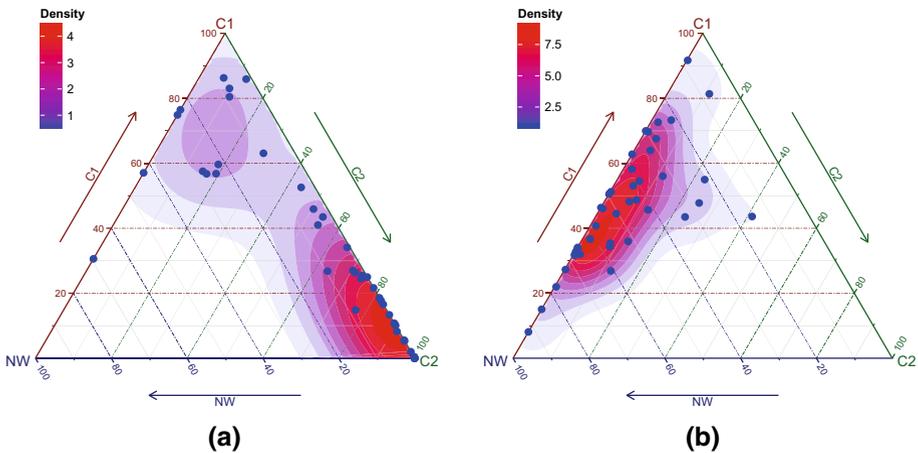


Fig. 4 Ternary plot showing the results of the tests performed for areas labeled as containing water (a) or not containing water (b). Each point represents one of the areas used for the tests. The axes show the percentage of pixels in the areas classified as no water (left vertex), water by a single classifier (top vertex), or water by both classifiers (right vertex)

The left graph displays how accurately the classifiers predict water when applied only to the training regions where water was present (positives regions, a.k.a. examples). Ideally, all regions would be classified in the lower right (C2) as this indicates that both models classify these areas as water only pixels. While the best results are achieved by points lying exactly on the right vertex, any classification along the right base (C2) is acceptable because it indicates that the water areas are identified as water by at least one classifier. Figure 4a shows that the classifiers have a high accuracy because most water areas are correctly classified. Greater than half of all the positive training regions were able to be correctly classified with more than 85% accuracy. Twelve training regions were classified as water by at least one classifier and only one region was classified incorrectly.

Figure 4b shows the accuracy of the classifiers when applied to the training regions in which water was not present (negative regions, a.k.a. counterexamples). In this case, the ideal classification lies on the left vertex (NW) because it represents that neither of the models identified water and agree that water is not present inside the training regions. While it very seldom happens that both classifiers identify the non-water areas as water, it is more common that one of the two classifiers does. This indicates that although the methodology is generally sound, there is an overall uncertainty associated with the classification of water by a single classifier.

4.3 Image classification

Following the performance assessment, the learned model is applied to classify all unlabeled pixels in the remaining images. The result of the classification of each image is a discrete output indicating the presence or absence of water within each image. Furthermore, because multiple classifiers are employed, a measure of confidence can be derived that is proportional to the cumulative number of classifications. Though only two classifiers are used in the current tests, the methodology can scale up to including a much larger number of classifiers which will lead to a finer measure of confidence.

Figure 5 shows a summary of the classification of all 6444 images using the described methodology. The three vertices in the figure represent the percentage of no water (left vertex), water classified by single classifier (top vertex), or water classified by both classifiers (right vertex). Each point inside the triangle shows how each of the 6444 images is classified as containing a percentage of pixels classified as no water, water by a single classifier, or water by both classifiers. Therefore, an image can be classified as 100% water by either model, C1 or C2, or no water (NW) present at all. For example, if a point were to be selected near the center of the graph at an intersection of RGB, it would represent a ratio of the three components (e.g. 40% blue, 40% red, 20% green). The density, each point representing an individual CAP image, is indicated by a choropleth gradient. The higher density of cases along the NW-C1 axis indicates that most images had a mix of no water or water classified by a single classifier. Relatively few images were classified as containing water by both classifiers. This is consistent with the observation that most images in the data contained a very modest amount of water (e.g. vegetated areas).

Each image is assessed based on the amount of water present. In aggregation, the amount of water present in a classified image can be geographically represented to generate a damage assessment map (Fig. 6). The situational awareness map provides an overview of the percent of pixels classified by both classifiers. The map shows the classification for the overall area in which images were available, including the case study area of Boulder County.

Points represented in dark blue depict images containing high value of pixels classified as water. For example, a responder may observe a large cluster of images classified with greater than 40% water north from the city of Denver. From the visual representation of

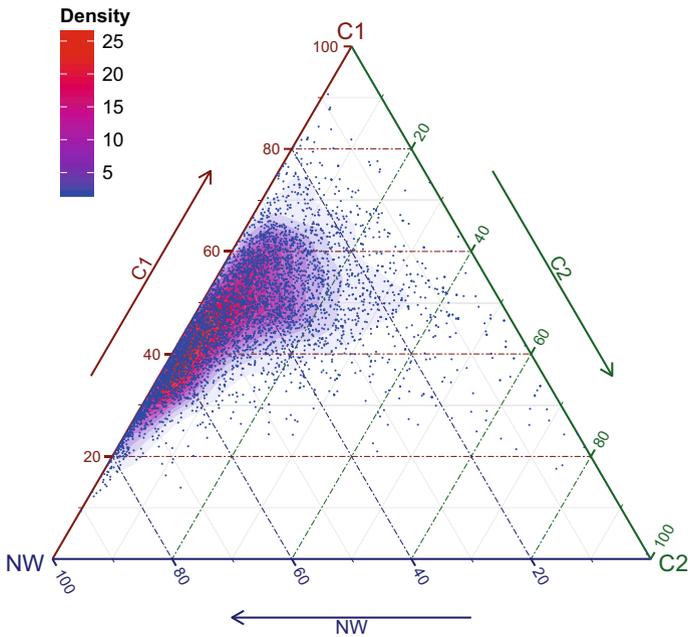


Fig. 5 Ternary plot showing the overall classification of all 6444 images. Each of the points represents an image. The axis indicates percentage of the image classified as no water (left vertex), or water by a single classifier (top vertex), or water by both classifiers (right vertex)

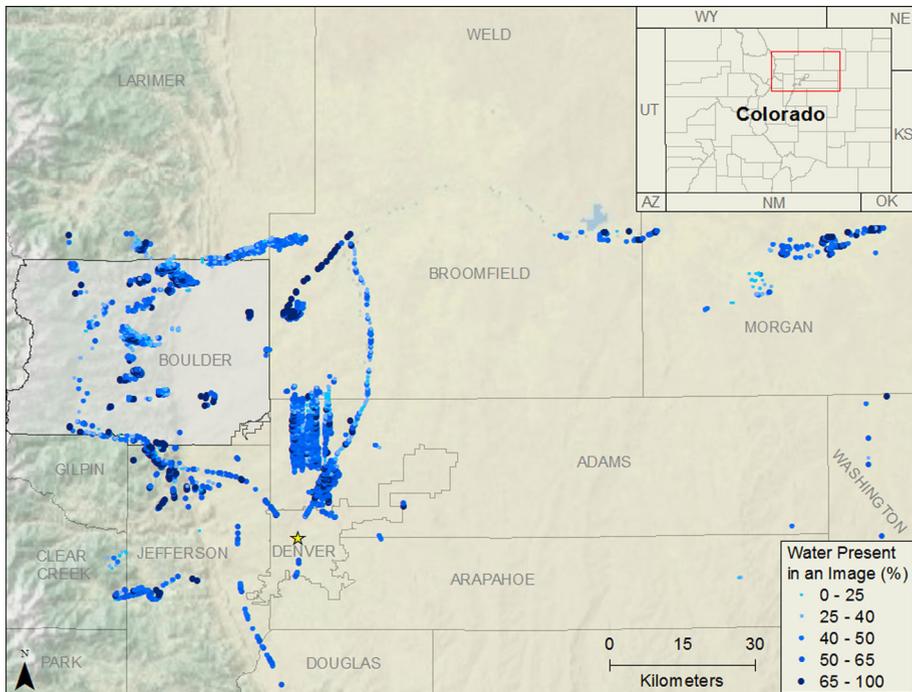


Fig. 6 Damage assessment map based on the classification of all 6,444 images for Boulder and surrounding counties. Each image is represented with a circle ranging from light to dark blue. Light blue indicates a low number of water classified pixels, and dark blue indicates a high number of water classified pixels

water presence/absence, emergency responders can quickly assess the areas that are likely to contain flooding damage and require immediate assistance. This knowledge can be used in near real time for situational awareness and to prioritize available resources.

Sample results of accurate water classification in RGB images are shown in Fig. 7. Here, the models are able to identify the presence of water regardless of the different color shades of the water. Furthermore, a clear contrast is shown between water areas that are continuously present throughout time and areas identified as water due to the damage caused by the Colorado flood. A distinction is made between roads and water allowing responders to have a complete understanding of road conditions.

Throughout the research process, multiple issues with the classification of RGB imagery were identified. These include the presence of clouds, shadows from physical features, and the acquisition angle. Sample results of a less accurate CAP classification are displayed in Fig. 8. Although the models are able to accurately identify areas where water is present, a high amount of uncertainty is also present. On the top images, one of the classifiers identifies forested areas as potential water pixels (teal), while the image below shows the misclassification of water due to the presence of shadows.

Overall, the proposed methodology adequately classified water regions providing a general idea of damaged or flooded regions within an image. Results show that this methodology is promising for classifying water and calculating damage assessment in RGB images. As shown by Fig. 6, a situational awareness map can be generated to provide

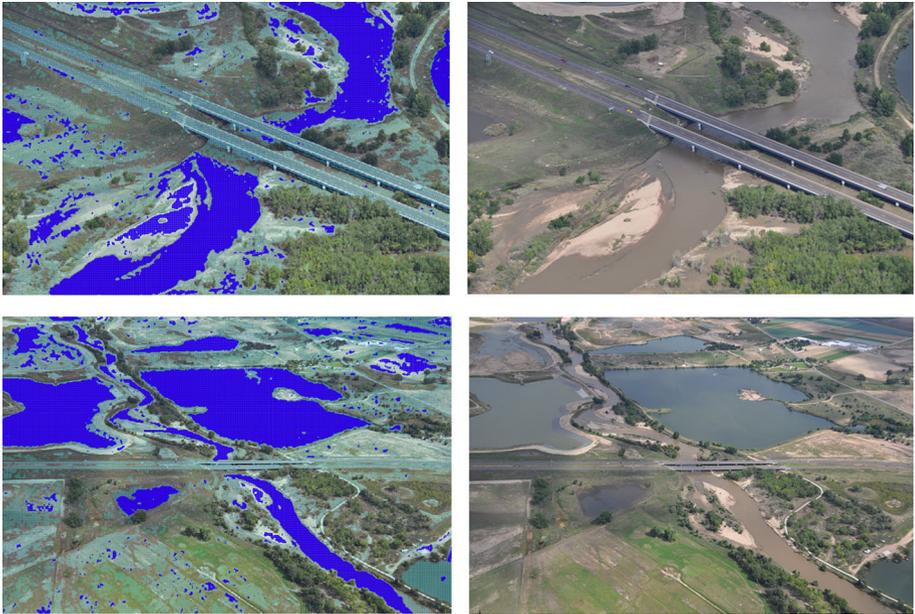


Fig. 7 Sample classification of water pixels by both classifiers (*blue*) or by a single classifier (*teal*). The algorithm was able to correctly classify most of the water areas

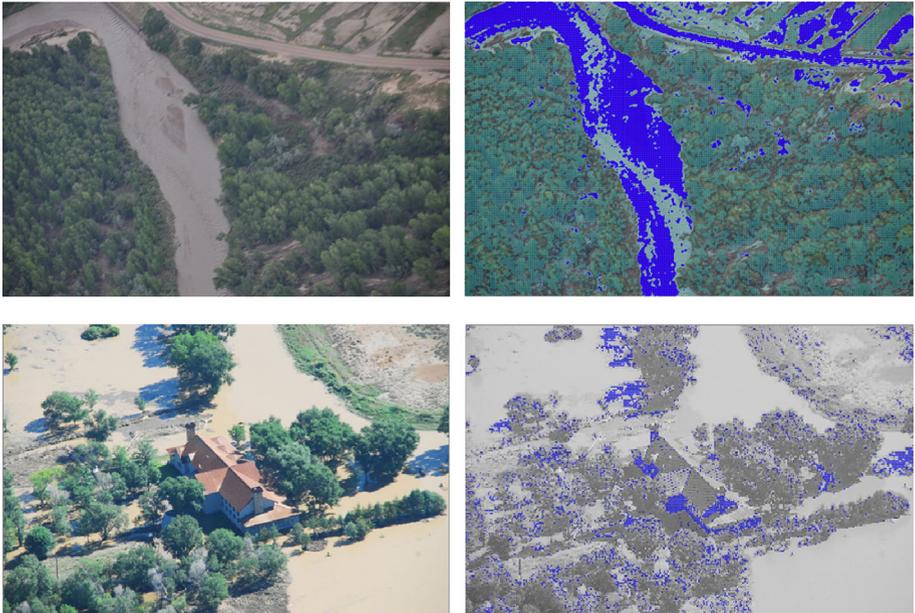


Fig. 8 Sample classification of water pixels by both classifiers (*blue*) or by a single classifier (*teal*). The algorithm was not able to correctly classify all of the water areas

an overview of the percent of water coverage in each CAP image and, in turn, depict regions where a high amount of water or damage may be present.

4.4 Analysis of images with different shadows, pixel intensities, and color

The ability to correctly classify unseen images depends both on the quality of the training areas and the ability of the algorithms to learn the attribute-value relations that characterize water pixels. Of particular interest is the classification of images that are significantly different from the training set, which have different shadows, pixel intensity, and colors. Figure 9 shows two set of classified images. Figure 9a and c is the original RGB CAP, and Fig. 9b and d is the same images classified according to the learned model.

The image classification shows that the algorithm is able to correctly identify water bodies in both scenes, albeit with some pixel misclassification. Figure 9b shows that the algorithm misinterprets the cloud shadows for water pixels (top right part of the image), similarly to what happened in Fig. 8. Figure 9a and c shows how the methodology performs well on rare images that are very different than those used for training. Specifically, these images are very dark; however, the difference in pixel intensity does not cause water misclassification. The wavelet transformation is able to extract texture information regardless of the color intensity, making the methodology resilient and able to classify data acquired under different conditions.

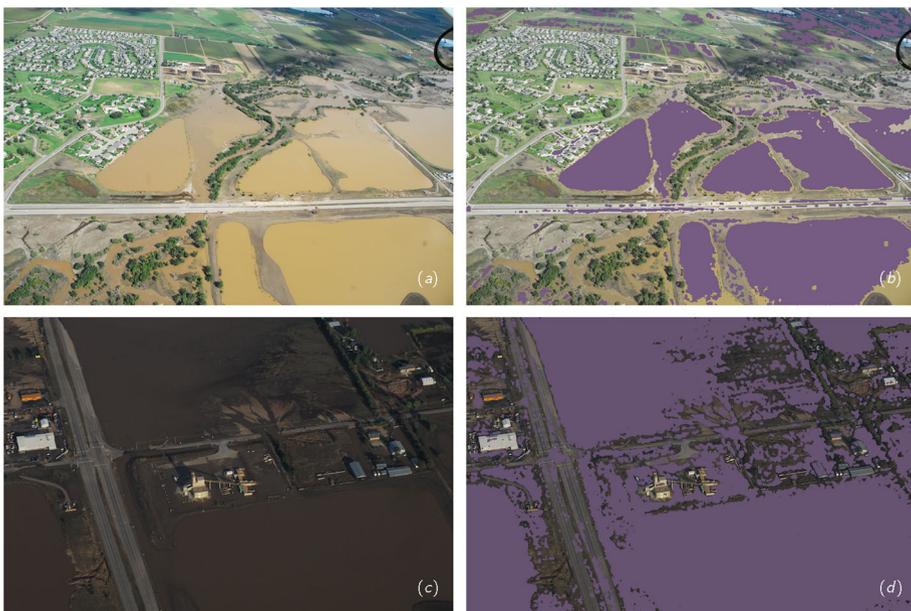


Fig. 9 Selected CAP images used to represent the methodology on different images with respect to shadows, pixel intensities, and *color*. Images (a) and (c) are the original RGB CAP, and images (b) and (d) are the classified CAP

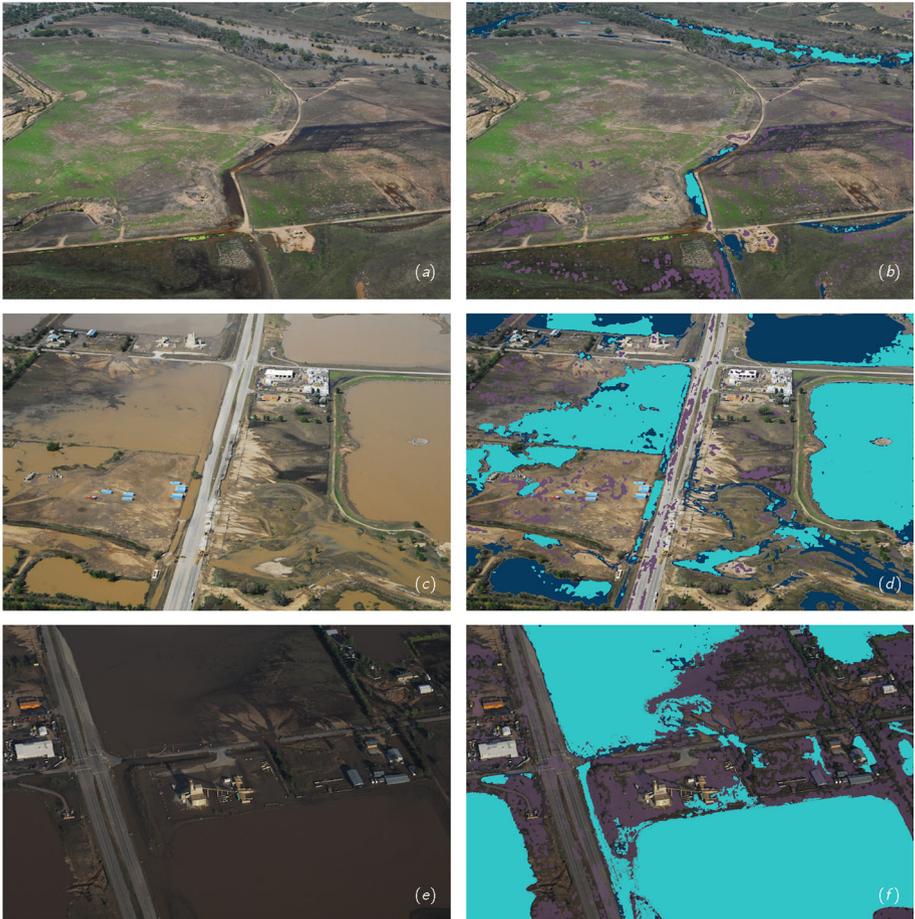


Fig. 10 Selected CAP images used to represent low, medium, and high accuracy of the algorithm. Images (a), (c), and (e) are the original RGB CAP. Images (b), (d), and (f) are the same images representing the control pixels (*dark blue*), classified pixels (*purple*), and the intersection of the two (*teal*)

4.5 Quantitative assessment of the results

The methodology was assessed quantitatively using a set of 11 scenes acquired in Bloomfield county, located NE of Boulder in the center of Fig. 6. Using the selected set, the intersection of water between the original RGB CAP image and classified image was calculated. The water pixels present in the original image (control pixels) were manually identified using visual interpretation and compared with the number of water pixels classified by the algorithm (classified pixels). The intersection (intersect pixels) is calculated as the common area between control and classified pixels. Additionally, the number of pixels misclassified as water (false positives) as well as the number of water pixels not identified (false negative) are computed. The total accuracy is computed as the ratio between the number of intersecting pixels and the control pixels in the RGB CAP images. The overall accuracy for the assessment over the 11 scenes is 67%, with values ranging between 33% and 98%. In most cases, the accuracy is higher than 67% (scene 6–11), and

Table 2 Predictive accuracy assessment for 11 selected scenes

Scene	Control pixels (X)	Classified pixels (Y)	Intersect pixels (Z)	False positives (Y-Z)	False negatives (X-Z)	Total accuracy (Z/X) (%)
1	5976	7275	1987	5288	3989	33
2	40,574	71,207	20,988	50,219	19,586	52
3	58,205	48,068	32,578	15,490	25,627	56
4	52,708	51,731	30,245	21,486	22,463	57
5	60,998	41,825	35,813	6012	25,185	59
6	65,972	62,367	45,821	16,546	20,151	69
7	61,170	51,335	43,242	8093	17,928	71
8	56,026	48,840	41,315	7525	14,711	74
9	49,731	81,247	43,316	37,931	6415	87
10	66,674	76,496	58,426	18,070	8248	88
11	74,257	104,948	72,564	32,384	1693	98
Average accuracy						67

the overall average is reduced by the single low score of 33%. Therefore, it can be concluded that the initial assessment shown in Fig. 6, suggesting that the set of 11 scenes acquired in Bloomfield county have a high percentage of water pixels, is true. Therefore, this map could have given an accurate initial assessment of regions with a higher amount of detected water. Overlaying this map with additional data of natural water bodies can lead to an assessment of which areas are most likely to have been impacted by the flood.

Three scenes were selected from the sample of 11 to represent instances of low, medium, and high accuracy as shown in Fig. 10. Figure 10a, c, and e is the original CAP images in RGB color. Each scene has a difference in color intensity, percentage of water pixels, and texture. Figure 10b, d, and f are CAP scenes with water shown with different colors. Dark blue represents the manually selected water control pixels (X), purple the pixels classified as water by the algorithm (Y), and teal the intersection (Z) of control and classified pixels ($X \cap Y$).

Out of the 6444 images that were used for this work many were images that contained minimal water and large vegetated areas. Figure 10 is an overall representation of the CAP images which may have a small amount of water, images in which the ground and water have the same color and contain a high amount of water, as well as rare images that have high pixel intensity where it is very difficult to distinguish water from the ground. Figure 10a and b show scene 1 listed on Table 2 with an overall accuracy of 33%. Figure 10c and d represents CAP images that have a relatively high percentage of water present. In this example, Fig. 10d represents a case with an overall estimated accuracy of 59% (scene 5) on Table 2. Lastly, Fig. 10e and f represents CAP images that have a relatively high percentage of water and dark color intensity. Figure 10f represents the last case, scene 11 from Table 2 with an overall estimated accuracy of 98%.

In Fig. 10b, there are areas at the top left and bottom right corners in which the algorithm fails to correctly classify water pixels. However, it correctly identifies the river located at the top right corner and the small accumulation of water in the center of the image. Although the algorithm overestimates pixels as water, it is still able to classify a fair portion of the control pixels. Several of the misclassified pixels are in areas that are

impossible to assess visually, and where water may indeed be present. In Fig. 10d, the classifier performs well at identifying water versus areas that have been eroded but are relatively the same in color, most likely because of the texture information extracted by the wavelet transformation. A similar example is shown in Fig. 10f where the algorithm correctly identifies nearly all of the manually identified control pixels.

There are cases when the algorithm overestimates water pixels such as in the lower left corner of Fig. 10b and the center of Fig. 10f. There are also cases when it underestimates and fails to identify water pixels such as in Fig. 10d. It is important to note that the goal of the algorithm is not to correctly identify every pixel in the image, but rather to generate a very quick automatic initial assessment with minimal expert supervision.

5 Conclusion

This study illustrates how freely available aerial imagery can aid in filling gaps commonly associated with traditional remote sensing techniques. The dependence of water features identified and extracted from RGB images using multiple classifiers shows a promising technique. The models were validated on a small test case to assess the quality of the learned models. This was accomplished by utilizing a selection of regions as training examples of flooded and non-flooded areas. Multiple classifiers were run in parallel to create a model describing water vs non-water regions. Furthermore, because multiple classifiers are employed, a measure of uncertainty can be derived from the number of classifiers that identify specific pixels as containing water. Wavelet transforms were used to compute scale coefficients and complement the RGB color space of the original images.

The ability to perform a quick and accurate analysis of flood conditions is paramount to aid disaster relief operations. Although the proposed methodology generates classifications which might include errors in a single image, the high spatial distribution of CAP images leads to a generally good first damage assessment, as shown in Fig. 6. Therefore, while water identification in individual images can be incorrect, the errors are generally small when aggregated over several thousand images.

This methodology improves upon current state-of-the-art approaches due to its use of freely available, but difficult-to-use, data sources in order to perform a more accurate and timely analysis of flood conditions. The combination of wavelet transformations and the use of multiple machine learning techniques provides a more accurate and robust image classification. This approach is also broadly applicable to problems outside of flood classification. It is expected that this method can be applied to the identification of fire burn scars or agricultural transitions in addition to this study case.

Acknowledgements Work performed under this project has been partially funded by the Office of Naval Research (ONR) Award #N00014-14-1-0208 (PSU #171570). We wish to thank Drs. Andris and Brooks for their comments that helped improve the present manuscript.

References

- Amolins K, Zhang Y, Dare P (2007) Wavelet based image fusion techniques an introduction, review and comparison. *ISPRS J Photogramm Remote Sens* 62(4):249–263
- Arivazhagan S, Ganesan L (2003) Texture classification using wavelet transform. *Pattern Recognit Lett* 24(9):1513–1521

- Aujol JF, Aubert G, Blanc-Féraud L (2003) Wavelet-based level set evolution for classification of textured images. *IEEE Trans Image Process* 12(12):1634–1641
- Baatz M, Schäpe A (2000) Multiresolution segmentation: an optimization approach for high quality multi-scale image segmentation. *Angewandte Geographische Informationsverarbeitung XII* 58:12–23
- Baaziz N, Abahmane O, Missaoui R (2010) Texture feature extraction in the spatial-frequency domain for content-based image retrieval. arXiv preprint [arXiv:10125208](https://arxiv.org/abs/10125208)
- Bai C, Zou W, Kpalma K, Ronsin J (2012) Efficient colour texture image retrieval by combination of colour and texture features in wavelet domain. *Electron Lett* 48(23):1463–1465
- Bharati MH, Liu JJ, MacGregor JF (2004) Image texture analysis: methods and comparisons. *Chemometr Intell Lab Syst* 72(1):57–71
- Bianconi F, Fernández A (2007) Evaluation of the effects of gabor filter parameters on texture classification. *Pattern Recognit* 40(12):3325–3335
- Breiman L, Friedman J, Stone CJ, Olshen RA (1984) Classification and regression trees. Wadsworth International Group, Belmont
- Campbell JB, Wynne RH (2011) Introduction to remote sensing, 5th ed. Guilford Press, New York
- Cervone G, Sava E, Huang Q, Schnebele E, Harrison J, Waters N (2016) Using twitter for tasking remote-sensing data collection and damage assessment: 2013 boulder flood case study. *Int J Remote Sens* 37(1):100–124
- Chang JD, Yu SS, Chen HH, Tsai CS (2010) Hsv-based color texture image classification using wavelet transform and motif patterns. *J Comput* 20(4):63–69
- Csurka G, Perronnin F (2011) An efficient approach to semantic segmentation. *Int J Comput Vis* 95(2):198–212
- Do MN, Vetterli M (2002) Wavelet-based texture retrieval using generalized gaussian density and kullback-leibler distance. *IEEE Trans Image Process* 11(2):146–158
- Dougherty G (2012) Pattern recognition and classification: an introduction. Springer Science & Business Media, Berlin
- Frazier PS, Page KJ et al (2000) Water body detection and delineation with landsat tm data. *Photogramm Eng Remote Sens* 66(12):1461–1468
- Gochis D, Schumacher R, Friedrich K, Doesken N, Kelsch M, Sun J, Ikeda K, Lindsey D, Wood A, Dolan B, et al (2014) The great Colorado flood of september 2013. *Bull Am Meteorol Soc*. doi:[10.1175/BAMS-D-13-00241.1](https://doi.org/10.1175/BAMS-D-13-00241.1)
- Goodchild MF, Glennon JA (2010) Crowdsourcing geographic information for disaster response: a research frontier. *Int J Digit Earth* 3(3):231–241
- Hadjimitsis DG (2007) The use of satellite remote sensing and gis for assisting flood risk assessment: a case study of the agriokalamini catchment area in paphos-cyprus. *Remote Sens Int Soc Optics Photonics* 67420Z–67420Z
- Han J, Kamber M, Pei J (2006) Data mining, southeast asia edition: concepts and techniques. Morgan Kaufmann, Burlington
- Haralick RM (1979) Statistical and structural approaches to texture. *Proc IEEE* 67(5):786–804
- Hein H, Feklistova L (2011) Computationally efficient delamination detection in composite beams using haar wavelets. *Mech Syst Signal Process* 25(6):2257–2270
- Huang K, Aviyente S (2008) Wavelet feature selection for image classification. *IEEE Trans Image Process* 17(9):1709–1720
- Im J, Jensen J, Tullis J (2008) Object-based change detection using correlation image analysis and image segmentation. *Int J Remote Sens* 29(2):399–423
- Islam MM, Sado K (2002) Development priority map for flood countermeasures by remote sensing data with geographic information system. *J Hydrol Eng* 7(5):346–355
- Joyce KE, Belliss SE, Samsonov SV, McNeill SJ, Glassey PJ (2009) A review of the status of satellite remote sensing and image processing techniques for mapping natural hazards and disasters. *Prog Phys Geogr*. doi:[10.1177/0309133309339563](https://doi.org/10.1177/0309133309339563)
- Lamovec P, Veljanovski T, Mikoš M, Oštir K (2013) Detecting flooded areas with machine learning techniques: case study of the selška sora river flash flood in september 2007. *J Appl Remote Sens* 7(1):073,564–073,564
- Lee J, Warner T (2006) Segment based image classification. *Int J Remote Sens* 27(16):3403–3412
- Li W, Liu K, Su H (2014) Wavelet-based nearest-regularized subspace for noise-robust hyperspectral image classification. *J Appl Remote Sens* 8(1):083,665–083,665
- Mallat SG (1989) A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans Pattern Anal Mach Intell* 11(7):674–693

- Mason DC, Horritt MS, Amico JTD, Scott TR, Bates PD (2007) Improving river flood extent delineation from synthetic aperture radar using airborne laser altimetry. *IEEE Trans Geosci Remote Sens* 45(12):3932–3943
- Materka A, Strzelecki M (1998) Texture analysis methods: a review. COST B11 report, Technical University of Lodz, Institute of Electronics, COST B11 report, Brussels
- Möller M, Lymburner L, Volk M (2007) The comparison index: a tool for assessing the accuracy of image segmentation. *Int J Appl Earth Obs Geoinf* 9(3):311–321
- Moody DI, Brumby SP, Rowland JC, Altmann GL (2014) Land cover classification in multispectral imagery using clustering of sparse approximations over learned feature dictionaries. *J Appl Remote Sens* 8(1):084,793–084,793
- Mueller M, Segl K, Kaufmann H (2004) Edge- and region-based segmentation technique for the extraction of large, man-made objects in high-resolution satellite imagery. *Patt Recognit* 37(8):1619–1628
- Muñoz X, Freixenet J, Cufi X, Martí J (2003) Strategies for image segmentation combining region and boundary information. *Patt Recognit Lett* 24(1):375–392
- Myint SW, Gober P, Brazel A, Grossman-Clarke S, Weng Q (2011) Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. *Remote Sens Environ* 115(5):1145–1161
- Nagasawa R et al (2014) Potential flood hazard assessment by integration of alos palsar and aster gdem: a case study for the hoachau commune, hoachau district, in central vietnam. *J Appl Remote Sens* 8(1):083,626–083,626
- Pantofaru C, Schmid C, Hebert M (2008) Object recognition by integrating multiple image segmentations. In: *Computer Vision—ECCV 2008*, Springer, pp 481–494
- Pittman E (2014) How Boulder, Colorado, prepared for ‘biblical’ rainfall. *Emerg Manag*
- Quinlan JR (1993) *C4.5: programs for machine learning*. Morgan Kaufmann, Burlington
- Rango A, Anderson AT (1974) Flood hazard studies in the mississippi river basin using remote sensing. *Wat Resour Bull* 10(5):1060–1081
- Rango A, Salomonson VV (1974) Regional flood mapping from space. *Water Resour Res* 10(3):473–484
- Recio JAR, Fernández LAR, Fernández-Sarriá A (2005) Use of gabor filters for texture classification of digital images. *Física de la Tierra* 17:47
- Sakamoto T, Van Nguyen N, Kotera A, Ohno H, Ishitsuka N, Yokozawa M (2007) Detecting temporal changes in the extent of annual flooding within the cambodia and the vietnamese mekong delta from modis time-series imagery. *Remote Sens Environ* 109(3):295–313
- Salamati N, Larlus D, Csurka G, Süsstrunk S (2014) Incorporating near-infrared information into semantic image segmentation. *arXiv preprint arXiv:14066147*
- Sanyal J, Lu X (2004) Application of remote sensing in flood management with special reference to monsoon asia: a review. *Nat Hazards* 33(2):283–301
- Schiewe J (2002) Segmentation of high-resolution remotely sensed data—concepts, applications and problems. *Int Archives Photogramm Remote Sens Spatial Inf Sci* 34(4):380–385
- Schnebele E, Cervone G (2013) Improving remote sensing flood assessment using volunteered geographical data. *Nat Hazards Earth Syst Sci* 13:669–677
- Schnebele E, Cervone G, Kumar S, Waters N (2014a) Real time estimation of the calgary floods using limited remote sensing data. *Water* 6(2):381–398
- Schnebele E, Cervone G, Waters N (2014b) Road assessment after flood events using non-authoritative data. *Nat Hazards Earth Syst Sci* 14(4):1007–1015
- Schumann G, Bates PD, Horritt MS, Matgen P, Pappenberger F (2009) Progress in integration of remote sensing-derived flood extent and stage data and hydraulic models. *Rev Geophys* 47(4):RG4001
- Shotton J, Winn J, Rother C, Criminisi A (2006) Textonboost: joint appearance, shape and context modeling for multi-class object recognition and segmentation. In: *Computer Vision—ECCV 2006*, Springer, pp 1–15
- Skarbnik N (2009) The importance of phase in image processing. Technion-Israel Institute of Technology, Faculty of Electrical Engineering
- Smith LC (1997) Satellite remote sensing of river inundation area, stage, and discharge: a review. *Hydrol Process* 11(10):1427–1439
- Steinberg D, Colla P (1995) *CART: tree-structured non-parametric data analysis*. Salford Systems, San Diego, CA
- Stow D, Hamada Y, Coulter L, Anguelova Z (2008) Monitoring shrubland habitat changes through object-based change identification with airborne multispectral imagery. *Remote Sens Environ* 112(3):1051–1061
- Vijayaraghavan C, Thirumalaivasan D, Venkatesan R (2012) Utilization of remote sensing and gis in managing disasters—a review. *Int J Sci Eng Res* 3(1):1–8

-
- Wang Y, Colby J, Mulcahy K (2002) An efficient method for mapping flood extent in a coastal floodplain using landsat tm and dem data. *Int J Remote Sens* 23(18):3681–3696
- Weidner U (2008) Contribution to the assessment of segmentation quality for remote sensing applications. *Int Archives Photogramm Remote Sens Spatial Inf Sci* 37(B7):479–484
- Winter S (2000) Location similarity of regions. *ISPRS J Photogramm Remote Sens* 55(3):189–200
- Zou W, Chi Z, Lo KC (2008) Improvement of image classification using wavelet coefficients with structured-based neural network. *Int J Neural Syst* 18(03):195–205