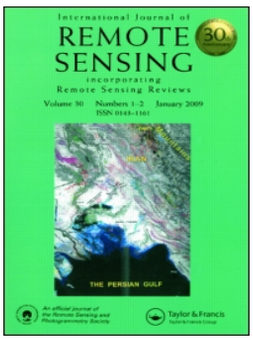

International Journal of Remote Sensing

ISSN: 0143-1161 (Print) 1366-5901 (Online) Journal homepage: http://www.tandfonline.com/loi/tres20

Impervious surface extraction in imbalanced datasets: integrating partial results and multi-temporal information in an iterative one-class classifier

Zewei Xu, Giorgos Mountrakis & Lindi J. Quackenbush

To cite this article: Zewei Xu, Giorgos Mountrakis & Lindi J. Quackenbush (2017) Impervious surface extraction in imbalanced datasets: integrating partial results and multi-temporal information in an iterative one-class classifier, International Journal of Remote Sensing, 38:1, 43-63, DOI: 10.1080/01431161.2016.1259677

To link to this article: http://dx.doi.org/10.1080/01431161.2016.1259677

INTERNATIONAL JOURNAL OF REMOTE SENSING, 2017
VOL. 38, NO. 1, 43–63
http://dx.doi.org/10.1080/01431161.2016.1259677

Taylor & Francis
Taylor & Francis Group

Impervious surface extraction in imbalanced datasets: integrating partial results and multi-temporal information in an iterative one-class classifier

Zewei Xu, Giorgos Mountrakis and Lindi J. Quackenbush

Department of Environmental Resources Engineering, SUNY College of Environmental Science and Forestry, Syracuse, NY, USA

ABSTRACT

Accurate urban land use/cover monitoring is an essential step towards a sustainable future. As a key part of the classification process, the characteristics of reference data can significantly affect classification accuracy and quality of produced maps. However, ideal reference data is not always readily available; users frequently have difficulty generating sufficient reference data for some classes given time, cost, data availability, expertise level, or other limitations. This study aims at dealing with this lack of sufficiently balanced reference data by presenting a modified hybrid one-class support vector data description (SVDD) model. The underlying hypothesis is that the lack of balanced reference data can be overcome through integration of partially extracted results and multi-temporal spectral information. The partially extracted results, defined as highly accurate classified pixels identified in previous algorithmic iterations, allow a gradual increase of the available training data. Furthermore, the method incorporates a voting system that integrates multi-temporal images using the SVDD algorithm. We applied this hybrid method to binary impervious classification of multi-temporal Landsat Thematic Mapper imagery from Central New York with imbalanced reference data. The proposed hybrid one-class SVDD model achieved a 5–6% improvement in overall accuracy and 0.05–0.09 in kappa than the typical one-class SVDD benchmark. While the method was tested on a single site (*albeit* with an unusually high reference dataset size of >870,000 pixels) we feel confident to suggest implementation of our methodology in other sites over the traditional method. This is because our approach automatically reverts to the traditional method when voting is inconsistent or there is a limited number of highly accurately classified pixels to assist future iterations. Future work could explore the quantity and temporal specificity (e.g. benefits of specific months) of the multi-temporal image selection and/or test other one-class classifiers.

ARTICLE HISTORY
Received 10 June 2016
Accepted 1 November 2016

1. Introduction

More than half of the world's population currently lives in cities (Madlener and Sunak 2011) and over the next several decades, urbanization will continue to accelerate to

CONTACT Giorgos Mountrakis gmountrakis@esf.edu Department of Environmental Resources Engineering, SUNY College of Environmental Science and Forestry, 1 Forestry Dr, Syracuse, NY 13210, USA

match population growth. The United Nations anticipates that by 2050 the proportion of people living in the urban environment will rise to 70% (United Nations 2008). The expansion of urban areas substantially changes surrounding land cover, particularly from pervious to impervious cover types. This can create significant problems including the loss of arable land, degradation of air quality, an increase of waste, and pollution of water in the urban environment. The spatial extent of impervious surface provides an important indicator of urbanization. Thus, accurate estimation of the distribution of impervious surface contributes to a wide range of ecosystem studies, such as hydrology, land use planning, and resource management (Schueler 1994; Luo and Mountrakis 2010; Weng 2012).

Compared to datasets derived from sources such as ground surveys, remotely sensed imagery is widely used in impervious surface extraction because of advantages such as cost efficiency and elimination of problems such as accessibility. Researchers have developed a wide range of classification models in order to accurately estimate the distribution of imperviousness. Regression models have been used to explore the relationship between impervious surfaces and other land cover types using remotely sensed data (Ridd 1995; Bauer, Loeffelholz, and Wilson 2005; Yang 2006; Jin and Mountrakis 2013). Spectral mixture analysis models have also been created to classify impervious surface (Roberts et al. 1998; Wu and Murray 2003; Braun and Herold 2004; Powell et al. 2007; Van de Voorde, De Roeck, and Canters 2009). A range of machine learning algorithms has been applied including decision tree models (Smith 2000; Herold, Koeln, and Cunningham 2003; Yang et al. 2003; Dougherty et al. 2004; Crane, Xian, and McMahon 2005), neural networks (Civco and Hurd 1997; Flanagan and Civco 2001; Iyer and Mohan 2002; Lee and Lathrop 2006), and Support Vector Machines (SVMs) (Mountrakis, Im, and Ogole 2011). Researchers have also applied contextual classification models that take into account labelling of neighbours to determine the most appropriate pixel class (Richards and Jia 2006; Luo and Mountrakis 2011; Mountrakis and Luo 2011).

Although existing methods were shown to be effective in a variety of different situations, supervised classification methods share a common premise: the generation of reference data is an essential step. However, generating reference data can be tedious since it typically involves substantial human input. Additionally, the generation of reference data often depends on photo interpretation or field assessment of points selected randomly. Ensuring adequate coverage of rare classes when an area has a disproportionately large area of a certain class, can be resource intensive as scattered high resolution imagery or field data would need to be collected and interpreted. With a random selection strategy in a highly imbalanced dataset, the selection of a higher proportion type is more probable than a lower proportion type. These challenges can lead to imbalanced reference data, which can directly impact classification quality (Foody et al. 2006; Muñoz-Marí et al. 2010).

One method employed to solve the reference data imbalance problem is using a data domain description algorithm, which is also called a one-class classifier. This approach works for binary classification problems where one class is sampled extensively, while the other class is severely undersampled (Tax and Duin 1999). The focus of domain description is to describe a set of objects, rather than distinguish between classes as in classification. This description should cover the class of objects represented by the

training set, and ideally should reject all other possible objects in the feature space. Although the data domain description is frequently used for detecting objects that differ significantly from the rest of the dataset, it is also a good method for one-class classification (Tax and Duin 1999).

There are several existing methods that use the data domain description concept. When the outlying patterns are assumed to follow a statistical distribution, this distribution can be estimated (Ritter and Gallegos 1997). Tarassenko, Hayton, and Brady (1995) used Parzen density estimation and a Gaussian mixture model to identify anomalies outside the normal class. Based on the support vector idea developed by Vapnik (1979), Schölkopf, Burges, and Smola (1999) also used a hyperplane to separate target objects from the origin with maximal margin. The SVM method has proved useful in document classification, texture segmentation, image classification, and ecological modelling (Tax and Duin 2002; Guo, Kelly, and Graham. 2005). However, this method also suffers drawbacks such as the sensitivity of the outcome to parameters that are difficult to tune (Manevitz and Yousef 2002; Muñoz-Marí et al. 2010).

Tax and Duin (1999) presented the SVM-inspired support vector data description (SVDD). SVDD obtains a spherically shaped boundary around a dataset and is analogous to the support vector classifier. The use of kernel functions makes this algorithm more flexible in depicting the boundary between classes and robust against outliers in the training set. In addition, the hypersphere used by SVDD can be further refined or tightened by using negative examples. Compared to other one-class classification methods, numerous tests have demonstrated a series of advantages for SVDD including smaller sample size requirements, and the ability to deal with sparse or complex datasets, especially when outlier information is used (Tax and Duin 1999; Muñoz-Marí et al. 2007; Sanchez-Hernandez, Boyd, and Foody 2007; Gurram, Kwon, and Han 2012; Khazai et al. 2012). The SVDD method has also been shown to be effective in terms of accuracy and robustness to high-dimensional problems when applied to different application domains, including facial expression analysis (Seo and Ko 2004; Zeng et al. 2006), speaker recognition (Dong, Wu, and H. Wang 2001), gene expression data clustering (Ji et al. 2008), image retrieval (Lai et al. 2004; Yu, DeMenthon, and Doermann 2008), remote-sensing image classification (Banerjee, Burlina, and Diehl 2006; Muñoz-Marí et al. 2007; Li, Song, and Xu 2011a; Li, Guo, and Elkan 2011b; Sakla et al. 2011), and hyperspectral image anomaly detection (Banerjee, Burlina, and Meth 2007).

Many studies in the remote-sensing field (e.g. Ji and Jensen 1999; Shackelford and Davis 2003; Lu and Weng 2006; Elmore and Guinn 2010; Lu, Moran, and Hetrick 2011a; Lu et al. 2011b) have demonstrated the feasibility of using multispectral satellite data to map impervious surface area in urban environments. Within these studies, a common source of image data are the Landsat missions for studies in North America. These satellites now provide multispectral information with near-global availability, with archives of some regions going back to 1972. Landsat data is commonly used for detection, quantification, and mapping of imperviousness patterns because of its repetitive data acquisition, wide availability, and accurate georeferencing procedures. Additionally, Landsat has relatively fine spatial resolution (30 m pixel size) compared to other free multispectral data sources such as the Moderate Resolution Imaging Spectroradiometer. In the USA, there are also many

state and federal orthophotography programmes with much higher spatial resolution than Landsat, but they lack Landsat's spectral richness and global coverage. Furthermore, multi-temporal Landsat data has proved to be more effective in land use/land cover classifications compared to single-date analyses. Guerschman et al. (2003) explored the use of multi-temporal Landsat TM data for the classification of various land cover types in the southwestern portion of the Argentine Pampas and achieved better results compared to using single-date imagery. Bauer, Yuan, and Sawaya (2003) combined multi-temporal Landsat images for land cover classification of the metropolitan area in Twin Cities of Minnesota and found that the combination of early summer with mid to late summer images provided higher classification accuracy than any of the single image dates. Similarly, Zheng, Campbell, and De Beurs (2012) used multi-temporal Landsat imagery to perform crop classification and achieved better classification results by incorporating multi-temporal features.

From a methodological perspective, iterative classification methods are advantageous over traditional ones because they can incorporate additional features generated from partially classified results; they are also able to self-adjust or update reference data during the iterative process. Partial results often contain additional features that can potentially enhance the performance of classifiers. There are several iterative classification strategies designed to achieve this goal. Luo and Mountrakis (2010) developed a hybrid classification model that used intermediate inputs derived from partial classification results on an impervious surface classification task with Landsat Enhanced Thematic Mapper Plus (ETM+) imagery. Their work suggested an average accuracy improvement of 3.6% by using intermediate inputs, which proved to be statistically significant. Jia et al. (2014) designed an automated approach for land cover updating by integrating iterative sample selection for training; they achieved a satisfactory classification accuracy of 83.1%. They enhanced the model by incorporating multi-temporal remote-sensing data using an iterative training sample selection method and achieved an accuracy increase of 3–4% (Jia et al. 2015). Hedjam et al. (2016) developed an iterative classification strategy by using ensembles of various machine learning algorithms and achieved better result compared to traditional methods. Guccione, Mascolo, and Appice (2015) implemented two classifiers that worked iteratively and exploited each other's decisions to improve the training phase resulting in an accuracy higher than basic classifierssuch as SVMs.

It is frequently difficult to generate balanced and sufficient reference data for many classification problems considering resource constraints (Johnson, Tateishi, and Hoan 2013; Fernández, García, and Herrera 2011). This paper provides a potential solution to imbalanced reference data usage through the development of an iterative process that incorporates multi-temporal imagery. Our scientific hypothesis explores potential substitution of training data from one class with the inclusion of multi-temporal imagery and temporary initial results obtained from an iterative, hierarchical classification. We employ a multi-temporal SVDD approach that simultaneously takes advantage of the algorithmic benefits of the one-class SVDD method and the rich information embedded in the multi-temporal Landsat images.

2. Study area and data sets

A 35 km × 27 km region located in Central New York State, USA containing the city of Syracuse was selected to evaluate the proposed classification approach. The proportion of impervious to pervious cover within the study site is about 1:3. This area, shown in Figure 1, contains land cover classes of bare soil, forest, grass, water, and impervious surface. To minimize processing and analysis time, six square subregions (in green, each with 190 pixels × 190 pixels) were selected to provide representatives of dense, medium and low impervious regions in the study area. Calibration samples (training and optimization) were randomly and equally selected within these six subregions.

2.1. *Landsat inputs*

This study used all Landsat 5 images of the study site acquired in 2000 with cloud coverage less than 20%. This included six Thematic Mapper (TM) scenes acquired on 2 April, 25 April, 4 May, 21 June, 27 October, and 3 November in 2000 with path 15 and row 30. While the scenes were screened for minimal cloud coverage, if a cloud appeared on any of the six scenes those pixels were excluded from the entire classification process. Because the six images were taken within a seven month period, the land cover was assumed to remain unchanged and thus the same reference data could be used for classification of all images. The multidate approach provided additional information to reduce potential class confusion (Henits, Jürgens, and Mucsi 2016).

Six bands (blue, green, red, near infrared (IR), and two short wave IR) from each TM image were utilized to provide the spectral information for the classifications. In

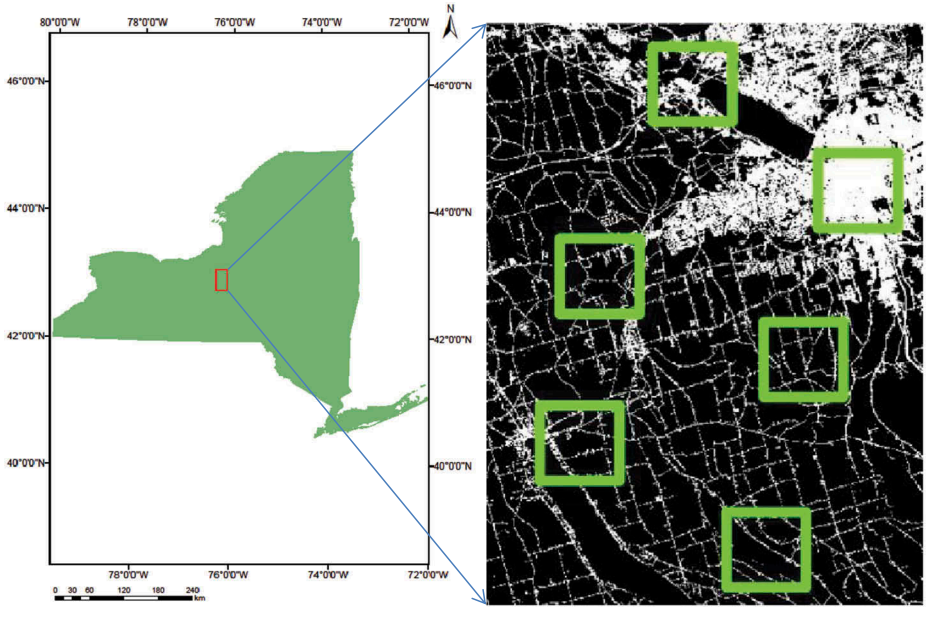

Figure 1. Study site in Central New York State. Enlarged area (35 km × 27 km) on the right shows impervious cover in white (reference image) overlaid with six representative subregions (green rectangles).

addition, three indices were calculated – the normalized difference vegetation index (NDVI) (Carlson and Ripley 1997), the normalized difference soil index (NDSI) (Wolf 2012), and the normalized difference water index (NDWI) (Gao 1996) – using the following equations:

$$\mathrm{NDVI} = \frac{(R_{\mathrm{NIR}} - R_{\mathrm{RED}})}{(R_{\mathrm{NIR}} + R_{\mathrm{RED}})}, \tag{1}$$

$$\mathrm{NDSI} = \frac{(R_{\mathrm{SWIR}} - R_{\mathrm{NIR}})}{(R_{\mathrm{SWIR}} + R_{\mathrm{NIR}})}, \tag{2}$$

$$\mathrm{NDWI} = \frac{(R_{\mathrm{RED}} - R_{\mathrm{SWIR}})}{(R_{\mathrm{RED}} + R_{\mathrm{SWIR}})}, \tag{3}$$

where R_{RED} is the reflectance of the red band, R_{NIR} is the near IR band reflectance, and R_{SWIR} is the shortwave infrared band 5 (1.55–1.75 μm) reflectance of the TM imagery.

In addition to spectral content, texture information has long been found effective in the classification process (Haralick, Shanmugam, and Dinstein 1973). While there are many texture measures that can be incorporated in image classification, for this study, the spectral bands and indices were supplemented by the standard deviation of each index within a local 3 × 3 pixel window. The standard deviation was selected because it was a simple measure of texture that was straightforward to implement and has been shown to be useful in impervious surface extraction (Luo and Mountrakis 2012). To summarize, all classifications used the six spectral bands, three indices (NDVI, NDSI, NDWI), and three texture inputs derived from the standard deviation of each index within a 3 × 3 pixel window for a total of 12 inputs per image.

2.2. *Reference data*

Reference data came from previous work performed by Luo and Mountrakis (2010). Luo and Mountrakis (2010) generated reference data using three band (near IR, red, and green) 1999 and 2001 digital orthophotography with 0.61 m pixel size, as well as a Landsat ETM+ scene acquired on April 2001. The ETM+ image was overlaid on the photography and on-screen interpretation and digitization techniques were utilized to generate a binary reference dataset of the entire 35 km × 27 km area. The strategy applied by Luo and Mountrakis (2010) was that Landsat pixels that included any impervious pixel on the orthophotography were assigned to the imperviousness class; all other pixels were labelled as pervious. Each of the six TM images used in this study was registered to this reference dataset. Since the reference data are in a binary form, classification was conducted using a binary strategy.

Table 1 summarizes the size of the training and optimization datasets applied for calibrating the different scenarios in the study. In the SVDD and hybrid SVDD methods, only pervious pixels were used for training. The rationale to use pervious data as training is that this is the dominant cover type in our study site, thus pervious training data are easier to acquire (e.g. centre of agricultural parcels, vegetation patches, or water). The two-class SVM classification was trained used 300 samples from the impervious class and 300 samples from the pervious class. All models require both pervious and impervious

Table 1. Calibration data (training and optimization) for two-class SVM, SVDD, and hybrid SVDD methods.

			Optimization pixel count		
Method	Class	Training pixel count	Scenario 1	Scenario 2	Scenario 3
Two-class SVM	Impervious	300	210	120	60
	Pervious	300	210	120	60
SVDD	Impervious	0	210	120	60
	Pervious	300	210	120	60
Hybrid SVDD	Impervious	0	210	120	60
	Pervious	300	210	120	60

Training data remains the same across all scenarios, while the size of the optimization datasets varies.

data for model parameter optimization. We selected varying sample sizes (210, 120, and 60 pixels) for training to explore how imbalanced datasets affect the optimization and especially to what extent different levels of imbalance impact the result for each class. In all cases, validation used the full reference dataset developed by Luo and Mountrakis (2010), with pixels used for calibration removed.

3. Methods

3.1. *Benchmark and proposed classification algorithms*

The objective of our work is to propose a hybrid method that increases the accuracy of the original one-class SVDD classification method that was developed to deal with imbalanced training data. The underlying classification method was based on SVMs, as a 15 year meta-analysis study found them to outperform other methods (Khatami, Mountrakis, and Stehman 2016). For direct accuracy assessment our proposed new hybrid SVDD method is compared to the traditional one-class SVDD (Tax and Duin 1999). For completeness, we also incorporate a two-class SVM (Vapnik 2000); however, our method should not be directly compared to that as the two-class SVM does not support imbalanced training data. Rather, the purpose of reporting the two-class SVM is to characterize the potential cost associated with imbalanced data.

3.1.1. *One-class SVDD*

The one-class classification algorithm was first presented by Schölkopf, Burges, and Smola (1999) to deal with the problem of imbalanced training data for detecting abnormalities when the collection of training data for the abnormal state is expensive or impossible to acquire. The method was refined by Tax and Duin (1999) whose support vector data description algorithm obtains a spherical boundary around the data in feature space, instead of the planar approach in SVM. The volume of the SVDD hypersphere is minimized to mitigate incorporation of outliers in the solution. During the minimization process, the algorithm creates a representational model of the training data. If newly encountered data are significantly different, they are labelled as being out of class. Tax and Duin (1999) recommend tightening the hypersphere by incorporating a kernel function. While there are several different choices of kernel functions, this study used the Gaussian RBF kernel. The RBF has only two free parameters, s the kernel width and the penalty parameter C, to be tuned and is shown to yield tighter boundaries than

other kernel choices (Tax and Duin 1999; Shawe-Taylor and Cristianini 2004; Tax and Duin 2004; Muñoz-Marí et al. 2007). During training, a set of data are used to delineate the hypersphere. The portion of the training dataset used depends on the user's selection of *C* and *s* (Chang and Lin 2011). For this study, the SVDD algorithm used to train the model was the one incorporated in the Libsvm MATLAB toolbox developed by Chang and Lin (2011). The one-class SVDD method provided a benchmark and was implemented using a multi-temporal image stack containing all bands and indices.

3.1.2. *Proposed multi-temporal hybrid SVDD*

The hybrid SVDD algorithm developed in this paper uses the SVDD model described in the prior section but extends its applicability to multi-temporal imagery through a voting process. A hierarchical classification is proposed (Mountrakis 2008; Mountrakis et al. 2009) to incorporate pixels that have higher classification confidence as training data for subsequent rounds of classification. In addition, intermediate statistics based on those initially classified pixels are calculated and added into the subsequent classifications (Luo and Mountrakis 2010, 2012).

The original training dataset contains six spectral bands, three spectral indices (NDVI, NDSI, and NDWI), and the standard deviation of these indices within a 3 × 3 pixels window. Three hundred pervious samples were used to separately train the one-class SVDD classifier for each image date. Parameter optimization took place by using a lower number of pervious and impervious training samples (discussed further in the following section).

The pervious/impervious binary results from the independent classification of each of the six dates were combined in a voting process (Figure 2). Pixels with complete agreement for either class were considered to be classified. Complete agreement was used to limit potential misclassifications, for example bare soil in the winter images spectrally mixing with the impervious class. These pixels (referred to as partial results)

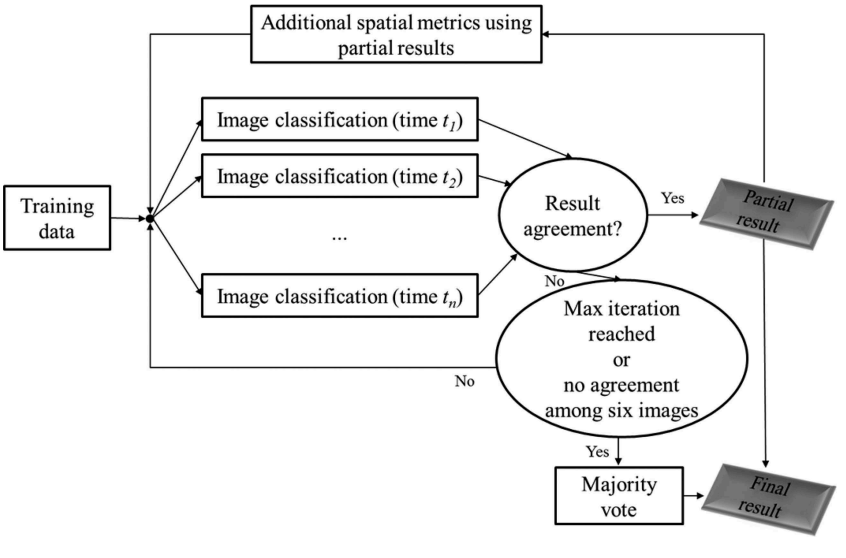

Figure 2. Flowchart of iterative classification.

were included in the final classification, but also served a different purpose: they became additional training data for the next classification iteration of the remaining unclassified pixels. The partial results were added to the training data as a new training set that includes both pervious and impervious training data. Optimization data are regenerated by excluding pixels that have already been classified in the training data and including a resampling process to keep the pervious/impervious distribution equal. By using the partial results, two additional input variables were created for each image, namely the distances to the nearest complete agreement pixel for each of the two classes. It is possible to perform a two-class SVM using data from the two classes derived from the first round of classification. However, we continued to use the SVDD approach because prior work has found that a one-class algorithm performs as well as the two-class SVM when sufficient training information from both classes has been added and is more computationally efficient when dealing with large training datasets (Li and Manikopoulos 2004; Senf, Chen, and Zhang 2006; Wu and Ye 2009).

The user's accuracy of the impervious surface from the classification of each image was multiplied by the voting score map to be an additional input variable and also applied to subsequent classifications. The voting and classification iterations continued until no additional pixels are classified. Any pixels that remained unclassified are then assigned to the class with the highest weighted sum vote based on the user's accuracy from each of the six images. We used user's accuracy rather than producer's accuracy because thematic map users and land use policy makers care more about the user's accuracy since it reveals the utility or accuracy of the final map product. In addition, our tests using optimization pixels indicated that creating weights based on user's accuracy was advantageous over producer's accuracy in the model in terms of accuracy. This suggests that the omission error (exclusion) by classifying the remaining pixels dominates over the commission error (inclusion), making the weights created from the user's accuracy more useful.

3.1.3. *Two-class SVM*

SVMs are a supervised non-parametric statistical learning technique first presented by Vapnik (1979). By learning from a training dataset, the SVM method generates a set of labelled data instances and then seeks to find hyperplanes that separate the dataset into a discrete number of classes consistent with the training examples. This supervised classification method requires training data from each target class in order to classify the entire dataset. The simplest SVM is the two-class SVM that uses a binary classification process to assign samples into one of two possible classes. Mountrakis, Im, and Ogole (2011) provide a review of SVM implementations in remote sensing.

Before training begins for an SVM classification a kernel function type is selected. In order to map the input vectors into a very high-dimensional feature space and better separate the training data, Schölkopf and Smola (2002) recommend applying a radial basis function (RBF) kernel. During training two parameters are optimized: the penalty parameter, C, and the kernel width, s (Schölkopf and Smola 2002). The penalty parameter provides a trade-off between the margin between classes and the number of target objects rejected (Lo and Wang 2012). For this study, we employed the Libsvm MATLAB toolbox (Chang and Lin 2011).

3.2. *Parameter optimization*

The two-class SVM and one-class SVDD algorithms with an RBF kernel require selection of two parameters: the penalty parameter and kernel width parameter. A grid search provides a good method to efficiently search for these parameters since it characterizes the searching range using two dimensions. The optimal values for the parameters occur at the intersection values on a two-dimensional grid (Hsu, Chang, and Lin 2003; Wang, Wu, and Zhang 2005). Reference data for both classes are required to perform the optimization.

3.2.1. *Selection of penalty parameter*

The penalty parameter C generally serves the same function in both SVM and SVDD algorithms, namely it allows the user to determine the trade-off between training error, in terms of the number of outliers excluded in training data, and the margin of the hyperplane. If C is too large, then the classification accuracy rate is very high in the training stage, but will be very low in the testing stage. If C is too small, then the classification accuracy rate may be unsatisfactory, making the model inapplicable.

We applied the parameter selection method developed by Theissler and Dear (2013) with an initial search range of parameter C set from 2^{-8} to 2^{8} using an exponential increment of 1. After using this first round of optimization to narrow the potential value for C, a finer increment of the exponent, 0.1, was used in a secondary search to find the best model. Since C is a weight parameter, it can be any fractional or integral number above zero (Lin et al. 2008; Theissler and Dear 2013).

3.2.2. *Selection of kernel width*

The kernel width parameter s has a stronger impact than C on classification outcomes, because its value influences the outcome of partitioning in the feature space. In SVM, s determines the complexity of the delineation of margins between classes. In SVDD, s determines the margin of the hypersphere and the number of support vectors that are used to generate the hypersphere. However, in both cases, excessive values for s lead to over-fitting, while a disproportionately small value can result in under-fitting (Lin et al. 2008; Theissler and Dear 2013). Because s is the width of the kernel, it can be any positive real number.

In this study, the initial search range of parameter s was set from 2^{-20} to 2^{20} using an increment of 1 on the exponent (Theissler and Dear 2013). As with the optimization for C, after the initial round of searching, a finer increment of the exponent, 0.1 was used to identify the best model. A grid search method is employed to identify the best combination of C and s in the iterations.

3.3. *Algorithmic calibration*

Calibration samples were randomly generated within the six subregions shown in Figure 1, which were selected as being representative of the entire study area. The number of samples was equally distributed among the six sites. The calibration dataset was separated in two different parts, one for training each algorithm and another for optimizing the parameters, in essence creating training and parameter optimization datasets. Accuracy was reported using a significantly expanded validation dataset, which is discussed in the next section.

3.4. *Accuracy assessment*

The reference dataset produced by Luo and Mountrakis (2010) provided the validation dataset, hence accuracy assessment was conducted using all cloud free pixels covering the entire 35 km × 27 km area, with calibration data excluded. Thus, the accuracy assessment for all three algorithms used 216,799 impervious and 660,869 pervious pixels. For the hybrid SVDD model a further accuracy assessment was incorporated to evaluate the accuracy of each intermediate classification based on the partially classified results. The accuracy assessment for the iteration steps was conducted based on the classified pixels with remaining unclassified pixels excluded from the accuracy calculation.

4. Results

The first section of the results presents the incremental and final classification results for the hybrid SVDD method. The second section compares the classification results of the SVDD benchmark method, the hybrid SVDD approach, and two-class SVM. In addition to evaluating the performance of the different classifiers, we also sought to understand the impact of the optimization sample size in the model performance using the aforementioned three scenarios.

4.1. *Iterative classification results of the hybrid SVDD model*

Tables 2–4 show the progress of the iterative classifications of the hybrid SVDD model for scenarios 1–3, respectively. These tables show the total number of classified pixels steadily increasing, while correspondingly, the overall accuracy (OA) and kappa decrease as more challenging pixels are classified with each iteration. When there is no additional increase in the number of classified pixels for an iteration, any remaining pixels are labelled. Prior to this final labelling, the proportion of classified pixels in scenarios 1, 2, and 3 showed significant variation with proportions of 75.32%, 64.24%, and 76.90%, respectively. Waske, Benediktsson, and Sveinsson (2009) found that overall accuracy and specific class accuracies were often compromised when imbalanced reference data are used. The different optimization scenarios presented in this study generally demonstrate good overall accuracy using the hybrid SVDD approach. However, the results do suggest that if the imbalance in the total

Table 2. Results from iterative hybrid SVDD classification for scenario 1 (210 testing points).

Scenario 1		Impervious			Pervious				
Iteration	Classified (%)	PA (%)	UA (%)	No. of pixels classified	PA (%)	UA (%)	No. of pixels classified	OA (%)	kappa
1	41.2	93.3	81.5	128,368	88.9	96.2	233,175	90.4	0.80
2	71.1	81.1	80.6	135,185	93.8	93.9	488,994	90.7	0.75
3	73.7	79.1	80.6	135,185	94.0	93.4	511,660	90.4	0.74
4	73.7	79.1	80.6	135,185	94.0	93.4	511,676	90.4	0.74
5	75.3	77.5	80.6	135,185	94.2	93.1	525,950	90.2	0.73
Final	100	62.6	79.6	170,454	94.7	88.5	707,214	86.8	0.62
No. of pixels	Total classified			170,454			707,214		
	Reference			216,799			660,869		

Table 3. Results from iterative hybrid SVDD classification in scenario 2 (120 testing points).

Scenario 2		Impervious			Pervious				
Iteration	Classified (%)	PA (%)	UA (%)	No. of pixels classified	PA (%)	UA (%)	No. of pixels classified	OA (%)	kappa
1	44.3	91.5	83.6	1157,59	92.3	96.2	272,772	92.1	0.82
2	63.2	89.1	74.8	151,435	88.8	95.6	403,248	88.9	0.73
3	64.0	87.8	73.9	151,874	88.4	95.1	409,472	88.3	0.72
4	64.2	87.8	73.7	153,290	88.2	95.1	409,972	88.1	0.72
5	64.2	87.8	73.3	153,680	88.0	95.1	410,159	87.9	0.71
Final	100	72.0	70.7	220,941	90.2	90.8	656,727	85.7	0.62
No. of pixels	Total classified		220,941				656,727		
	Reference		216,799				660,869		

Table 4. Results from iterative hybrid SVDD classification in scenario 3 (60 testing points).

Scenario 3		Impervious			Pervious				
Iteration	Classified (%)	PA (%)	UA (%)	No. of pixels classified	PA (%)	UA (%)	No. of pixels classified	OA (%)	kappa
1	43.8	91.0	83.9	115,759	92.6	96.1	272,772	92.2	0.82
2	69.3	81.1	81.4	151,435	94.1	94.0	403,248	91.0	0.75
3	76.9	69.6	78.0	151,874	93.6	90.4	409,472	87.7	0.66
4	76.9	69.6	77.9	153,290	93.6	90.4	409,972	87.7	0.67
Final	100	61.3	75.1	220,941	93.3	88.0	656,727	85.4	0.58
No. of pixels	Total classified		177,041				700,627		
	Reference		216,799				660,869		

number of reference samples does not reflect the proportions of the two classes, for example, in scenario 1 and 3, the larger proportion class can dominate the smaller class, an issue that was also observed by Fernández, García, and Herrera (2011). When the remaining pixels were labelled in the last step the accuracy assessment of final classified maps showed less variation across the three scenarios. The producer's accuracy (PA) and user's accuracy (UA) generally decreased with each iteration, except for the PA of the pervious class, which was not impacted by the iterative classification. This suggests that the omission error is well controlled in the classification, which can be attributed to the training data coming from the pervious class.

4.2. *Comparing the proposed hybrid SVDD model with the SVDD benchmark and two-class SVM*

Table 5 contrasts the proposed hybrid SVDD with the existing SVDD method. The accuracy assessment used the validation dataset with 216,799 impervious pixels and 660,869 pervious pixels. The hybrid SVDD model outperforms the traditional one-class SVDD model by 5–6% in overall accuracy and 0.05–0.09 in kappa. As expected the two-class SVM considerably outperformed the one-class SVDD method. Performance slightly improves as the dataset size increases. The hybrid SVDD model also offers comparable accuracy to the two-class SVM, despite the imbalanced dataset in the hybrid SVDD model suggesting that in cases where imbalanced datasets are the only training option, our method offers a viable alternative.

Table 5. Classification results comparison between benchmark, two-class SVM and hybrid SVDD model.

Scenario		Impervious		Pervious		OA	kappa	No. of pixels	
		PA (%)	UA (%)	PA (%)	UA (%)	(%)		Impervious	Pervious
1 (210 points)	SVDD benchmark	79.8	56.7	80.1	92.3	80.0	0.53	305,239	572,429
	Hybrid SVDD model	62.6	79.6	94.7	88.5	86.8	0.62	170,454	707,214
	Two-class SVM	87.2	63.6	83.6	95.2	84.5	0.63	297,321	580,347
2 (120 points)	SVDD benchmark	79.8	56.7	80.1	92.3	80.0	0.53	305,239	572,429
	Hybrid SVDD model	72.0	70.7	90.2	90.8	85.7	0.62	220,941	656,727
	Two-class SVM	87.0	63.0	83.3	95.1	84.2	0.62	299,157	578,512
3 (60 points)	SVDD benchmark	79.8	56.7	80.1	92.3	80.0	0.53	305,239	572,429
	Hybrid SVDD model	61.3	75.1	93.3	88.0	85.4	0.58	177,041	700,627
	Two-class SVM	87.4	61.5	82.1	95.2	83.4	0.61	307,934	569,734

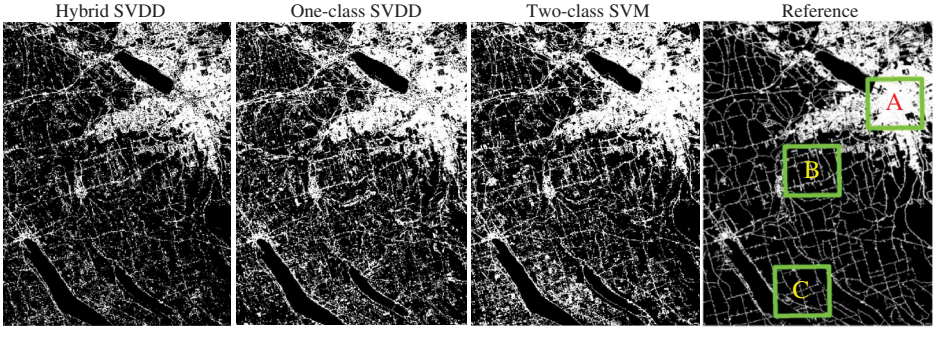

Figure 3. Iterative classification results and reference data highlighting three sites (A, B, C corresponding to downtown, suburban, and rural areas, respectively) with varying proportions of impervious pixels.

Figure 3 visually compares the classification result from scenario 2 (120 sample points) using the hybrid SVDD model, the SVDD benchmark method and the two-class SVM with the reference dataset. The hybrid SVDD model provides clearer delineation of the road network, boundaries of the city and suburban area. The baseline SVDD and SVM methods overestimate the impervious coverage by mistakenly classifying many patches in suburban areas into the impervious class.

Figure 4 shows enlargements of difference maps for sites A, B, and C as shown in Figure 3. This figure highlights three locations: one in the downtown area A, one suburban area B and one rural area C. The hybrid SVDD model underestimates the number of impervious pixels in the downtown area of Syracuse and misses some road extensions in suburban areas. However, when compared to the classified images of the other two methods, the hybrid SVDD model seems to have better overall performance in impervious class estimation, which was also reflected in the user's accuracy estimation of the impervious class in Table 5.

5. Discussion and conclusions

Generating reference data can be tedious since it typically involves substantial human input. Additionally, the generation of reference data from one class is often much easier

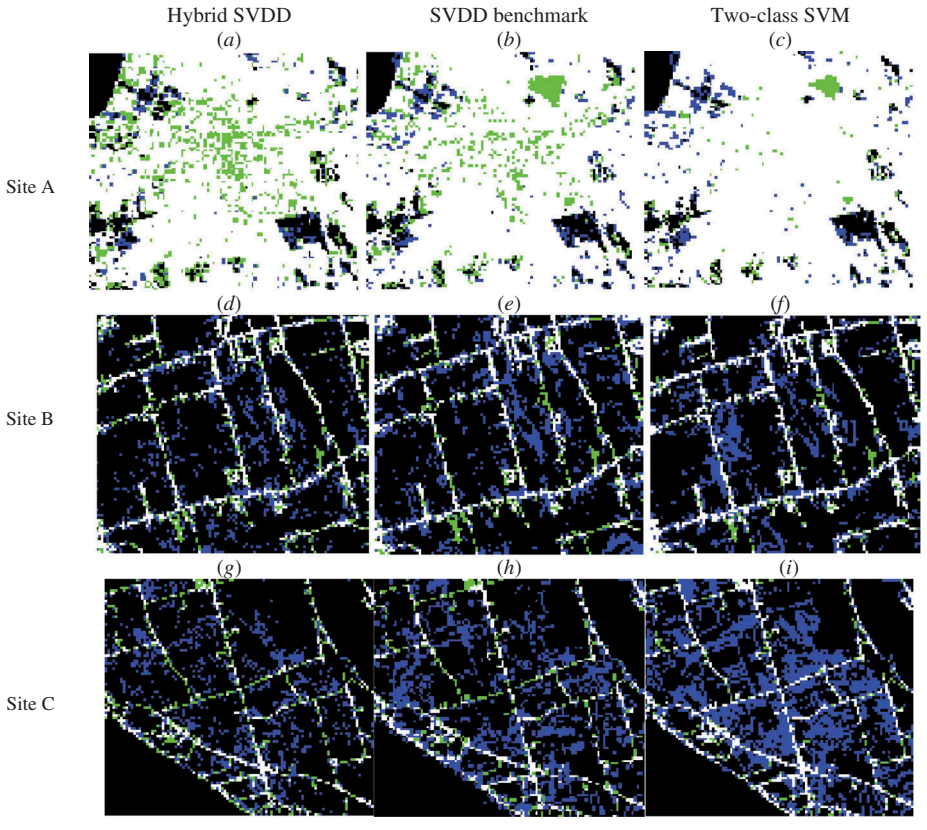

Figure 4. Difference maps comparing the three classification methods to the reference image: (*a*), (*d*), (*g*) differences between the hybrid SVDD and reference image; (*b*), (*e*), (*h*) differences between the SVDD benchmark and reference image; (*c*), (*f*), (*i*) differences between the two-class SVM and reference image. Correctly classified impervious pixels are white; correctly classified pervious class are black; green pixels are misclassified impervious pixels; blue pixels are misclassified pervious pixels.

than other classes due to different proportions or inaccessibility of some regions (Johnson, Tateishi, and Hoan 2013; Fernández, García, and Herrera 2011). The resultant imbalanced reference data can greatly hinder the performance of many standard classifiers including decision trees, neural networks, and SVMs (Japkowicz and Stephen 2002; Kang and Cho 2006; Waske, Benediktsson, and Sveinsson 2009).

Fernández, García, and Herrera (2011) categorized solutions for imbalanced dataset classification into two types: data sampling and algorithmic modification. Data sampling approaches modify training instances to produce a balanced distribution that allows classifiers to perform similarly to a standard classification, while algorithmic modification methods seek a computational solution, for example adaptation of learning methods, to deal with imbalance issues. Most prior methods focused on solving the issue of data imbalance using techniques that fall into one of these two categories. Waske, Benediktsson, and Sveinsson (2009) resampled imbalanced data and developed a multiple SVM classifier system, which enhanced the accuracy of certain underrepresented land cover types, but at a cost of reducing the overall accuracy and accuracy of the most

of other types. Lee and Lee (2012) developed a hybrid method from an algorithmic standpoint using analysis of variance, fuzzy C-means, and bacterial foraging optimization to do the classification and found it outperformed other classifiers in terms of overall accuracy. Serpico and Bruzzone (1997) proposed a three-phase neural network technique to avoid misrepresentation of minority classes with successful results compared to typical neural network approaches.

Our approach incorporates both data sampling and algorithmic modification for imbalanced dataset classification. We used a stratified sampling strategy by identifying areas of differing density of the less common cover type in generating calibration data, which is believed to be more representative of the population than random sampling from the population. We also imbed iterative classification in our method, an approach first presented by Luo and Mountrakis (2010) for impervious surface extraction. Luo and Mountrakis (2010) designed a hybrid classifier that utilized manipulated results from prior classification levels to enhance accuracy in subsequent classification steps. They also presented a multi-process classification model that consisted of *a priori* and *a posteriori* classifiers for impervious surface extraction (Mountrakis and Luo 2011). This study extends these prior studies by designing a new method that iteratively adjusts the hypervolume used to separate the feature space by updating training data and incorporating new knowledge from a partially classified result using multi-temporal dataset. This differs from traditional classification where the training dataset remains constant and the prior work presented by Luo and Mountrakis (2010) who classified single date imagery with neural networks. By incorporating an iterative, multi-spectral approach, our hybrid SVDD method produced 5.5–7% higher overall accuracy than the SVDD benchmark and offers comparable accuracy to the two-class SVM, despite the imbalanced dataset in the hybrid SVDD model. These results suggest that our method offers a viable alternative in cases where imbalanced datasets are the only training option.

We expect our results to be generalizable in other sites for several reasons. First, the manually extracted reference dataset is unusually large for a remote-sensing study (>870,000 points), 1–2 orders of magnitude higher than comparable studies. Additionally, within our study area, we have landscape heterogeneity with a transition from highly urban to rural, varying distributions of vegetation, agriculture and water, and strong topographic variability. Finally, and most importantly, our method reverts back to the traditional SVDD method if it detects inconsistency in the temporal labels or there are a limited number of pixels classified with high accuracy. Waske, Benediktsson, and Sveinsson (2009) found the overall accuracy and specific class accuracies were often compromised when used an imbalanced reference data compared to balanced datasets; our proposed method tackles this issue. Although the different optimization scenarios presented generally demonstrate good overall accuracy using the hybrid SVDD approach, they do suggest that if the imbalance in the total number of reference samples is too high, for example, in scenario 3, the difference in the proportions between the two classes increases. This translates into domination of the larger proportion class with the smaller proportional class losing its representation, which was also highlighted by Fernández, García, and Herrera (2011). Finally, even though not explicitly tested, our study found that the SVDD algorithm was more computationally efficient than the SVM approach, which reflects prior work (Li and Manikopoulos 2004; Senf, Chen, and Zhang 2006; Wu and Ye 2009).

In this paper, the hybrid SVDD model is tested on impervious/pervious classification; future studies could explore other land use/cover class types and the quantity and temporal specificity (e.g. benefits of specific months) of the multi-temporal image selection. Furthermore, other algorithms that share similar conditions with the SVDD could be tested, for example the one-class SVM (OCSVM) proposed by Schölkopf, Burges, and Smola (1999). Also, bootstrapping or cross-validation can also be added into the accuracy assessment process to further test the performance of the developed algorithm. While there are many avenues for future enhancement, this study provides a novel building platform to address the frequent problem of reference dataset imbalance.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Giorgos Mountrakis http://orcid.org/0000-0001-5958-8134

References

Banerjee, A., P. Burlina, and C. Diehl. 2006. "A Support Vector Method for Anomaly Detection in Hyperspectral Imagery." *IEEE Transactions on Geoscience and Remote Sensing* 44: 2282–2291. doi:10.1109/TGRS.2006.873019.

Banerjee, A., P. Burlina, and R. Meth. 2007. "Fast Hyperspectral Anomaly Detection via SVDD." *Proceedings of IEEE International Conference on Image Processing* 4: 101–104.

Bauer, M., B. Loeffelholz, and B. Wilson. 2005. "Estimation, Mapping and Change Analysis of Impervious Surface Area by Landsat Remote Sensing." In *Proceedings of Pecora 16 Conference, American Society of Photogrammetry and Remote Sensing Annual Conference,* Sioux Falls, SD, October 23–27, CD-ROM.

Bauer, M. E., F. Yuan, and K. E. Sawaya. 2003. "Multi-Temporal Landsat Image Classification and Change Analysis of Land Cover in the Twin Cities (Minnesota) Metropolitan Area." In *Proceedings, MultiTemp-2003, Second International Workshop on the Analysis of Multi-Temporal Remote Sensing Images*, July 16–18, 368–375. Singapore: World Scientific Publishing Co. Z.

Braun, M., and M. Herold. 2004. "Mapping Imperviousness Using NDVI and Linear Spectral Unmixing of ASTER Data in the Cologne-Bonn Region (Germany)." In Proceedings of the SPIE 10th International Symposium on Remote Sensing, Barcelona, Spain, September 8–12. Accessed March 20, 2010. http://www.geogr.uni-jena.de/~c5hema/pub/spie_paper_mhfinal_4.pdf.

Carlson, T. N., and D. A. Ripley. 1997. "On the Relation between NDVI, Fractional Vegetation Cover, and Leaf Area Index." *Remote Sensing of Environment* 62: 241–252. doi:10.1016/S0034-4257(97)00104-1.

Chang, C.-C., and C.-J. Lin. 2011. "LIBSVM: A Library for Support Vector Machines." *ACM Transactions on Intelligent Systems and Technology* 2: 1–27. doi:10.1145/1961189.

Civco, D. L., and J. D. Hurd. 1997. "Impervious Surface Mapping for the State of Connecticut." In *Proceedings of the ASPRS Annual Conference*, Seattle, WA, 124–135.

Crane, M., G. Xian, and C. McMahon. 2005. "Estimation of Sub-Pixel Impervious Surface Using Landsat and ASTER Imagery for Assessing Urban Growth." In Proceedings of the ISPRS Joint Conference: 5th International Symposium on Remote Sensing and Data Fusion over Urban Areas (URBAN 2005) 5th International Symposium on Remote Sensing of Urban Areas (URS 2005), Tempe, AZ, March 14–16.

Dong, X., Z. Wu, and W. F. H. Wang. 2001. "Support Vector Domain Description for Speaker Recognition." In *Neural Networks for Signal Processing XI, 2001. Proceedings of the 2001 IEEE Signal Processing Society Workshop*, 481–488. New York: IEEE.

Dougherty, M., R. L. Dymond, S. J. Goetz, C. A. Jantz, and N. Goulet. 2004. "Evaluation of Impervious Surface Estimates in a Rapidly Urbanizing Watershed." *Photogrammetric Engineering & Remote Sensing* 70: 1275–1284. doi:10.14358/PERS.70.11.1275.

Elmore, A. J., and S. M. Guinn. 2010. "Synergistic Use of Landsat Multispectral Scanner with GIRAS Land-Cover Data to Retrieve Impervious Surface Area for the Potomac River Basin in 1975." *Remote Sensing of Environment* 114: 2384–2391. doi:10.1016/j.rse.2010.05.004.

Fernández, A., S. García, and F. Herrera. 2011. "Addressing the Classification with Imbalanced Data: Open Problems and New Challenges on Class Distribution." In *Hybrid Artificial Intelligent Systems*, 1–10. Berlin, Heidelberg: Springer.

Flanagan, M., and D. L. Civco. 2001. "Subpixel Impervious Surface Mapping." In *Proceedings of the 2001 ASPRS Annual Convention*, St. Louis, MO, 23.

Foody, G. M., A. Mathur, C. Sanchez-Hernandez, and D. S. Boyd. 2006. "Training Set Size Requirements for the Classification of a Specific Class." *Remote Sensing of Environment* 104: 1–14. doi:10.1016/j.rse.2006.03.004.

Gao, B.-C. 1996. "NDWI—A Normalized Difference Water Index for Remote Sensing of Vegetation Liquid Water from Space." *Remote Sensing of Environment* 58: 257–266. doi:10.1016/S0034-4257(96)00067-3.

Guccione, P., L. Mascolo, and A. Appice. 2015. "Iterative Hyperspectral Image Classification Using Spectral–Spatial Relational Features." *IEEE Transactions on Geoscience and Remote Sensing* 53 (7): 3615–3627. doi:10.1109/TGRS.2014.2380475.

Guerschman, J. P., J. M. Paruelo, C. D. Bella, M. C. Giallorenzi, and F. Pacin. 2003. "Land Cover Classification in the Argentine Pampas Using Multi-Temporal Landsat TM Data." *International Journal of Remote Sensing* 24 (17): 3381–3402. doi:10.1080/0143116021000021288.

Guo, Q., M. Kelly, and C. H. Graham. 2005. "Support Vector Machines for Predicting Distribution of Sudden Oak Death in California." *Ecological Modelling* 182: 75–90. doi:10.1016/j.ecolmodel.2004.07.012.

Gurram, P., H. Kwon, and T. Han. 2012. "Sparse Kernel-Based Hyperspectral Anomaly Detection." *IEEE Geoscience and Remote Sensing Letters* 9: 943–947. doi:10.1109/LGRS.2012.2187040.

Haralick, R. M., K. Shanmugam, and I. Dinstein. 1973. "Textural Features for Image Classification." *IEEE Transactions on Systems, Man and Cybernetics, SMC* 3: 610–621. doi:10.1109/TSMC.1973.4309314.

Hedjam, R., M. Kalacska, M. Mignotte, and H. Z. Nafchi. 2016. "Iterative Classifiers Fusion Model for Change Detection in Remote Sensing Imagery." *change (ie, binary change detection, change or no-change)* 2: 11.

Henits, L., C. Jürgens, and L. Mucsi. 2016. "Seasonal Multitemporal Land-Cover Classification and Change Detection Analysis of Bochum, Germany, Using Multitemporal Landsat TM Data." *International Journal of Remote Sensing* 37: 1–16. doi:10.1080/01431161.2015.1125558.

Herold, N., G. Koeln, and D. Cunningham. 2003. "Mapping Impervious Surfaces and Forest Canopy Using Classification and Regression Tree (CART) Analysis." In *ASPRS Annual Conference Proceedings*, Anchorage, AK.

Hsu, C. W., C. C. Chang, and C. J. Lin. 2003. *A Practical Guide to Support Vector Classification*. Technical Report, University of National Taiwan, Department of Computer Science and Information Engineering, Taipei, Taiwan, July 1–12.

Iyer, S. V., and B. K. Mohan. 2002. "Urban Land Use Monitoring Using Neural Network Classification." *Proceedings of the IEEE Symposium on Geoscience and Remote Sensing* 5: 2959–2961.

Japkowicz, N., and S. Stephen. 2002. "The Class Imbalance Problem: A Systematic Study." *Intelligent Data Analysis* 6 (5): 429–449.

Ji, M., and J. R. Jensen. 1999. "Effectiveness of Subpixel Analysis in Detecting and Quantifying Urban Imperviousness from Landsat Thematic Mapper Imagery." *Geocarto International* 14 (4): 33–41. doi:10.1080/10106049908542126.

Ji, R., L. Ding, W. Min, and J. Liu. 2008. "The Application of SVDD in Gene Expression Data Clustering." In *Bioinformatics and Biomedical Engineering, ICBBE. The 2nd International Conference*, 371–374. Shanghai: IEEE.

Jia, K., Q. Li, X. Wei, L. Zhang, X. Du, Y. Yao, and X. Wang. 2015. "Multi-Temporal Remote Sensing Data Applied in Automatic Land Cover Update Using Iterative Training Sample Selection and Markov Random Field Model." *Geocarto International* 30 (8): 882–893. doi:10.1080/10106049.2014.997310.

Jia, K., S. Liang, X. Wei, L. Zhang, Y. Yao, and S. Gao. 2014. "Automatic Land-Cover Update Approach Integrating Iterative Training Sample Selection and a Markov Random Field Model." *Remote Sensing Letters* 5 (2): 148–156. doi:10.1080/2150704X.2014.889862.

Jin, H., and G. Mountrakis. 2013. "Integration of Urban GrowthModelling Products with Image-Based Urban Change Analysis." *International Journal of Remote Sensing* 34: 5468–5486. doi:10.1080/01431161.2013.791760.

Johnson, B. A., R. Tateishi, and N. T. Hoan. 2013. "A Hybrid Pansharpening Approach and Multiscale Object-Based Image Analysis for Mapping Diseased Pine and Oak Trees." *International Journal of Remote Sensing* 34 (20): 6969–6982. doi:10.1080/01431161.2013.810825.

Kang, P., and S. Cho. October 2006. "EUS Svms: Ensemble of Under-Sampled Svms for Data Imbalance Problems." In *Neural Information Processing*, Hong Kong, 837–846. Berlin, Heidelberg: Springer.

Khatami, R., G. Mountrakis, and S. Stehman. 2016. "A Meta-Analysis of Remote Sensing Research on Supervised Pixel-Based Land-Cover Image Classification Processes:General Guidelines for Practitioners and Future Research." *Remote Sensing of Environment* 177: 89–100. doi:10.1016/j.rse.2016.02.028.

Khazai, S., A. Safari, B. Mojaradi, and S. Homayouni. 2012. "Improving the SVDD Approach to Hyperspectral Image Classification." *IEEE Geoscience and Remote Sensing Letters* 9: 594–598. doi:10.1109/LGRS.2011.2176101.

Lai, C., D. M. J. Tax, R. P. W. Duin, E. Pekalska, and P. Paclik. 2004. "A Study on Combining Image Representations for Image Classification and Retrieval." *International Journal of Pattern Recognition and Artificial Intelligence* 18: 867–890. doi:10.1142/S0218001404003459.

Lee, C.-Y., and Z.-J. Lee. 2012. "A Novel Algorithm Applied to Classify Unbalanced Data." *Applied Soft Computing* 12 (8): 2481–2485. doi:10.1016/j.asoc.2012.03.051.

Lee, S., and R. G. Lathrop. 2006. "Subpixel Analysis of Landsat ETM+ Using Self-Organizing Map (SOM) Neural Networks for Urban Land Cover Characterization." *IEEE Transactions on Geoscience and Remote Sensing* 44: 1642–1654. doi:10.1109/TGRS.2006.869984.

Li, L., and C. N. Manikopoulos. 2004. "Windows NT One-class Masquerade Detection." In *Proceedings from the Fifth Annual IEEE Information Assurance Workshop*, 82–87. West Point, NY: SMC.

Li, P., B. Song, and H. Xu. 2011a. "Urban Building Damage Detection from Very High Resolution Imagery by One-Class SVM and Shadow Information." In IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Vancouver, Canada, July 24–29, 1409–1412.

Li, W., Q. Guo, and C. Elkan. 2011b. "A Positive and Unlabeled Learning Algorithm for One-Class Classification of Remote-Sensing Data." *IEEE Transactions on Geoscience and Remote Sensing* 49: 717–725. doi:10.1109/TGRS.2010.2058578.

Lin, S.-W., Z.-J. Lee, S.-C. Chen, and T.-Y. Tseng. 2008. "Parameter Determination of Support Vector Machine and Feature Selection Using Simulated Annealing Approach." *Applied Soft Computing* 8: 1505–1512. doi:10.1016/j.asoc.2007.10.012.

Lo, C.-S., and C.-M. Wang. 2012. "Support Vector Machine for Breast MR Image Classification." *Computers & Mathematics with Applications* 64: 1153–1162. doi:10.1016/j.camwa.2012.03.033.

Lu, D., G. Li, E. Moran, M. Batistella, and C. C. Freitas. 2011b. "Mapping Impervious Surfaces with the Integrated Use of Landsat Thematic Mapper and Radar Data: A Case Study in an Urban–Rural Landscape in the Brazilian Amazon." *ISPRS Journal of Photogrammetry and Remote Sensing* 66: 798–808. doi:10.1016/j.isprsjprs.2011.08.004.

Lu, D., E. Moran, and S. Hetrick. 2011a. "Detection of Impervious Surface Change with Multitemporal Landsat Images in an Urban–Rural Frontier." *ISPRS Journal of Photogrammetry and Remote Sensing* 66: 298–306. doi:10.1016/j.isprsjprs.2010.10.010.

Lu, D., and Q. Weng. 2006. "Use of Impervious Surface in Urban Land-Use Classification." *Remote Sensing of Environment* 102: 146–160. doi:10.1016/j.rse.2006.02.010.

Luo, L., and G. Mountrakis. 2010. "Integrating Intermediate Inputs from Partially Classified Images within a Hybrid Classification Framework: An Impervious Surface Estimation Example." *Remote Sensing of Environment* 114: 1220–1229. doi:10.1016/j.rse.2010.01.008.

Luo, L., and G. Mountrakis. 2011. "Converting Local Spectral and Spatial Information from a Priori Classifiers into Contextual Knowledge for Impervious Surface Classification." *ISPRS Journal of Photogrammetry and Remote Sensing* 66: 579–587. doi:10.1016/j.isprsjprs.2011.03.002.

Luo, L., and G. Mountrakis. 2012. "A Multi-Process Model of Adaptable Complexity for Impervious Surface Detection." *International Journal of Remote Sensing* 33: 365–381. doi:10.1080/01431161.2010.532177.

Madlener, R., and Y. Sunak. 2011. "Impacts of Urbanization on Urban Structures and Energy Demand: What Can We Learnfor Urban Energy Planning and Urbanization Management?" *Sustainable Cities and Society* 1: 45–53. doi:10.1016/j.scs.2010.08.006.

Manevitz, L. M., and M. Yousef. 2002. "One-Class Svms for Document Classification." *The Journal of Machine Learning Research* 2: 139–154.

Mountrakis, G. 2008. "Next Generation Classifiers: Focusing on Integration Frameworks." *Photogrammetric Engineering and Remote Sensing* 74: 1178–1181.

Mountrakis, G., J. Im, and C. Ogole. 2011. "Support Vector Machines in Remote Sensing: A Review." *ISPRS Journal of Photogrammetry and Remote Sensing* 66: 247–259. doi:10.1016/j.isprsjprs.2010.11.001.

Mountrakis, G., and L. Luo. 2011. "Enhancing and Replacing Spectral Information with Intermediate Structural Inputs: A Case Study on Impervious Surface Detection." *Remote Sensing of Environment* 115: 1162–1170. doi:10.1016/j.rse.2010.12.018.

Mountrakis, G., R. Watts, L. Luo, and J. Wang. 2009. "Developing Collaborative Classifiers Using an Expert-Based Model." *Photogrammetric Engineering & Remote Sensing* 75: 831–843. doi:10.14358/PERS.75.7.831.

Muñoz-Marí, J., F. Bovolo, L. Gómez-Chova, L. Bruzzone, and G. Camps-Valls. 2010. "Semisupervised One-Class Support Vector Machines for Classification of Remote Sensing Data." *IEEE Transactions on Geoscience and Remote Sensing* 48: 3188–3197. doi:10.1109/TGRS.2010.2045764.

Muñoz-Marí, J., G. Camps-Valls, L. Gómez-Chova, and J. Calpe-Maravilla. 2007. "Combination of One-Class Remote Sensing Image Classifiers." In IEEE International Geoscience and Remote Sensing Symposium, Barcelona, 23–28 July, 1509–1512.

Powell, R. L., D. A. Roberts, P. E. Dennison, and L. L. Hess. 2007. "Sub-Pixel Mapping of Urban Land Cover Using Multiple Endmember Spectral Mixture Analysis: Manaus, Brazil." *Remote Sensing of Environment* 106: 253–267. doi:10.1016/j.rse.2006.09.005.

Richards, J. A., and X. Jia. 2006. *Remote Sensing Digital Image Analysis*. 4th ed. Berlin, Heidelberg: Springer.

Ridd, M. K. 1995. "Exploring a V-I-S (Vegetation-Impervious Surface-Soil) Model for Urban Ecosystem Analysis through Remote Sensing: Comparative Anatomy for Cities." *International Journal of Remote Sensing* 16: 2165–2185. doi:10.1080/01431169508954549.

Ritter, G., and M. T. Gallegos. 1997. "Outliers in Statistical Pattern Recognition and an Application to Automatic Chromosome Classification." *Pattern Recognition Letters* 18: 525–539. doi:10.1016/S0167-8655(97)00049-4.

Roberts, D. A., M. Gardner, R. Church, S. Ustin, G. Scheer, and R. O. Green. 1998. "Mapping Chaparral in the Santa Monica Mountains Using Multiple Endmember Spectral Mixture Models." *Remote Sensing of Environment* 65: 267–279. doi:10.1016/S0034-4257(98)00037-6.

Sakla, W., A. Chan, J. Ji, and A. Sakla. 2011. "An SVDD-Based Algorithm for Target Detection in Hyperspectral Imagery." *IEEE Geoscience and Remote Sensing Letters* 8: 384–388. doi:10.1109/LGRS.2010.2078795.

Sanchez-Hernandez, C., D. S. Boyd, and G. M. Foody. 2007. "One-Class Classification for Mapping a Specific Land-Cover Class: SVDD Classification of Fenland." *IEEE Transactions on Geoscience and Remote Sensing* 45: 1061–1073. doi:10.1109/TGRS.2006.890414.

Schölkopf, B., C. J. C. Burges, and A. J. Smola. 1999. *Advances in Kernel Methods—Support Vector Learning*. Cambridge, MA: MIT Press.

Schölkopf, B., and A. Smola. 2002. *Learning With Kernels-Support Vector Machines, Regularization, Optimization and Beyond*. Cambridge, MA: MIT Press.

Schueler, T. R. 1994. "The Importance of Imperviousness." *Watershed Protection Techniques* 1: 100–110.

Senf, A., X. W. Chen, and A. Zhang. 2006. "Comparison of One-Class SVM and Two-Class SVM for Fold Recognition." In *Neural Information Processing*, Hong Kong, 140–149. Berlin, Heidelberg: Springer.

Seo, J., and H. Ko. 2004. "Face Detection Using Support Vector Domain Description in Color Images." *Proceedings of IEEE Acoustics, Speech, and Signal Processing (ICASSP'04) International Conference* 5: 5–729.

Serpico, S. B., and L. Bruzzone. August 1997. "Training of Neural Networks for Classification of Imbalanced Remote-Sensing Data." In *Geoscience and Remote Sensing, 1997. IGARSS'97. Remote Sensing-A Scientific Vision for Sustainable Development. 1997 IEEE International*, edited by Tammy I. Stein, 1202–1204. Vol. 3. Singapore: IEEE.

Shackelford, A. K., and C. H. Davis. 2003. "A Combined Fuzzy Pixel-Based and Object-Based Approach for Classification of High-Resolution Multispectral Data over Urban Areas." *IEEE Transactions on Geoscience and Remote Sensing* 41: 2354–2364. doi:10.1109/TGRS.2003.815972.

Shawe-Taylor, J., and N. Cristianini. 2004. *Kernel Methods for Pattern Analysis*. Cambridge, UK: Cambridge University Press.

Smith, A. J. 2000. "Subpixel Estimates of Impervious Surface Cover Using Landsat TM Imagery." MA scholarly paper. University of Maryland, College Park.

Tarassenko, L., P. Hayton, and M. Brady. 1995. "Novelty Detection for the Identification of Masses in Mammograms." *Proceedings of the Fourth International Conference on Artificial Neural Networks* 409: 442–447.

Tax, D. M. J., and R. P. W. Duin. 1999. "Support Vector Domain Description." *Pattern Recognition Letters* 20: 1191–1199. doi:10.1016/S0167-8655(99)00087-2.

Tax, D. M. J., and R. P. W. Duin. 2004. "Support VectorData Description." *Machine Learning* 54: 45–66. doi:10.1023/B:MACH.0000008084.60811.49.

Tax, D. M. J., and R. P. W. Duin. 2002. "Uniform Object Generation for Optimizing One-Class Classifiers." *The Journal of Machine Learning Research* 2: 155–173.

Theissler, A., and I. Dear. 2013. "Autonomously Determining the Parameters for SVDD with RBF Kernel from a One-Class Training Set." In *Proceedings of the WASET International Conference on Machine Intelligence* 7(7): 1135–1143.

United Nations. 2008. *Department of Economic and Social Affairs Population Division. World Urbanization Prospects: The 2007 Revision. No. 216*. New York: United Nations Publications.

Van de Voorde, T., T. De Roeck, and F. Canters. 2009. "A Comparison of Two Spectral Mixture Modelling Approaches for Impervious Surface Mapping in Urban Areas." *International Journal of Remote Sensing* 30: 4785–4806. doi:10.1080/01431160802665918.

Vapnik, V. 1979. *Estimation of Dependences Based on Empirical Data*, 5165–5184. Vol. 27. Moscow: Nauka. (in Russian) (English translation: New York: Springer Verlag, 1982).

Vapnik, V. N. 2000. *The Nature of Statistical Learning Theory*. New York: Springer.

Wang, J., X. Wu, and C. Zhang. 2005. "Support Vector Machines Based on K-Means Clustering for Real-Time Business Intelligence Systems." *International Journal of Business Intelligence and Data Mining* 1: 54–64. doi:10.1504/IJBIDM.2005.007318.

Waske, B., J. A. Benediktsson, and J. R. Sveinsson. 2009. " Classifying Remote Sensing Data with Support Vector Machines and Imbalanced Training Data." In *Multiple Classifier Systems*, 375–384. Berlin, Heidelberg: Springer.

Weng, Q. 2012. "Remote Sensing of Impervious Surfaces in the Urban Areas: Requirements, Methods, and Trends." *Remote Sensing of Environment* 117: 34–49. doi:10.1016/j.rse.2011.02.030.

Wolf, A. F. 2012. "Using Worldview-2 Vis-NIR Multispectral Imagery to Support Land Mapping and Feature Extraction Using Normalized Difference Index Ratios." In *SPIE Defense, Security, and Sensing, International Society for Optics and Photonics, 2012*, edited by S. S. Shen and P. E. Lewis, vol. 8390, Baltimore, MD, April 23–27. doi:10.1117/12.917717.

Wu, C., and A. T. Murray. 2003. "Estimating Impervious Surface Distribution by Spectral Mixture Analysis." *Remote Sensing of Environment* 84: 493–505. doi:10.1016/S0034-4257(02)00136-0.

Wu, M., and J. Ye. 2009. "A Small Sphere and Large Margin Approach for Novelty Detection Using Training Data with Outliers." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 31: 2088–2092. doi:10.1109/TPAMI.2009.24.

Yang, L., C. Huang, C. G. Homer, B. K. Wylie, and M. J. Coan. 2003. "An Approach for Mapping Large-Area Impervious Surfaces: Synergistic Use of Landsat-7 ETM+ and High Spatial Resolution Imagery." *Canadian Journal of Remote Sensing* 29: 230–240. doi:10.5589/m02-098.

Yang, X. 2006. "Estimating Landscape Imperviousness Index from Satellite Imagery." *IEEE Geoscience and Remote Sensing Letters* 3: 6–9. doi:10.1109/LGRS.2005.853929.

Yu, X., D. DeMenthon, and D. Doermann. 2008. "Support Vector Data Description for Image Categorization from Internet Images." In Proceedings of the 19th international conference on pattern recognition (ICPR'08), December 8–11, Tampa, FL.

Zeng, Z., F. Yun, G. I. Roisman, Z. Wen, Y. Hu, and T. S. Huang. 2006. "One Class Classification for Spontaneous Facial Expression Analysis." In *Automatic Face and Gesture Recognition, 7th International Conference*, 281–286. Southampton: IEEE.

Zheng, B., J. B. Campbell, and K. M. De Beurs. 2012. "Remote Sensing of Crop Residue Cover Using Multi-Temporal Landsat Imagery." *Remote Sensing of Environment* 117: 177–183. doi:10.1016/j.rse.2011.09.016.