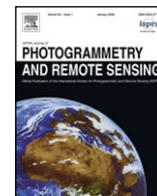




Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

Review article

Support vector machines in remote sensing: A review

Giorgos Mountrakis*, Jungho Im, Caesar Ogole

Department of Environmental Resources Engineering, SUNY College of Environmental Science and Forestry, 1 Forestry Dr, Syracuse, NY 13210, USA

ARTICLE INFO

Article history:

Received 6 June 2010

Received in revised form

17 September 2010

Accepted 1 November 2010

Available online 3 December 2010

Keywords:

Support vector machines

Review

Remote sensing

SVM

SVMs

ABSTRACT

A wide range of methods for analysis of airborne- and satellite-derived imagery continues to be proposed and assessed. In this paper, we review remote sensing implementations of support vector machines (SVMs), a promising machine learning methodology. This review is timely due to the exponentially increasing number of works published in recent years. SVMs are particularly appealing in the remote sensing field due to their ability to generalize well even with limited training samples, a common limitation for remote sensing applications. However, they also suffer from parameter assignment issues that can significantly affect obtained results. A summary of empirical results is provided for various applications of over one hundred published works (as of April, 2010). It is our hope that this survey will provide guidelines for future applications of SVMs and possible areas of algorithm enhancement.

© 2010 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). Published by Elsevier B.V. All rights reserved.

1. Introduction

Remotely-sensed data are used in numerous applications. Typically, an image classification process is initiated to convert data into meaningful information. Unfortunately, image classification is not a trivial task. As noted by Chi et al. (2008), classification of remote sensing data is particularly daunting because most of the supervised learning schemes require sufficiently large amount of training samples, yet definition and acquisition of reference data is often a critical problem. Various classification techniques, both parametric and non-parametric, have been developed and used in different contexts – remote sensing inclusive.

Previous reviews, such as that by Plaza et al. (2009), focused on recent developments in methodologies for processing a specific type of imagery, for example hyperspectral images. The review provided in this paper follows the algorithmic perspective rather than image characteristics. More specifically, we focus on applications of support vector machines (SVMs) in remote sensing. The motivation to carry out this study comes from different sources. First, SVMs are not as well-known as other classifiers (e.g., decision trees, variants of neural networks) in the general remote sensing community, yet they can match if not exceed the performance of established methods. Second, their performance

gains seem well-suited for remote sensing applications, where a limited amount of reference data is often provided. Third, even though the method is not widely popular, in recent years there has been a significant increase in SVM works on remote sensing problems suggesting this review is current and appropriate.

This review focuses on recent research papers (available by April, 2010) published in eight major journals of remote sensing, namely, ISPRS Journal of Photogrammetry and Remote Sensing, Remote Sensing of Environment, Photogrammetric Engineering & Remote Sensing, IEEE Transactions on Geoscience and Remote Sensing, IEEE Geoscience and Remote Sensing Letters, International Journal of Remote Sensing, Canadian Journal of Remote Sensing and GIScience and Remote Sensing. A limited number of research papers relevant to the thematic point and thus included in this review came from additional sources. The selected papers represent a wide range of: (i) applications from coal reserve detection to urban growth monitoring, (ii) resolutions from sub-meter to several kilometers pixel size, (iii) spectral resolution from single to hundreds of bands, and (iv) comparative methods from maximum likelihood classifiers to neural networks. For completeness, we first recap on the basics of SVM methodology before diving into specific works. Relevant papers are then summarized, while juxtaposition of general patterns enables us to derive conclusions and recommendations for further investigations.

2. Overview of support vector machines

Support vector machines (SVMs) is a supervised non-parametric statistical learning technique, therefore there is no assumption

* Corresponding address: Department of Environmental Resources Engineering, SUNY College of Environmental Science and Forestry, 419 Baker Hall, 1 Forestry Dr, Syracuse, NY 13210, USA. Tel.: +1 (315) 470 4824; fax: +1 (315) 470 6958.

E-mail address: gmountrakis@esf.edu (G. Mountrakis).

URL: <http://www.aboutgis.com> (G. Mountrakis).

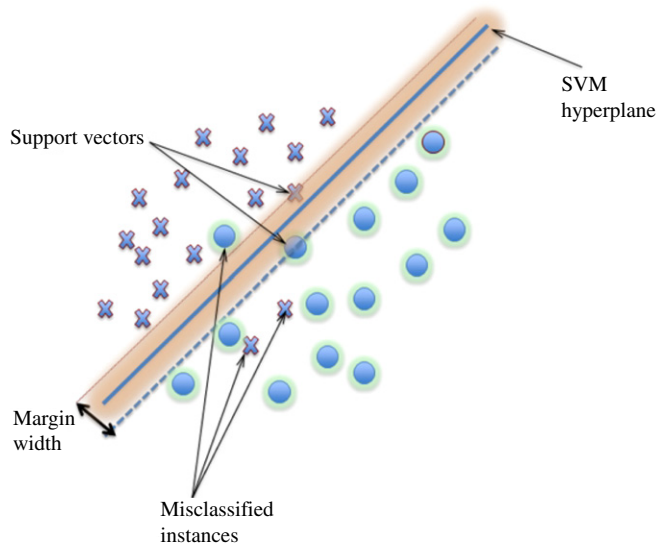


Fig. 1. Linear support vector machine example.
Source: adapted from Burges (1998).

made on the underlying data distribution. In its original formulation (Vapnik, 1979) the method is presented with a set of labeled data instances and the SVM training algorithm aims to find a hyperplane that separates the dataset into a discrete predefined number of classes in a fashion consistent with the training examples. The term optimal separation hyperplane is used to refer to the decision boundary that minimizes misclassifications, obtained in the training step. Learning refers to the iterative process of finding a classifier with optimal decision boundary to separate the training patterns (in potentially high-dimensional space) and then to separate simulation data under the same configurations (dimensions) (Zhu and Blumberg, 2002).

In its simplest form, SVMs are linear binary classifiers that assign a given test sample a class from one of the two possible labels. An instance of a data sample to be labeled in the case of remote sensing classification is normally the individual pixel derived from the multi-spectral or hyperspectral image. Such a pixel is represented as a pattern vector, and for each image band, it consists of a set of numerical measurements. Elements of the feature vector may also include other discriminative variable measurements based on pixel spatial relationships such as texture. Fig. 1 illustrates a simple scenario of a two-class separable classification problem in a two-dimensional input space. An important generalization aspect of SVMs is that frequently not all the available training examples are used in the description and specification of the separating hyperplane. The subset of points that lie on the margin (called support vectors) are the only ones that define the hyperplane of maximum margin.

The implementation of a linear SVM assumes that the multi-spectral feature data are linearly separable in the input space. In practice, data points of different class memberships (clusters) overlap one another. This makes linear separability difficult as the basic linear decision boundaries are often not sufficient to classify patterns with high accuracy. Techniques and workarounds such as the soft margin method (Cortes and Vapnik, 1995) and the kernel trick are used to solve the inseparability problem by introducing additional variables (called slack variables) in SVM optimization and mapping (using a suitable mathematical function) the non-linear correlations into a higher (Euclidean or the Hilbert) space, respectively. A kernel function typically needs to fulfill Mercer's Theorem in order to be a valid kernel in SVMs (Scholkopf and Smola, 2001). The choice of a kernel function often has a bearing on the results of analysis. Furthermore, typical remote sensing

problems usually involve identification of multiple classes (more than two). Adjustments are made to the simple SVM binary classifier to operate as a multi-class classifier using methods such as one-against-all, one-against-others, and directed acyclic graph (Knerl et al., 1990).

SVMs are particularly appealing in the remote sensing field due to their ability to successfully handle small training data sets, often producing higher classification accuracy than the traditional methods (Mantero et al., 2005). The underlying principle that benefits SVMs is the learning process that follows what is known as structural risk minimization. Under this scheme, SVMs minimize classification error on unseen data without prior assumptions made on the probability distribution of the data. Statistical techniques such as maximum likelihood estimation usually assume that data distribution is known a priori. Burges (1998) in a well-organized SVM tutorial described a simple experiment to illustrate an advantage of SVMs in an image recognition problem. In that demonstration, the performance of a basic multi-way SVM-based recognizer was assessed on image classification in the presence of prior knowledge. The accuracy turned out to be approximately the same if the pixels were first shuffled, with each image instance suffering the same random permutation. Yet, when the act of 'vandalism' (or removal of prior knowledge) took place, SVM still outperformed even the best neural networks. This discovery is particularly appealing in remote sensing applications since data acquired from remotely sensed imagery usually have unknown distributions, and methods such as Maximum Likelihood Estimation (MLE) that assume a multivariate normal data model do not necessarily match that assumption. Even if the data, whose dimensionality is assumed to match the number of spectral bands, were normally distributed, the assumption that the distribution can be described using a bell-shaped (Gaussian) function ceases to be sound, since the concentration of data in higher dimensional space tends to be in the tails (Fauvel et al., 2009). This phenomenon will continue to be encountered in remote sensing as new sensors increase spectral resolution and therefore data dimensionality.

There is also another interesting concept that serves as a key attraction to SVMs. Commonly described by many authors under the notion of overfitting (Montgomery and Peck, 1992), yet variously referred to by others as bias-variance tradeoff (Geman et al., 1992) or capacity control (Guyon et al., 1992), SVM-based classification has been known to strike the right balance between accuracy attained on a given finite amount of training patterns and the ability to generalize to unseen data.

Alongside the benefits derived from the SVM formulation there are also several challenges. The major setback concerning the applicability of SVMs is the choice of kernels. Although many options are available, some of the kernel functions may not provide optimal SVM configuration for remote sensing applications. Empirical evidence indicates that kernels such as radial basis function and polynomial kernels applied on SVM-based classification of satellite image data produce different results (Zhu and Blumberg, 2002). A good explanation on SVM kernels and their functionality is presented in numerous papers (e.g., Kavzoglu and Colkesen, 2009). From the non-expert user point of view, SVM theory is a bit intimidating, particularly due to the fact that the more efficient SVM variants often incorporate some difficult to understand concepts. This limits effective cross-disciplinary applications of SVMs.

Numerous SVM tutorials are available (such as Cortes and Vapnik (1995) and Burges (1998)), but none of these contains an exhaustive discussion on the increasing number of newly proposed variants of SVMs. In the remote sensing field a good starting point would be a textbook by Tso and Mather (2009) that provides a review of the entire field of classification methods for remotely sensed data, including SVMs. For those interested in rule extraction from SVMs a recent computer science review is available (Barakat

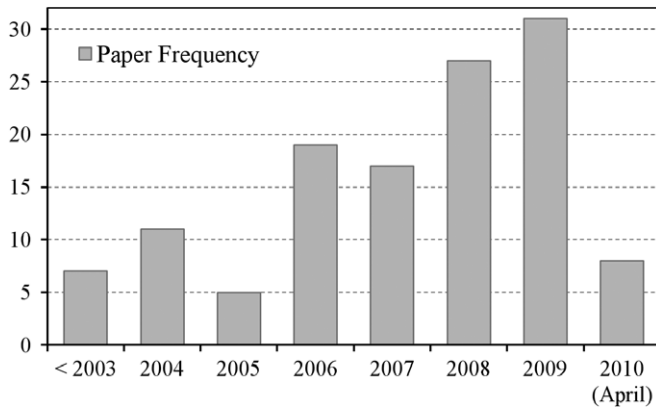


Fig. 2. Growth of SVMs popularity in remote sensing over the past decade.

and Bradley, in press). Chen and Ho (2008) provide an excellent general reference for statistical learning in remote sensing. It should be noted that for this review the term SVM is inclusive of the traditional SVM method as well as SVM-based variants, since most of the latter still heavily rely on the standard SVM method.

3. Brief overview

Support vector machines (SVMs) have recently found numerous applications in remote sensing. For this review we identified 108 relevant papers, with more than half published in the last 2.5 years (Fig. 2). This increasing trend is expected to grow, making this a critical time for a review of existing work.

The SVM papers included a wide range of remote sensing application domains and sensors. A summary of this diverse group is presented in Fig. 3. Satellite sensors are preferred, especially multispectral ones. There is some limited interest in change detection (10% of the papers), a pattern that is expected to significantly increase as the Landsat archive is now freely available. There is an almost equal split between high and medium resolution sensors, mostly related to a strong preference to Ikonos and Landsat imagery, but also to high resolution airborne sensors.

4. SVM works focusing on algorithmic advancements

This section summarizes SVM advancements that were achieved during the past decade. Papers that merely contrasted SVM performance with other methods or papers incorporating SVMs for a specific application are discussed in the next section.

4.1. Classification

SVMs are typically a supervised classifier, which requires training samples. Literature shows that SVMs are not relatively sensitive to training sample size and scientists have improved SVMs to successfully work with limited quantity and quality of training samples. For example, Foody and Mathur (2004b) showed that only a quarter of the original training samples acquired from SPOT HRV satellite imagery was sufficient to produce an equally high accuracy for a two-crop classifier. Mantero et al. (2005) estimated probability density of thematic classes using an SVM. The SVM-based approach used a recursive procedure to generate prior probability estimates for known and unknown classes by adapting the Bayesian minimum-error decision rule. The approach was tested using synthetic data and two optical sensor data (i.e., Daedalus ATM and Landsat TM) and confirmed method effectiveness, especially when the availability of ground reference data was limited. Transductive inference learning theory was incorporated into an SVM for remote sensing classification in Bruzzone et al. (2006). Their SVM-based approach defines the separating hyperplane according to a process that integrates the unlabeled samples together with the training samples. Experiments showed that the proposed method was effective, particularly for a set of ill-posed remote sensing classification problems due to the limited training samples. Foody and Mathur (2006) proposed a focus on mixed pixel training samples over more tedious, conventional pure pixel acquisition, assuming an SVM classifier. The analysis of a three-waveband multispectral SPOT HRV image showed the benefits of mixed pixel sampling on a crop type classification task. Foody et al. (2006) evaluated four dataset reduction methods for a one-class problem

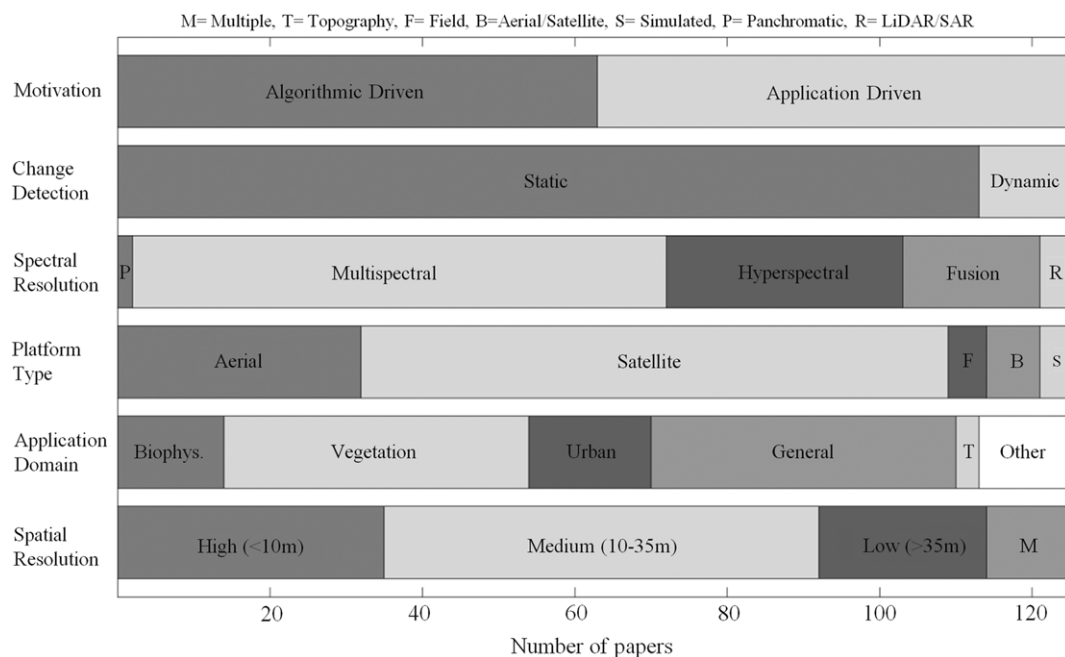


Fig. 3. Summary statistics of selected works.

(cotton vs. others) using SVMs and LISS-III data and found that significant data reduction was feasible ($\sim 90\%$) with minimal information loss.

Sahoo et al. (2007) investigated the incorporation of localized, highly sensitive transformations to capture subtle changes in hyperspectral signatures. They compared the so called S-transform to classifiers without it and found encouraging results. The implementation algorithm was an SVM that showed additional robustness to small data samples in a geological classification. Blanzieri and Melgani (2008) investigated a local k -nearest neighbor adaptation to formulate localized variants of SVM approaches. Their results indicated substantial improvements, especially with the integration of non-linear kernel functions. Tuia and Camps-Valls (2009) addressed the issue of kernel predetermination by proposing a regularization method that identifies kernel structure through analysis of unlabeled samples. Camps-Valls et al. (2010) proposed an improved methodology for assessing kernel independence in various imagery types using the Hilbert–Schmidt independence criterion. Marconcini et al. (2009) discussed the incorporation of spatial information through composite kernels finding substantial improvements however with an additional computation cost. Camps-Valls et al. (2008) proposed a methodological framework using composite kernels for multi-temporal classification of remote sensing data from different sources. The method was tested using both synthetic and real optical Landsat TM data and found that the cross-information composite kernel was the best in general, but a simple summation kernel also showed similar performance. Composite kernels that take advantage of the properties of Mercer's kernels were further discussed in their prior work (Camps-Valls et al., 2006c). Chi et al. (2008) proposed a method, called primal SVM that is capable of differentiating land covers using a reasonably small amount of training examples. Their method sought to replace the regularization-based approach previously employed in SVMs. The primal SVM formulation makes it possible to optimize directly on the primal representation, and therefore limits the number of samples. Evaluation was performed using Hyperion imagery of the Okavango Delta (in Botswana) for vegetation classification. Primal SVM yielded competitive accuracy values as the state-of-art alternative algorithms trained on larger datasets. Gómez-Chova et al. (2008) investigated the addition of a regularization term on the geometry of both labeled and unlabeled samples that was based on graph Laplacian, leading to a Laplacian SVM variant. This semisupervised classification method offers improvements when compared with traditional SVMs, especially in small training datasets and underlying complex problems. Castillo et al. (2008) proposed a modified version of the SVM classifier, called bootstrapped SVM. The training strategy adapted in the bootstrapped SVM is such that an incorrectly classified training sample in a given learning step is removed from the training pool, re-assigned a correct label, and re-introduced into the training set in the subsequent training cycles. The key result was the ability to capture data variability in a highly biased binary dataset, only 0.05% of the total number of training pixels were needed to achieve about the same accuracy level as the standard SVM. An interesting SVM adaptation was proposed by Wang and Jia (2009), where the space between support vectors is considered to provide a soft classification in addition to the traditional hard classification. Demir and Erturk (2009) offered an improvement to hyperspectral SVM classifiers by incorporating border training samples in a two step classification process. Song et al. (2005) proposed an SVM adaptation for Landsat-based vegetation monitoring. The SVM ν parameter was tackled through an integration of one and two class SVM sequential classification steps.

Mathur and Foody (2008b) investigated methods for efficient reduction of field data. They concluded that for cropland mapping equivalent classification results can be obtained with a third of

the original dataset assuming SVM methods are used for the classification process. At the 24 m ground pixel size acquired by the LISS-III sensor the reduced dataset yielded a small 1.34% accuracy loss at 90.66%.

Integration of a genetic algorithm (GA) and SVM for remote sensing classification was evaluated with a limited availability of training samples in Ghoggali et al. (2009). The experimental results revealed again an ability to improve classification accuracy with a small training sample size. However, the computational load was significant mainly due to the slow GA convergence. Ghoggali and Melgani (2008) integrated genetic training into SVM classification in order to incorporate land cover transition rules in multi-temporal classification. The results indicated a mixed performance, however the algorithmic flexibility and humanly intuitive process suggest promising future work. Bruzzone and Persello (2009) proposed a novel context-sensitive semi-supervised SVM classification model, which can be successfully utilized when some of training data are not reliable. Their model explores the contextual information of the neighboring pixels of each training sample and improves the unreliable training data. They tested their model using Ikonos and Landsat TM data and compared the results with those based on some of the widely used classification algorithms such as the standard SVM, a progressive semi-supervised SVM, maximum likelihood and k -nearest neighbor. The proposed SVM algorithm outperformed the other classification models in terms of robustness and effectiveness, particularly when non-fully reliable training samples were used. Huang and Zhang (2010) compared multi-SVM methods with traditional vector stacking techniques on high resolution urban mapping.

Su (2009) investigated training data reduction using a hierarchical clustering analysis and Multiangle Imaging SpectroRadiometer (MISR) satellite data (250 m–1.1 km, 17 products) on a vegetation classification problem. It was shown that a two thirds reduction of the dataset size was possible without significant accuracy degradation in SVM and maximum likelihood classifier (MLC) methods. Gomez-Chova et al. (2010) proposed a method to increase classification reliability and accuracy by combining labeled and unlabeled pixels using clustering and the mean map kernel. They tested their approach to classify clouds using Envisat's Medium Resolution Imaging Spectrometer (MERIS) data. They found that their method was particularly successful when sample selection bias (i.e., training and test data follow different distributions) exists.

Selecting an optimum SVM method for remote sensing classification is not an easy task. Foody and Mathur (2004a) proposed a single multiclass SVM classification method while typical multiclass SVMs are based mainly on the use of multiple binary analyses. They compared their approach with other classification methods such as discriminant analysis, decision trees, and neural networks, and found that the SVM-based approach outperformed the other methods with different sizes of training samples. Bazi and Melgani (2006) investigated the most appropriate feature subspace and model selection based on a genetic optimization framework using three feature selection methods including steepest ascent, recursive feature elimination technique, and the radius margin bound minimization method. They used two criteria, the simple support vector count and the radius margin bound, to identify an optimum SVM-based classification system for hyperspectral remote sensing data. The genetically optimized SVM using the support vector count as a criterion resulted in the best performance for both simulated and real-world AVIRIS hyperspectral data. Mathur and Foody (2008a) evaluated the performance of SVMs in non-binary classification tasks. Their results indicated their proposed one shot SVM classifier outperformed the binary-based multiple classifiers in terms of obtained accuracy but also initial parameterization.

SVMs have also been used for feature selection. Pal (2006) investigated methods for feature selection based on SVMs. Citing the

unreasonably large computational requirements as a major disadvantage of exhaustive search methods in practical applications, the researchers justified the use of a non-exhaustive search procedure in selecting features with high discriminating power from large search spaces. SVM-based methods combined with GA were compared with the random forest feature selection method in land cover classification problems with hyperspectral data and small benefits were identified. Zhang and Ma (2009) addressed the issue of feature selection in SVM approaches. They implemented a modified recursive SVM approach to classify hyperspectral AVIRIS data. The reduced dimensionality returned slightly better results, however their method has higher computational demands compared with others. On the same subject Archibald and Fann (2007) provided an interesting integration of feature selection within the SVM classification approach. They achieved comparable accuracy while significantly reducing the computational load.

Some studies improved the performance of SVM-based classification through algorithm and/or data fusion. Zhang et al. (2006) proposed a pixel shape index describing the contextual information of nearby pixels and evaluated its usability for land cover classification using QuickBird data based on SVMs. The pixel shape indices were combined with transformed spectral bands such as principal component analysis or independent component analysis. They found that integration of spectral and shape features as well as the transformed spectral components in an SVM were able to improve classification accuracy. Waske and Benediktsson (2007) classified multi-sensor (SAR, Landsat TM, and SPOT) and multi-temporal data through data fusion based on SVMs. Their method was based on the decision fusion of multiple SVMs that were individually trained on the different data sources. Their approach outperformed the other methods including maximum likelihood, decision trees, and a typical SVM. Mitra et al. (2004) proposed an active learning-based approach to reduce the selected support vectors. Their semi-supervised method gradually creates clusters based on interactive user input. Their method yielded better results than a typical SVM, however the authors caution on the algorithm's sensitivity on user-provided erroneous labeling. Zhang and Ma (2008) proposed an SVM variant, the Potential SVM as an alternative for multispectral image classification. The Potential SVM is an attractive variant due to its ability to handle non-Mercer kernels and its mathematical formulation that addresses SVM scalability issues. Tests on very high (0.1 m) and medium (30 m) resolution indicated equal or better accuracy than the traditional SVM, while offering faster simulation times due to support vector reduction. A fusion approach to classification using extended morphological profiles was proposed in Fauvel et al. (2008). They evaluated the approach using high spatial/spectral resolution ROSIS data in urban areas based on SVM classification. Ensemble methods for multiple SVM integration were evaluated by Pal (2008). Two popular integration techniques such as boosting (alternating observation weight) and bagging (alternating observations) were tested using Landsat ETM+ data for an agricultural classification. The findings suggest that an optimized ensemble method could lead to improved results, though further testing is suggested as others have found contradictory results. Chen et al. (2009) proposed an improved classification method by stacking multiple hierarchical SVM classifiers. The method also incorporates discrimination information of two feature spaces (i.e., magnitude and shape). Experiments showed that the method with the generalization ability and the use of multiple feature spaces was effective for hyperspectral image classification. Chen et al. (2008) investigated the integration of SVMs with pairwise decision trees on hyperspectral data. The one-against-one SVM adaptation provided similar results to their proposed method, which was attributed to the hierarchical structure of the decision tree-based method. Demir and Ertürk (2007) implemented a relevance vector machine (RVM) approach, which was originally proposed by Tipping (2000, 2001),

for vegetation mapping using hyperspectral imagery. Obtained results indicated a significantly lower usage of classification vectors, however a lower accuracy rate was obtained compared to the typical SVM approach. The authors suggest their method as a viable solution for real time classification due to the highly efficient simulation times. Tuia et al. (2009) combined morphological filters and SVMs to conduct land use classification using high spatial resolution QuickBird panchromatic images. They tested multiple morphology-based features and found that simple morphological features generated with opening and closing operators resulted in the best performance. Muñoz-Marí et al. (2007) provided an interesting comparison of available one-class classifiers for both single and multiple class remote sensing problems. They also investigated a one-class classifier called support vector domain description (SVDD) that is particularly attractive in the presence of incomplete training data. Tan et al. (2007) proposed a new technique combining entropy decomposition and SVM for classification. The approach was tested using multi-temporal SAR images for rice monitoring. Their approach was especially useful when retrieving polarimetric information for each class resulting in good separation between classes. Tarabalka et al. (2009) proposed a new classification scheme emphasizing both spectral and spatial characteristics of hyperspectral images. Their method combined the pixel-wise SVM classification results and the segmentation map based on partitional clustering using the majority voting strategy. The approach was specifically useful when large spatial structures were included in data or when different classes had dissimilar spectral responses and a comparable number of pixels.

Although SVMs are typically employed for supervised classification tasks, they have also been used for unsupervised classification in combination with other techniques. For example, Bovolo et al. (2008) combined an SVM and a selective Bayesian thresholding approach for unsupervised change detection. They used a selective Bayesian thresholding to delineate pseudo-training samples and conducted binary change detection (i.e., change vs. no change) using the samples based on an SVM approach. Their method outperformed the change vector analysis (CVA)-based method with the expectation-maximization algorithm, but required much longer computational time due to the model-selection strategy to identify an optimum structure of their model. Mukhopadhyay and Maulik (2009) integrated a multi-objective fuzzy clustering scheme with an SVM for unsupervised classification. Their method identified high-confidence points from certain clusters to train the SVM classifier. The method was tested using several satellite images (i.e., SPOT, Landsat TM, and IRS) and concluded that their method was more effective when compared to other methods such as neural networks, *k*-nearest neighbor, and fuzzy *c*-means.

4.2. Regression

In addition to classification tasks, SVMs have been advanced to solve regression problems, where in essence a continuous prediction output is expected. A multiple estimator system for biophysical parameter estimation from remote sensing data was proposed by Bruzzone and Melgani (2005). They particularly focused on incorporation of SVMs into the system and combination with multilayer perceptron (MLP) neural networks. They simulated different operational conditions with SVM and MLP and pointed out that their system increased the robustness of the estimation process; it provided accuracy very close to that of the best estimator included in the ensemble based on the experiment of chlorophyll concentration estimation using MERIS data. Camps-Valls et al. (2006a) investigated a RVM-based approach, a variant of SVMs, in order to lessen the uncertainty inherent in handling satellite-derived and *in-situ* measurements of oceanic chlorophyll concentration. RVM

incorporated prior knowledge of the problem and proved to be useful in quantifying chlorophyll concentrations based on ocean surface reflectance. The technique was reported to be less sensitive to parameter selection; it also provided a considerable high accuracy despite sparsity in the solution space. The authors recommended use of RVMs in other applications that involve estimation of biophysical parameters using remotely sensed data because of its robustness to even small amount of training data and low sensitivity to free parameter setting. Camps-Valls et al. (2006b) further investigated an SVM variant providing a more robust regression model for ocean chlorophyll concentration that proved successful, especially with limited training samples.

A synthetic algorithm of wavelets and SVMs was developed to predict evapotranspiration by Kaheil et al. (2008). A range of remote sensing-derived input variables such as MODIS LAI, MODIS emissivity, and spectral data of Landsat TM and ASTER were fed into the algorithm to produce a spatial distribution of evapotranspiration at the finest spatial resolution of the input data. Moser and Serpico (2009) proposed an automatic parameter optimization method for SVM regression of land and sea surface temperatures. They tested their method using AVHRR and MSG satellite images with synchronous *in situ* measurements and compared with typical grid-search based optimization methods such as cross-validation and hold-out. The proposed method resulted in similar accuracy with the other methods, but much more efficient than them, particularly when a high number of training samples was available.

Regression SVM has also been used for data generation and fusion. Zheng et al. (2008) proposed a multiscale mapped least-squares SVM (LS-SVM) to sharpen multispectral bands using a higher resolution panchromatic band. QuickBird data were used in the experiments and multiscale Gaussian radial basis function kernels were incorporated. Their method was compared with other fusion algorithms such as discrete wavelet transform, curvelet transform, atrous wavelet transform, extended fast IHS and found that both their method and the atrous wavelet transform resulted in the best performance. Shi et al. (2009) also used an LS-SVM approach to generate a digital surface model (DSM) from Light Detection And Ranging (LiDAR) data. Assessed visually and quantitatively against the radial basis function (fastRBF) and triangulation technique, LS-SVM was found to be more effective in terms of noise reduction, computational efficiency and accuracy in DSM generation. This reformulation of the standard SVM based on regression models bears similarity with regularization networks and Gaussian processes. LS-SVM incorporates pixel neighborhood and topographic analyses. As such, basic principles of differential geometry play a key role in generation of gradient and curvature equations, and in other such related tasks. Readers interested in SVM-based regression models could find useful information in Smola and Schölkopf (2004).

5. Application-oriented SVM papers

This section presents papers where the incorporated SVMs were not highly customized; instead the focus was their evaluation under a given task. Where applicable, results from works that contrasted SVMs with other methods are mentioned.

5.1. Biophysical tasks

SVMs have been used in remote sensing-based estimation and monitoring of biophysical parameters such as chlorophyll concentration, gross primary product, and evapotranspiration. For example, Kwiatkowska and Fargion (2003) employed SVMs to cross-calibrate the global chlorophyll concentration from different satellite sensors (i.e., SeaWiFS and MODIS). The goal was to

extrapolate this cross-calibration knowledge on to other data in a different spatio-temporal domain and to identify representative products for the global chlorophyll concentration. They revealed there were significant discrepancies between the different sensor products; there is a high dependency on sensor calibrations and operational characteristics. Bazi and Melgani (2007) estimated chlorophyll concentrations in coastal waters based on a particle swarm optimization and SVM techniques using MERIS and SeaBAM data. They found that their method was more effective than the typical SVM and less sensitive to training sample size. Sun et al. (2009) investigated *in situ* hyperspectral measurements to estimate chlorophyll concentration in Lake Taihu using SVMs. They first identified the best three-wavelength factor using an iterative optimization and used them as inputs to an SVM to estimate chlorophyll concentration. Their approach proved more accurate than the typical linear regression models.

Knudby et al. (2010) studied reef fish richness, diversity and biomass using Ikonos images and predictive modeling. SVMs were compared with five other methods and performed almost equally to the highest ranked ensemble algorithms. Clevers et al. (2007) tested an SVM-based band shaving technique to reduce dimensionality in hyperspectral datasets. The application domain was grassland biomass estimation and three bands were identified as sufficient for field studies. Durbha et al. (2007) assessed leaf area index extraction for Multiangle Imaging SpectroRadiometer (MISR) satellite data. They proposed an adjusted support vector regression method which included parameter regularization.

An SVM-based model was used to calculate global ocean primary productivity (Tang et al., 2008); it was found to be more accurate than the vertically generalized production (VGPM) approach due to its ability to identify the nonlinear relationship between the ocean's primary productivity and other parameters. The problem was particularly difficult because of the sparse nature of the data. SVMs performed better than the traditional VGPM making them appealing for undersampled applications such as oceanographic studies. Yang et al. (2007) modeled continental gross primary product using MODIS and other sources; SVMs were the underlying algorithmic methodology.

Xie et al. (2008) implemented SVR to calculate the moisture transport in oceanic environments using MISR and they found SVR outperforming linear regression and backpropagation neural networks. Yang et al. (2006) estimated evapotranspiration by combining MODIS and AmeriFlux data using SVMs at the continental scale. SVMs outperformed neural networks and multiple regression.

5.2. Land cover land use tasks

5.2.1. Vegetation/agriculture

In an early work, Gualtieri and Crompton (1998) evaluated SVM performance on vegetation classification. Hyperspectral AVIRIS imagery was used and results suggested SVM superiority over prior classifiers developed on the same dataset. Keramitsoglou et al. (2006) focused on vegetation mapping using Ikonos imagery. They contrasted SVMs with Kernel-based spatial Re-Classification (KRC) and RBF neural networks and found that even though SVMs showed slightly less robustness in the classification results over the RBF, their training time was considerably lower suggesting improved applicability. KRC also performed well but not as high as the SVM and RBF methods. Knorn et al. (2009) evaluated binary forest classification using SVMs in a spatial sequence of Landsat scenes. The major goal was to assess chain classification accuracy, which proved accurate even for lengthy sequences (e.g., six images) as long as image overlapping portions represented well the different features on the ground.

Huang et al. (2008b) performed SVM-based classification to assess in forest classification accuracy the influence of the

slope/aspect of the terrain, solar elevation and azimuth and the relative position of the trees. A 3.6% gain in overall classification accuracy was realized after topographic correction. Lardeux et al. (2009) used SVMs to classify dense tropical vegetation with SAR data. SVMs resulted in about 20% higher classification accuracy than the Wishart classification approach. They pointed out that SVMs can perform much better than the typical Wishart approach when radar data do not follow the Wishart distribution. SVMs were also evaluated against decision tree classifiers in a study that involved mapping of dynamic semi-natural habitat systems – technically known as fenlands (Boyd et al., 2006). In this problem, SVMs were implemented as a binary classifier, to classify the data into fen and ‘other’, while the ‘other’ class contained samples of several ground features. The performance of SVMs was slightly higher than that of the decision tree classifier.

Looking into forest species classification, Dalponte et al. (2008) used SVMs and data fusion of hyperspectral (AISA) and LiDAR data. SVMs outperformed Gaussian maximum likelihood classification and *k*-NN technique. They pointed out that the incorporation of LiDAR variables generally improved the classification performance and the first return data was the highest contributing factor. Heikkinen et al. (2010) applied a simulated optical radiation model to evaluate the tree species classification using an airborne four band sensor system. They employed SVMs with different kernel functions and found that the four bands were not sufficient to get successful classification results; the Mahalanobis kernel provided the best accuracy. Dalponte et al. (2009), in their study on hyperspectral image acquisition and analysis, focused on the choice of spectral resolution and associated method of classification. The study goal was to classify complex forest scenarios. Simulated data (consisting of degraded band sizes from 4.6 to 36.8 nm) was used to analyze the role of spectral resolution on classification accuracy in an investigation to determine the trade-off between spectral and spatial resolution. SVM-based classification resulted in higher accuracies than all other classifiers for all spectral bands simulation. The authors attributed this to the effectiveness of SVMs in managing the complex hyperspectral classification. Another interesting conclusion was that different classifiers exhibited variable behavior with respect to spectral resolution.

With a focus on forest degradation, Cao et al. (2009a) proposed a burn index using MODIS data. An SVM method was implemented as part of an iterative classifier targeting burn scar mapping. The results were accurate when compared with Landsat-derived reference data, however the method is also constrained by hotspot identification accuracy and the presence of clouds. Liu et al. (2006) investigated the use of high resolution (GSD 1 m), four band aerial photography for forest disease monitoring. Their findings indicated that a spatial-temporal contextual approach improved the initial classification results obtained with an SVM method. Using images captured by Landsat TM/ETM+ between 1988 and 2007, Kuemmerle et al. (2009) applied an SVM classifier to detect illegal logging in the Ukrainian Carpathians. The classification problem focused on mapping forest cover change in the subregions. Although no comparative assessments were carried out, SVM proved a very useful method in delineating forest/non-forest cover maps for all the stated time periods. Another study by Huang et al. (2008a) used Landsat TM and Landsat ETM+ images with a focus on developing an automated solution to forest cover change detection. The classification took place using SVMs and an extensive evaluation over multiple sites indicated an accuracy of approximately 90%.

Su and Huang (2009) conducted a study in southern New Mexico to evaluate the effect of different linkage techniques on classification accuracy for semiarid vegetation mapping. Four different linkage techniques were tested to calculate distances

between clusters in an attempt to create a hierarchical structure of the training dataset and reduce its size. Results indicated that a reduced dataset of approximately 20% of the original size could provide comparable classification accuracy. Su et al. (2009) discussed further the application of SVMs on MISR imagery to detect semi-arid vegetation areas, where SVMs performed significantly better than MLC.

Undertaking a crop classification task, Wilson et al. (2004) investigated salt marsh and crop plants that have been exposed to heavy metal or petroleum toxicity with control treatments using in situ spectroradiometer measurements. They used two classification methods, SVMs and logistic discrimination based on partial least squares compression, and found that the SVM-based method was superior. SVMs were also implemented for crop classification using HyMap hyperspectral imagery in Camps-Valls et al. (2004). SVMs outperformed typical neural networks in terms of accuracy, simplicity, and robustness. They also found that SVMs were not as sensitive to training sample size, and SVMs were able to successfully detect noisy bands. Hyperspectral image data of a cornfield, acquired through airborne mission (Compact Airborne Spectrographic Imager) was used in conjunction with the SVM method in automatic detection of weeds and nitrogen in the field (Karimi et al., 2006). The discriminant features were based on the general remote sensing principle: corn exhibits different spectral responses depending on the type or method of weed control used and nitrogen application rates. Waske and van der Linden (2008) segmented multi-sensor data (SAR and TM) at multi-levels and pre-classified each individual level of segmentation using SVM for crop classification. The pre-classification results were then fused to create a final classification output with an SVM and random forests as decision rules. They pointed out that it was more appropriate to define the kernel functions for each data source and level separately; their multiple classifier system improved the performance compared to a single classifier approach since the individual errors of multiple data sources at different aggregation scales were diverse and uncorrelated.

5.2.2. Impervious surfaces/Urban areas

Huang and Zhang (2009) targeted road extraction from Ikonos imagery. The underlying idea was to integrate spectral and shape characteristics at multiple scales. In every scale an SVM method was implemented and later results from each scale were fused leading to improved centerline extraction. Another road extraction work using Ikonos imagery was published by Song and Civco (2004). SVM methods were developed to create a binary road layer, that was further processed with shape-assisted and vectorization procedures. The SVM method yielded a lower classification error compared to the Gaussian maximum likelihood approach, a finding which the investigators attributed to the assumption that class signatures (feature groups) follow a normal distribution may not always be appropriate. Inglada (2007) implemented SVMs to classify man-made features (e.g., bridges, roads, roundabouts) from 2.5 m SPOT 5 imagery. Classifier robustness and resilience to variability in illumination and changes in spectral bands was achieved by incorporating invariant geometric features. The results were reasonable (~80% accuracy) considering the complexity of the underlying problem. SVMs were also used to classify bridges from Ikonos high-resolution panchromatic image data. Luo et al. (2007) used a simple yet effective contextual idea: bridges, in general, are adjacent to water and water is usually darker than other objects. Additional steps followed from this assumption. Gauss Markov Random Field SVM, an SVM adjustment that incorporates texture properties, was used to enhance classification performance of the traditional SVM. Higher overall accuracy and kappa values were recorded.

ASTER imagery was used as input to an SVM-facilitated processing technique in an investigation to determine SVM's suitability for mapping urban areas (Zhu and Blumberg, 2002). Different results were obtained depending on image resolution and on SVM kernel choices, namely the polynomial kernel and radial basis function (RBF). The RBF-based SVM yielded higher performance with respect to convergence speed; better classification precision on the sample data was achieved using an SVM based on the polynomial kernel. Esch et al. (2009) combined single date Landsat images to derive useful information about industrial, residential and transportation-related areas in Germany. SVMs were found to be effective in automatic estimation of impervious surfaces. The authors conclude that automatic extraction of urban areas is a difficult problem owing to the wide range of surface materials, and the heterogeneity of the classes. Brown et al. (1999) evaluated SVMs with linear spectral mixture models (LSMM) for land use subpixel analysis. Their analysis on Landsat data for binary urban classification revealed under certain circumstances the LSMM is identical to linear SVMs. Walton (2008) compared urban subpixel classification performance from random forests, rule-based regression and SVMs using a Landsat image. Results indicated that the rule-based regression using Cubist provided improved accuracy and training time. Watanachaturaporn et al. (2008) found that SVM methods outperformed backpropagation and radial basis function neural networks, maximum likelihood and decision trees. Imagery from Indian's Linear Imaging Self-scanning Sensor (LISS) III was used (23.5 m pixel size, 4 bands) for an urban-driven classification.

The effects of off-nadir collection and vegetation cover on urban classification were investigated using hyperspectral, 4 m GSD aerial images (Linden and Hostert, 2009). SVMs were employed for the classification process but were not compared with other methods as it was outside the scope of their study. Confronted with the common challenge of selecting an appropriate set of parameter values, Cao et al. (2009b) used SVMs to overcome the setbacks of empirical trial and error methods in extracting urban areas from available samples of Defense Meteorological Satellite Program – Optical LineScan (DMSP-OLS) and SPOT-derived NDVI data. The study employed Chinese city datasets (apparently because of the rapid urbanization of the study area) and the problem was reduced to a non-threshold binary classification. Being non-parametric, SVMs proved to be a better choice for constructing a region-growing algorithm that semi-automatically discriminated urban pixels from any other type of background data. The main attraction was the ability of SVMs to achieve higher accuracy using a small number of training samples.

Nemmour and Chibani (2006) studied urban change detection using Landsat scenes and found SVM methods outperformed backpropagation neural networks. Interestingly, they found that user-defined SVM parameters did not have a significant influence in the SVM superiority. Another urban change detection study used multi-source data from Landsat TM/ETM+, European Remote Sensing Satellite (ERS) 1 and 2, and Advanced Synthetic Aperture Radar (ASAR) onboard the Environmental Satellite (ENVI-SAT) to map urban footprints in 1990, 2000 and 2006 (Griffiths et al., 2010). The classification method employed used SVMs and the authors developed an SVM-based forward feature selection procedure to rank input variable contribution. Licciardi et al. (2009) presented the five awarded algorithms useful for the classification of high resolution hyperspectral data over urban areas at the 2008 Data Fusion Contest. They found that SVMs were extremely useful for classification of hyperspectral data and decision fusion using multiple algorithms would be a way to go for future research regarding remote sensing classification.

5.2.3. General land cover land use tasks

Starting with high resolution imagery, Li et al. (2010) proposed an SVM-based classifier using QuickBird data. A scene segmentation algorithm was integrated with the SVM object classifier leading to better performance. It is also noted that the SVM classifier is highly dependent on the segmentation process, a typical drawback of object-based classifiers. Linear support vector machines were reported to be useful in classification of hyperspectral remote sensing data whose elements had been extracted using a technique called kernel principal component analysis (KPCA) (Fauvel et al., 2009). Although only the basic SVM was employed in the set of experiments, the improved feature provided a significant clue on the effectiveness of SVMs especially when applied on reliably clean datasets. Warner and Nerry (2009) performed a study to determine the effectiveness of thermal infrared data in land cover classification. An SVM classifier turned out to be an effective method at handling the complex distributions of the heterogeneous land cover classes that characterized the study area (Strasbourg, France). In their conclusion, the authors suggest that the inclusion of a single broad thermal band increased classification accuracy by as much as 20% for simulated Ikonos bands and provided a 4% improvement when hyperspectral VNIR and SWIR data were used.

In yet another remote sensing application, Huang et al. (2008c) presented an algorithmic fusion methodology to improve processing of very high-resolution (VHR) satellite imagery using a wavelet transformation. In justification of their undertaking, scientists cited that VHR images are characterized by complex multi-scale spectral and spatial information, therefore rendering the traditional fixed, single-band, single-window approach less efficient. The more relevant multi-scale spectral-spatial features were classified using support vector machines. Mladinich (2010) assessed three commercial software packages for object-based binary classification (disturbed vs. non-disturbed areas) over high resolution imagery (1 m). The ENVI software package was one of the three tools, and it incorporated an SVM algorithm adjusted from the library for support vector machines (LIBSVM). Results across the three tools were comparable, with Definiens classification showing higher consistency. In a bid to compare SVMs against maximum likelihood and backpropagation artificial neural networks, Pal and Mather (2005) experimented on Landsat 7 ETM+ and hyperspectral data. Results suggested SVM superiority as input dimensionality increased and as dataset size decreased.

Moving towards medium resolution imagery (15–30 m pixel size), in one of the earlier investigations Huang et al. (2002) provided an accuracy evaluation of SVMs versus three other classifiers, namely a MLC, a three-layer (input, hidden and output) backpropagation neural network classifier (NN), and a decision tree classifier (DTC). Variations of SVM classification results with different kernel configurations were also compared. The results showed that SVMs had the highest accuracy, followed by DTC and then MLC. The authors attributed the SVM high classification accuracy to its ability to locate an optimal separating hyperplane. It was also stated that while the SVM performance was influenced by choice of parameter sets, the results of NN and DTC classification, too, were affected by the classifier configurations. For example, NN behavior is affected by the network's structure and random initializations and DTC is affected by the degree of pruning. Dixon and Candade (2008) performed an algorithmic comparison between a MLC, a backpropagation NN and an SVM-based classifier. A statistical assessment on a Landsat scene showed clear deficiencies for the MLC method, however results were similar for NN and SVM classifiers. The authors noted the training speed as an advantage for the SVM method, while admitting that the relatively low dimensionality did not allow them to fully explore their investigation.

Another study to compare SVMs and neural network classifiers using Landsat imagery was undertaken by Candade (2004). A major conclusion drawn was that a small number of training samples is sufficient to find the support vectors for near-optimal SVM learning. In the land use application domain using a Landsat scene, the study uncovered that SVM performed better than the backpropagation neural network not only in terms of classification accuracy but also when training times were compared. Three different SVM kernels (the polynomial, radial basis function and linear kernels) were analyzed for their performance. Overfitting and local minima were cited as the underlying cause of the relatively poor performance of neural networks on small training samples. Classification of Landsat 5 TM imagery was assessed from Tenerife (Canary Islands). This is an inherently difficult practical problem for ground truth data collection posed by the complex topographic relief. Keuchel et al. (2003) compared the classification accuracy of SVMs, maximum likelihood and iterated conditional modes (ICM). SVM and ICM methods outperformed maximum likelihood, however the authors suggest caution should be exercised in parameter (SVM) and iteration number (ICM) assignments. Kavzoglu and Colkesen (2009) contrasted SVMs with radial basis and polynomial functions and MLC using Landsat ETM+ and ASTER imagery. In both image types the superiority of SVMs was underlined.

Melgani and Bruzzone (2004) classified AVIRIS hyperspectral data using SVMs and compared the results with those using radial basis function neural networks and the K -nearest neighbor classifier. They found SVMs outperformed the other methods and concluded SVMs are a valid and effective alternative to conventional pattern recognition approaches to hyperspectral remote sensing data. In a study involving land cover update analysis conducted by Marcal et al. (2005) Advanced Space-borne Thermal Emission and Reflectance Radiometer (ASTER) imagery from Vale de Sousa region (northwest of Portugal) was used to compare the effectiveness of various classification methods including the SVMs, k -nearest neighbor (k -NN), logistic discrimination (LD) and training data-driven fuzzy classifiers. The SVM and LD classifiers produced higher overall accuracy than k -NN and the fuzzy classifiers. Kumar et al. (2007) in recognition of the fact that the proportion of mixed pixels in remote sensing images increases as spatial resolution decreases, proposed a method to deal with data fuzziness. The approach, called full fuzzy method, was tested on a land cover mapping problem in India using the LISS-III sensor. The full fuzzy scheme involved SVM-based sub-pixel analysis at all three different stages. Performance variation with different distance metrics (e.g., Mahalanobis and Euclidean norm) was investigated. SVMs with the Euclidean norm gave the highest accuracy, outperforming a corresponding variant of k -means clustering algorithm.

At coarser spatial resolutions a study was undertaken to evaluate the discriminatory power of two vegetation indices (the global vegetation index and terrestrial chlorophyll index) obtained from MERIS for general land cover mapping (Dash et al., 2007). Although a moderate level of accuracy was achieved using discriminant analysis method, a repetition of the experiment using an SVM technique revealed that the latter methodology resulted in a 6% gain in overall accuracy. Carrão et al. (2008) investigated the incorporation of multi-temporal MODIS data for a general 500 m LCLU classification with an SVM method as the underlying classifier.

5.3. Other tasks

Remote sensing data from the Himalayas (Nepal) were used to study soil erosion processes in tectonically active orogens (Andermann and Gloaguen, 2009). This research employed SVMs to provide a classification into land use, erosion and geomorphological processes. Although the maximum likelihood classifier yielded

higher classification accuracy, the authors point out that SVM results could be improved by selecting suitable values of the user-chosen SVM parameters. Zebedin et al. (2006) implemented SVM methods as part of a complex three dimensional reconstruction task using aerial, multi-spectral, high resolution data. Using the freely available NOAA/AVHRR satellite image data Gautam et al. (2008) tackled the problem of creating an automatic detector of coal field fire spots. The SVM method was used successfully to further refine detection results by removing points falsely highlighted by the threshold-based methodology in the regions deemed suspect. Rock glacier detection was undertaken using Landsat and SRTM terrain data (Brenning, 2009). Eleven different classifiers were tested and the SVM-based method did not rank highly when compared with the other methods.

SVMs have also been used for pure pixels (endmembers) identification. Brown et al. (2000) compared a linear SVM with linear spectral mixture models to identify pure pixels using Landsat TM data. They found that the SVM framework is more appropriate for nonlinear and/or empirical mixture modeling because SVMs can handle spectral confusion of pure pixels appropriately. Filippi and Archibald (2009) investigated SVMs to extract endmembers from hyperspectral data and pointed out that SVM-based endmember extraction has advantages in terms of efficiency and accuracy and is not sensitive to noise.

An integrative approach to information mining from large image datasets was proposed by Li and Narayanan (2004) based on SVMs. This framework was aimed at enabling users to make complex queries that would extend information search criteria beyond image metadata and actually access image content, a process called content based image retrieval. The proposed system architecture provided three components, namely, the image processing module, database module and graphical user interface. The backend image processing intelligence incorporated an SVM method to facilitate land cover mapping from a set of Landsat TM images in eastern Nebraska. An identified challenge was the integration methodology that would yield optimal results. Melgani (2006) proposed two methods to reconstruct cloud-contaminated remote sensing data using a sequence of multi-temporal multispectral images. The first method was based on the expectation-maximization algorithm to implement the contextual prediction process. The second method used a single non-linear predictor based on SVMs. Both methods outperformed the other methods based on compositing algorithms for cloud removal, and the first method was slightly better than the second method. Finally, Mazzoni et al. (2007) described one of the few SVM-based operational remote sensing classifiers using MISR imagery. There is also an interesting reference to an SVM-based classifier running onboard NASA's EO-1 spacecraft, as part of the Autonomous Sciencecraft Experiment (Mazzoni et al., 2005a,b) to automatically detect degraded images (e.g., from clouds) and avoid further processing and transmission on the satellite's platform.

SVMs have also been used in landmine detection. Potin et al. (2006) utilized ground-penetrating radar data (GPSARs) to detect buried landmines. They developed an abrupt change detection algorithm based on SVM, which was effective in reducing the clutter noise to improve the landmine detection. Jin and Zhou (2007) introduced a fuzzy hypersphere SVM (FHSSVM) based on the reduced features using the sequential forward floating selection method. They tested the FHSSVM for detecting landmines using rail GPSAR and found their method significantly improved the performance of landmine detection in different scenarios.

6. Discussion and concluding remarks

This review discussed important contributions of SVM-based works in remote sensing. In order to summarize efficiently the

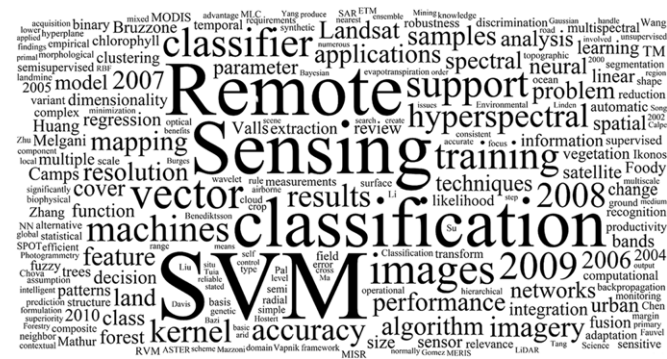


Fig. 4. Textual summary of this review.

content we resorted to a textual summary of frequently appearing single terms in this document. Fig. 4 displays such a visual representation, where higher frequency results in a larger font size. Looking beyond expected terms such as SVM, classification and remote sensing we also see the trends of recently published works (2008, 2009). In addition, Landsat is the prevailing sensor, while forest and land use applications show a significant presence. From the algorithmic perspective there is a significant discussion on the kernel and feature selection and their consequences to accuracy. Even though the focus is on classification tasks, there are worth-mentioning regression applications. Finally, the majority of comparative methods are neural networks, followed by maximum likelihood and decision trees.

Most of the findings show that there is empirical evidence to support the theoretical formulation and motivation behind SVMs. The most important characteristic is SVM's ability to generalize well from a limited amount and/or quality of training data. Compared to alternative methods such as backpropagation neural networks, SVMs can yield comparable accuracy using a much smaller training sample size. This is in line with the “support vector” concept that relies only on a few data points to define the classifier's hyperplane. This property has been exploited and has proved to be very useful in many of the applications we have seen thus far, mainly because acquisition of ground truth for remote sensing data is generally an expensive process.

SVMs offer additional benefits in contrast to alternative classification models, such as neural networks. They are resilient to getting trapped in local minima because of the convexity of the cost function which enables the classifier to consistently identify the optimal solution. In other words, SVM deals with quadratic problems hence it always gets to the global minimum. An added advantage is that there is no need for repeating classifier training using different random initializations or architectures. Furthermore, being non-parametric, SVMs do not assume a known statistical distribution of the data to be classified. This is particularly useful because the data acquired from remotely sensed imagery usually have unknown distributions. This allows SVMs to outperform techniques based on maximum likelihood classification because normality does not always give a correct assumption of the actual pixels distribution in each class (Su et al., 2009).

On the other hand, the majority of the studies uncovered common limitations to SVM methodologies, for example selection of SVM key parameters such as the kernel functions. To elaborate further, choosing a small value for the kernel width parameter (i.e. the kernel footprint in that multi-dimensional space) may lead to overfitting, while large kernel width values may lead to oversmoothing. This problem is not restricted to SVM methods, rather it is a general drawback of kernel-based approaches (e.g., radial basis function neural networks). Choice of the parameter value (usually denoted by C), which controls the trade-off between maximizing the margin and minimizing the training

error, is also an important consideration in SVM application. There exist no established heuristics for selection of these SVM parameters which frequently leads to a trial-and-error approach.

It has also been reported that the ‘one-against-the rest’ strategy for SVM multi-class classification can be problematic as it may result in unclassified instances of data and therefore lower classification accuracies (Pal and Mather, 2005). Moreover, SVM approaches frequently map input data to higher dimensional spaces in order to discern patterns. As dimensionality increases in addition to potential separability of patterns SVMs exhibit typical dimensionality issues such as outlier behavior and increased computational demands. This is a critical drawback especially for hyperspectral analysis where the dimensionality of the original data is high and kernel mapping is more vulnerable to dimensionality problems.

Moreover, SVMs are not optimized to deal with the inherent problem of noisy data; outlier effects are commonly encountered in remotely sensed data. Measurement errors due to limited precision of image acquisition instruments, and atmospheric and topographic distortions are some of the causes of such impurities. The quality of both training and test patterns are important in construction (training) and evaluation of automatic classification, recognition and detection systems. The performance of an SVM classifier can dramatically decrease with a relatively small number of mislabeled examples. Perhaps, more investigations into the potential of some of the relatively untapped lower level noise reduction techniques such as morphological image processing could provide a remedy to the problem of denoising. Citing one of the developments aimed at addressing this problem, Huang et al. (2008b) proposed a revised radiometric correction algorithm to counter the undesirable effects of atmospheric and topographic effect on data. Inglada (2007), supported by empirical evidence, similarly argued that higher number of geometric image features enhances multi-way characterization of objects that naturally have many different geometric properties. Also, pointed out by Dash et al. (2007), the choice of dataset source could help in remedying this hindrance by allowing the reduction in the size of the training set required.

There is significant room for extension of SVMs to address current pitfalls. For example, Foody (2008) assessed a relevance vector machine approach as a way to address the need to define the parameter C . RVMs are considered as a Bayesian treatment alternative to SVMs and have several advantages including probabilistic predictions, automatic estimation of parameters, and the arbitrary kernel functions. The authors argued that the new method leads to reduced sensitivity to the hyperparameter settings, thereby making the use of non-Mercer kernels possible. Furthermore, RVMs allow for fuzzy (or sub-pixel) classification of data making it possible to have a probabilistic output.

Typical SVM comparative assessment has not been as wide-reaching. Of particular interest would be comparison/fusion with algorithms such as self-organizing maps (Kohonen, 1997) that address efficiently high dimensionality problems and have already found fruitful ground in remote sensing (e.g., Hong et al., 2006; Gonçalves et al., 2008). In addition, integration with methodologies that deal more naturally with multi-class problems without the SVM complexity may further advance SVM understanding, for example a learning vector quantization system (Schneider et al., 2009).

In a nutshell, we can conclude that SVM classifiers, characterized by self-adaptability, swift learning pace and limited requirements on training size have proven a fairly reliable methodology in intelligent processing of data acquired through remote sensing. Past applications of the method on both real-world data and simulated environments have shown that SVMs exhibit superiority over most of the alternative algorithms — a big motivation and promise for future advances.

Acknowledgements

Support was provided by the National Science Foundation (award GRS-0648393), by the National Aeronautics and Space Administration (awards NNX08AR11G, NNX09AK16G) and by the Syracuse Center of Excellence CARTI Program.

References

- Andermann, C., Gloaguen, R., 2009. Estimation of erosion in tectonically active orogenies. Example from the Bhotekoshi catchment, Himalaya (Nepal). *International Journal of Remote Sensing* 30 (12), 3075–3096.
- Archibald, R., Fann, G., 2007. Feature selection and classification of hyperspectral images with support vector machines. *IEEE Geoscience and Remote Sensing Letters* 4 (4), 674–677.
- Barakat, N., Bradley, A.P., Rule extraction from support vector machines: a review. *Neurocomputing* (in press), doi:10.1016/j.neucom.2010.02.016.
- Bazi, Y., Melgani, F., 2006. Toward an optimal SVM classification system for hyperspectral remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing* 44 (11), 3374–3385.
- Bazi, Y., Melgani, F., 2007. Semisupervised PSO-SVM regression for biophysical parameter estimation. *IEEE Transactions on Geoscience and Remote Sensing* 45 (6), 1887–1895.
- Blanzieri, E., Melgani, F., 2008. Nearest neighbor classification of remote sensing images with the maximal margin principle. *IEEE Transactions on Geoscience and Remote Sensing* 46 (6), 1804–1811.
- Bovolo, F., Bruzzone, L., Marconcini, M., 2008. A novel approach to unsupervised change detection based on a semisupervised SVM and a similarity measure. *IEEE Transactions on Geoscience and Remote Sensing* 46 (7), 2070–2082.
- Boyd, D.S., Sanchez-Hernandez, C., Foody, G.M., 2006. Mapping a specific class for priority habitats monitoring from satellite sensor data. *International Journal of Remote Sensing* 27 (13), 2631–2644.
- Brenning, A., 2009. Benchmarking classifiers to optimally integrate terrain analysis and multispectral remote sensing in automatic rock glacier detection. *Remote Sensing of Environment* 113 (1), 239–247.
- Brown, M., Gunn, S.R., Lewis, H.G., 1999. Support vector machines for optimal classification and spectral unmixing. *Ecological Modelling* 120 (2–3), 167–179.
- Brown, M., Lewis, H.G., Gunn, S.R., 2000. Linear spectral mixture models and support vector machines for remote sensing. *IEEE Transactions on Geoscience and Remote Sensing* 38 (5), 2346–2360.
- Bruzzone, L., Chi, M., Marconcini, M., 2006. A novel transductive SVM for semisupervised classification of remote-sensing images. *IEEE Transactions on Geoscience and Remote Sensing* 44 (11), 3363–3373.
- Bruzzone, L., Melgani, F., 2005. Robust multiple estimator systems for the analysis of biophysical parameters from remotely sensed data. *IEEE Transactions on Geoscience and Remote Sensing* 43 (1), 159–174.
- Bruzzone, L., Persello, C., 2009. A novel context-sensitive semisupervised SVM classifier robust to mislabeled training samples. *IEEE Transactions on Geoscience and Remote Sensing* 47 (7), 2142–2154.
- Burges, C.J.C., 1998. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery* 2 (2), 121–167.
- Camps-Valls, G., Gomez-Chova, L., Calpe-Maravilla, J., Martin-Guerrero, J.D., Soria-Olivas, E., Alonso-Chorda, L., Moreno, J., 2004. Robust support vector method for hyperspectral data classification and knowledge discovery. *IEEE Transactions on Geoscience and Remote Sensing* 42 (7), 1530–1542.
- Camps-Valls, G., Gómez-Chova, L., Muñoz-Mari, J., Vila-Francés, J., Amorós-López, J., Calpe-Maravilla, J., 2006a. Retrieval of oceanic chlorophyll concentration with relevance vector machines. *Remote Sensing of Environment* 105 (1), 23–33.
- Camps-Valls, G., Bruzzone, L., Rojo-Alvarez, J.L., Melgani, F., 2006b. Robust support vector regression for biophysical variable estimation from remotely sensed images. *IEEE Geoscience and Remote Sensing Letters* 3 (3), 339–343.
- Camps-Valls, G., Gomez-Chova, L., Munoz-Mari, J., Vila-Francés, J., Calpe-Maravilla, J., 2006c. Composite kernels for hyperspectral image classification. *IEEE Geoscience and Remote Sensing Letters* 3 (1), 93–97.
- Camps-Valls, G., Gomez-Chova, L., Munoz-Mari, J., Rojo-Alvarez, J.L., Martinez-Ramon, M., 2008. Kernel-based framework for multitemporal and multisource remote sensing data classification and change detection. *IEEE Transactions on Geoscience and Remote Sensing* 46 (6), 1822–1835.
- Camps-Valls, G., Mooij, J., Scholkopf, B., 2010. Remote sensing feature selection by kernel dependence measures. *IEEE Geoscience and Remote Sensing Letters* 7 (3), 587–591.
- Candade, N., 2004. Multispectral classification of Landsat images: a comparison of support vector machine and neural network classifiers. *ASPRS Annual Conference Proceedings*, Denver, Colorado.
- Cao, X., Chen, J., Matsushita, B., Imura, H., Wang, L., 2009a. An automatic method for burn scar mapping using support vector machines. *International Journal of Remote Sensing* 30 (3), 577–594.
- Cao, X., Chen, J., Imura, H., Higashi, O., 2009b. A SVM-based method to extract urban areas from DMSP-OLS and SPOT VGT data. *Remote Sensing of Environment* 113 (10), 2205–2209.
- Carrão, H., Gonçalves, P., Caetano, M., 2008. Contribution of multispectral and multitemporal information from MODIS images to land cover classification. *Remote Sensing of Environment* 112 (3), 986–997.
- Castillo, C., Chollett, I., Klein, E., 2008. Enhanced duckweed detection using bootstrapped SVM classification on medium resolution RGB MODIS imagery. *International Journal of Remote Sensing* 29 (19), 5595–5604.
- Chen, H., Ho, P., 2008. Statistical pattern recognition in remote sensing. *Pattern Recognition* 41 (9), 2731–2741.
- Chen, J., Wang, C., Wang, R., 2008. Combining support vector machines with a pairwise decision tree. *IEEE Geoscience and Remote Sensing Letters* 5 (3), 409–413.
- Chen, J., Wang, C., Wang, R., 2009. Using stacked generalization to combine SVMs in magnitude and shape feature spaces for classification of hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing* 47 (7), 2193–2205.
- Chi, M., Feng, R., Bruzzone, L., 2008. Classification of hyperspectral remote-sensing data with primal SVM for small-sized training dataset problem. *Advances in Space Research* 41 (11), 1793–1799.
- Clevers, J.G.P.W., van der Heijden, G.W.A.M., Verzaakov, S., Schaepman, M.E., 2007. Estimating grassland biomass using SVM band shaving of hyperspectral. *Data Photogrammetric Engineering & Remote Sensing* 73 (10), 1141–1148.
- Cortes, C., Vapnik, V., 1995. Support-vector networks. *Machine Learning* 20 (3), 273–297.
- Dalponte, M., Bruzzone, L., Gianelle, D., 2008. Fusion of hyperspectral and LIDAR remote sensing data for classification of complex forest areas. *IEEE Transactions on Geoscience and Remote Sensing* 46 (5), 1416–1427.
- Dalponte, M., Bruzzone, L., Vesco, L., Gianelle, D., 2009. The role of spectral resolution and classifier complexity in the analysis of hyperspectral images of forest areas. *Remote Sensing of Environment* 113 (11), 2345–2355.
- Dash, J., Mathur, A., Foody, G.M., Curran, P.J., Chipman, J.W., Lillesand, T.M., 2007. Land cover classification using multi-temporal MERIS vegetation indices. *International Journal of Remote Sensing* 28 (6), 1137–1159.
- Demir, B., Ertürk, S., 2007. Hyperspectral image classification using relevance vector machines. *IEEE Geoscience and Remote Sensing Letters* 4 (4), 586–590.
- Demir, B., Ertürk, S., 2009. Clustering-based extraction of border training patterns for accurate SVM classification of hyperspectral images. *IEEE Geoscience and Remote Sensing Letters* 6 (4), 840–844.
- Dixon, B., Candade, N., 2008. Multispectral landuse classification using neural networks and support vector machines: one or the other, or both? *International Journal of Remote Sensing* 29 (4), 1185–1206.
- Durbha, S.S., King, R.L., Younan, N.H., 2007. Support vector machines regression for retrieval of leaf area index from multiangle imaging spectroradiometer. *Remote Sensing of Environment* 107 (1–2), 348–361.
- Esch, T., Himmler, V., Schorch, G., Thiel, M., Wehrmann, T., Bachofer, F., Conrad, C., Schmidt, M., Dech, S., 2009. Large-area assessment of impervious surface based on integrated analysis of single-date Landsat-7 images and geospatial vector data. *Remote Sensing of Environment* 113 (8), 1678–1690.
- Fauvel, M., Benediktsson, J.A., Chanussot, J., Sveinsson, J.R., 2008. Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles. *IEEE Transactions on Geoscience and Remote Sensing* 46 (11), 3804–3814.
- Fauvel, M., Chanussot, J., Benediktsson, J.A., 2009. Kernel principal component analysis for the classification of hyperspectral remote sensing data over urban areas. *EURASIP Journal on Advances in Signal Processing Article ID 783194*.
- Filippi, A.M., Archibald, R., 2009. Support vector machine-based endmember extraction. *IEEE Transactions on Geoscience and Remote Sensing* 47 (3), 771–791.
- Foody, G.M., Mathur, A., 2004a. A relative evaluation of multiclass image classification by support vector machines. *IEEE Transactions on Geoscience and Remote Sensing* 42 (6), 1335–1343.
- Foody, G.M., 2008. SVM-based multi-class classification of remotely sensed data. *International Journal of Remote Sensing* 29 (6), 1817–1823.
- Foody, G.M., Mathur, A., 2004b. Toward intelligent training of supervised image classifications: directing training data acquisition for SVM classification. *Remote Sensing of Environment* 93 (1–2), 107–117.
- Foody, G.M., Mathur, A., 2006. The use of small training sets containing mixed pixels for accurate hard image classification: training on mixed spectral responses for classification by a SVM. *Remote Sensing of Environment* 103 (2), 179–189.
- Foody, G.M., Mathur, A., Sanchez-Hernandez, C., Boyd, D.S., 2006. Training set size requirements for the classification of a specific class. *Remote Sensing of Environment* 104 (1), 1–14.
- Gautam, R.S., Singh, D., Mittal, A., Sajin, P., 2008. Application of SVM on satellite images to detect hotspots in Jharia coal field region of India. *Advances in Space Research* 41 (11), 1784–1792.
- Geman, S., Bienenstock, E., Doursat, R., 1992. Neural networks and the bias/variance dilemma. *Neural Computation* 4 (1), 1–58.
- Ghoggali, N., Melgani, F., 2008. Genetic SVM approach to semisupervised multitemporal classification. *IEEE Geoscience and Remote Sensing Letters* 5 (2), 212–216.
- Ghoggali, N., Melgani, F., Bazi, Y., 2009. A multiobjective genetic SVM approach for classification problems with limited training samples. *IEEE Transactions on Geoscience and Remote Sensing* 47 (6), 1707–1718.
- Gomez-Chova, L., Camps-Valls, G., Bruzzone, L., Calpe-Maravilla, J., 2010. Mean map kernel methods for semisupervised cloud classification. *IEEE Transactions on Geoscience and Remote Sensing* 48 (1), 207–220.
- Gómez-Chova, L., Camps-Valls, G., Muñoz-Mari, J., Calpe, J., 2008. Semisupervised image classification with Laplacian support vector machines. *IEEE Geoscience and Remote Sensing Letters* 5 (3), 336–340.
- Goncalves, M.L., Netto, M.L.A., Costa, J.A.F., Zullo, J.U., Nior, J., 2008. An unsupervised method of classifying remotely sensed images using Kohonen self-organizing maps and agglomerative hierarchical clustering methods. *International Journal of Remote Sensing* 29 (11), 3171–3207.

- Griffiths, P., Hostert, P., Gruebner, O., Linden, S., 2010. Mapping megacity growth with multi-sensor data. *Remote Sensing of Environment* 114 (2), 426–439.
- Gualtieri, J.A., Crompton, R.F., 1998. Support vector machines for hyperspectral remote sensing classification. In: *Proceedings of the 27th AIPR Workshop: Advances in Computer Assisted Recognition*, Washington, DC, 27 October. SPIE, Washington, DC, pp. 221–232.
- Guyon, I., Vapnik, V., Boser, B., Solla, S.A., 1992. Capacity control in linear classifiers for pattern recognition. In: *First IAPR International Conference on Pattern Recognition*. IEEE Computer Society Press, pp. 385–388.
- Heikkinen, V., Tokola, T., Parkkinen, J., Korpela, I., Jaaskelainen, T., 2010. Simulated multispectral imagery for tree species classification using support vector machines. *IEEE Transactions on Geoscience and Remote Sensing* 48 (3), 1355–1364.
- Hong, Y., Chiang, Y.-M., Liu, Y., Hsu, K.-L., Sorooshian, S., 2006. Satellite-based precipitation estimation using watershed segmentation and growing hierarchical self-organizing map. *International Journal of Remote Sensing* 27 (23), 5165–5184.
- Huang, C., Davis, L.S., Townshend, J.R.G., 2002. An assessment of support vector machines for land cover classification. *International Journal of Remote Sensing* 23 (4), 725–749.
- Huang, C., Song, K., Kim, S., Townshend, J.R.G., Davis, P., Masek, J.G., Goward, S.N., 2008a. Use of a dark object concept and support vector machines to automate forest cover change analysis. *Remote Sensing of Environment* 112 (3), 970–985.
- Huang, H., Gong, P., Clinton, N., Hui, F., 2008b. Reduction of atmospheric and topographic effect on Landsat TM data for forest classification. *International Journal of Remote Sensing* 29 (19), 5623–5642.
- Huang, X., Zhang, L., Li, P., 2008c. A multiscale feature fusion approach for classification of very high resolution satellite imagery based on wavelet transform. *International Journal of Remote Sensing* 29 (20), 5923–5941.
- Huang, X., Zhang, L., 2010. Comparison of vector stacking, multi-SVMs fuzzy output, and multi-SVMs voting methods for multiscale VHR urban mapping. *IEEE Geoscience and Remote Sensing Letters* 7 (2), 261–265.
- Huang, X., Zhang, L., 2009. Road centreline extraction from high-resolution imagery based on multiscale structural features and support vector machines. *International Journal of Remote Sensing* 30 (8), 1977–1987.
- Inglada, J., 2007. Automatic recognition of man-made objects in high resolution optical remote sensing images by SVM classification of geometric image features. *ISPRS Journal of Photogrammetry and Remote Sensing* 62 (3), 236–248.
- Jin, T., Zhou, Z., 2007. Ultrawideband synthetic aperture radar landmine detection. *IEEE Transactions on Geoscience and Remote Sensing* 45 (11), 3561–3573.
- Kaheil, Y.H., Rosero, E., Gill, M.K., McKee, M., Bastidas, L.A., 2008. Downscaling and forecasting of evapotranspiration using a synthetic model of wavelets and support vector machines. *IEEE Transactions on Geoscience and Remote Sensing* 46 (9), 2692–2707.
- Karimi, Y., Prasher, S.O., Patel, R.M., KIMB, S.H., 2006. Application of support vector machine technology for weed and nitrogen stress detection in corn. *Computers and Electronics in Agriculture* 51 (1–2), 99–109.
- Kavzoglu, T., Colkesen, I., 2009. A kernel functions analysis for support vector machines for land cover classification. *International Journal of Applied Earth Observation and Geoinformation* 11 (5), 352–359.
- Keramitsoglou, I., Sarimveis, H., Kiranoudis, C.T., Kontoes, C., Sifakis, N., Fitoka, E., 2006. The performance of pixel window algorithms in the classification of habitats using VHSR imagery. *ISPRS Journal of Photogrammetry and Remote Sensing* 60 (4), 225–238.
- Keuchel, J., Naumann, S., Heiler, M., Siegmund, A., 2003. Automatic land cover analysis for Tenerife by supervised classification using remotely sensed data. *Remote Sensing of Environment* 86 (4), 530–541.
- Knerr, S., Personnaz, L., Dreyfus, G., 1990. Single-layer learning revisited: a stepwise procedure for building and training a neural network. In: *NATO ASI Series, Springer, Algorithms, Architectures and Applications*.
- Knorn, J., Rabe, A., Radeloff, V.C., Kuemmerle, T., Kozak, J., Hostert, P., 2009. Land cover mapping of large areas using chain classification of neighboring Landsat satellite images. *Remote Sensing of Environment* 113 (5), 957–964.
- Knudby, A., LeDrew, E., Brenning, A., 2010. Predictive mapping of reef fish species richness, diversity and biomass in Zanzibar using IKONOS imagery and machine-learning techniques. *Remote Sensing of Environment* 114 (6), 1230–1241.
- Kohonen, T., 1997. *Self Organizing Maps*, 2nd ed. Springer-Verlag, Berlin.
- Kuemmerle, T., Chaskovskyy, T.K.O., Knorn, J., Radeloff, V.C., Kruhlov, I., Keeton, W.S., Hostert, P., 2009. Forest cover change and illegal logging in the Ukrainian Carpathians in the transition period from 1988 to 2007. *Remote Sensing of Environment* 113 (6), 1194–1207.
- Kumar, A., Ghosh, S.K., Dadhwal, V.K., 2007. Full fuzzy land cover mapping using remote sensing data based on fuzzy *k*-means and density estimation. *Canadian Journal of Remote Sensing* 33 (2), 81–87.
- Kwiatkowska, E.J., Fargion, G.S., 2003. Application of machine-learning techniques toward the creation of a consistent and calibrated global chlorophyll concentration baseline dataset using remotely sensed ocean color data. *IEEE Transactions on Geoscience and Remote Sensing* 41 (12), 2844–2860.
- Lardeux, C., Frison, P.L., Tison, C., Souyris, J.C., Stoll, B., Fruneau, B., Rudant, J.P., 2009. Support vector machine for multifrequency SAR polarimetric data classification. *IEEE Transactions on Geoscience and Remote Sensing* 47 (12), 4143–4152.
- Li, H., Gu, H., Han, Y., Yang, J., 2010. Object-oriented classification of high-resolution remote sensing imagery based on an improved colour structure code and a support vector machine. *International Journal of Remote Sensing* 31 (6), 1453–1470.
- Li, J., Narayanan, R.M., 2004. Integrated spectral and spatial information mining in remote sensing imagery. *IEEE Transactions on Geoscience and Remote Sensing* 42 (3), 673–685.
- Licciardi, G., Pacifici, F., Tuia, D., Prasad, S., West, T., Giacco, F., Thiel, C., Inglada, J., Christophe, E., Chanussot, J., Gamba, P., 2009. Decision fusion for the classification of hyperspectral data: outcome of the 2008 GRS-S data fusion contest. *IEEE Transactions on Geoscience and Remote Sensing* 47 (11), 3857–3865.
- Linden, S., Hostert, P., 2009. The influence of urban structures on impervious surface maps from airborne hyperspectral data. *Remote Sensing of Environment* 113 (11), 2298–2305.
- Liu, D., Kelly, M., Gong, P., A., 2006. Spatial-temporal approach to monitoring forest disease spread using multi-temporal high spatial resolution imagery. *Remote Sensing of Environment* 101 (2), 167–180.
- Luo, J., Ming, D., Liu, W., Shen, Z., Wang, M., Sheng, H., 2007. Extraction of bridges over water from IKONOS panchromatic data. *International Journal of Remote Sensing* 28 (16), 3633–3648.
- Mantero, P., Moser, G., Serpico, S.B., 2005. Partially supervised classification of remote sensing images through SVM-based probability density estimation. *IEEE Transactions on Geoscience and Remote Sensing* 43 (3), 559–570.
- Marcal, A.R.S., Borges, J.S., Gomes, J.A., Pinto Da Costa, J.F., 2005. Land cover update by supervised classification of segmented ASTER images. *International Journal of Remote Sensing* 26 (7), 1347–1362.
- Marconcini, M., Camps-Valls, G., Bruzzone, L., 2009. A composite semisupervised SVM for classification of hyperspectral images. *IEEE Geoscience and Remote Sensing Letters* 6 (2), 234–238.
- Mathur, A., Foody, G.M., 2008a. Multiclass and binary SVM classification: implications for training and classification users. *IEEE Geoscience and Remote Sensing Letters* 5 (2), 241–245.
- Mathur, A., Foody, G.M., 2008b. Crop classification by support vector machine with intelligently selected training data for an operational application. *International Journal of Remote Sensing* 29 (8), 2227–2240.
- Mazzoni, D., Tang, N., Doggett, T., Chien, S., Greeley, R., Cichy, B., 2005a. Learning classifiers for science event detection in remote sensing imagery. In: *Proceedings of the 8th International Symposium on Artificial Intelligence, Robotics and Automation in Space (i-SAIRAS 2005)*.
- Mazzoni, D.M., Horváth, A., Garay, M.J., Tang, B., Davies, R., 2005b. A MISR cloud-type classifier using reduced support vector machines. In: *Proceedings of the Eighth Workshop on Mining Scientific and Engineering Datasets*, 2005 SIAM International Conference on Data Mining.
- Mazzoni, D., Garay, M.J., Davies, R., Nelson, D., 2007. An operational MISR pixel classifier using support vector machines. *Remote Sensing of Environment* 107 (1–2), 149–158.
- Melgani, F., 2006. Contextual reconstruction of cloud-contaminated multitemporal multispectral images. *IEEE Transactions on Geoscience and Remote Sensing* 44 (2), 442–455.
- Melgani, F., Bruzzone, L., 2004. Classification of hyperspectral remote sensing images with support vector machines. *IEEE Transactions on Geoscience and Remote Sensing* 42 (8), 1778–1790.
- Mitra, P., Shankar, B.U., Pal, S., 2004. Segmentation of multispectral remote sensing images using active support vector machines. *Pattern Recognition Letters* 25 (9), 1067–1074.
- Mladinich, C.S., 2010. An evaluation of object-oriented image analysis techniques to identify motorized vehicle effects in semi-arid to arid ecosystems of the american west. *GIScience & Remote Sensing* 47 (1), 53–77.
- Montgomery, D.C., Peck, E.A., 1992. *Introduction to Linear Regression Analysis*, 2nd ed. Wiley, New York.
- Moser, G., Serpico, S.B., 2009. Automatic parameter optimization for support vector regression for land and sea surface temperature estimation from remote sensing data. *IEEE Transactions on Geoscience and Remote Sensing* 47 (3), 909–921.
- Mukhopadhyay, A., Maulik, U., 2009. Unsupervised pixel classification in satellite imagery using multiobjective fuzzy clustering combined with SVM classifier. *IEEE Transactions on Geoscience and Remote Sensing* 47 (4), 1132–1138.
- Muñoz-Mari, J., Bruzzone, L., Camps-Valls, G., 2007. A support vector domain description approach to supervised classification of remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing* 45 (8), 2683–2692.
- Nemmour, H., Chibani, Y., 2006. Multiple support vector machines for land cover change detection: an application for mapping urban extensions. *ISPRS Journal of Photogrammetry and Remote Sensing* 61 (2), 125–133.
- Pal, M., 2006. Support vector machine-based feature selection for land cover classification: a case study with DAIS hyperspectral data. *International Journal of Remote Sensing* 27 (14), 2877–2894.
- Pal, M., 2008. Ensemble of support vector machines for land cover classification. *International Journal of Remote Sensing* 29 (10), 3043–3049.
- Pal, M., Mather, P.M., 2005. Support vector machines for classification in remote sensing. *International Journal of Remote Sensing* 26 (5), 1007–1011.
- Plaza, A., Benediktsson, J.A., Boardman, J.W., Brazile, J., Bruzzone, L., Camps-valls, G., Chanussot, J., Fauvel, M., Gamba, P., Gualtieri, A., Marconcini, M., Tilton, J.C., Trianni, G., 2009. Recent advances in techniques for hyperspectral image processing. *Remote Sensing of Environment* 113 (1), S110–S122.
- Potin, D., Vanheege, P., Duflos, E., Davy, M., 2006. An abrupt change detection algorithm for buried landmines localization. *IEEE Transactions on Geoscience and Remote Sensing* 44 (2), 260–272.
- Sahoo, B.C., Oommen, T., Misra, D., Newby, G., 2007. Using the one-dimensional *s*-transform as a discrimination tool in classification of hyperspectral images. *Canadian Journal of Remote Sensing* 33 (6), 551–560.
- Schneider, P., Biehl, M., Hammer, B., 2009. Adaptive relevance matrices in learning vector quantization. *Neural Computation* 21 (12), 3532–3561.

- Scholkopf, B., Smola, A.J., 2001. *Learning with Kernels*. The MIT Press.
- Shi, W., Zheng, S., Tian, Y., 2009. Adaptive mapped least squares SVM-based smooth fitting method for DSM generation of LIDAR data. *International Journal of Remote Sensing* 30 (21), 5669–5683.
- Smola, A.J., Schölkopf, B., 2004. A tutorial on support vector regression. *Statistics and Computing* 14 (3), 199–222.
- Song, M., Civco, D., 2004. Road extraction using SVM and image segmentation. *Photogrammetric Engineering & Remote Sensing* 70 (12), 1365–1371.
- Song, X., Cherian, G., Fan, G., 2005. A ν -insensitive SVM approach for compliance monitoring of the conservation reserve program. *IEEE Geoscience and Remote Sensing Letters* 2 (2), 99–103.
- Su, L., 2009. Optimizing support vector machine learning for semi-arid vegetation mapping by using clustering-based analysis. *ISPRS Journal of Photogrammetry and Remote Sensing* 64 (4), 407–413.
- Su, L., Huang, Y., 2009. Support vector machine (SVM) classification: comparison of linkage techniques using a clustering-based method for training data selection. *GIScience & Remote Sensing* 46 (4), 411–423.
- Su, L., Huang, Y., Chopping, M.J., Rango, A., Martonchik, J.V., 2009. An empirical study on the utility of BRDF model parameters and topographic parameters for mapping vegetation in a semi-arid region with MISR imagery. *International Journal of Remote Sensing* 30 (13), 3463–3483.
- Sun, D., Li, Y., Wang, Q., 2009. A unified model for remotely estimating chlorophyll a in Lake Taihu, China, based on SVM and in situ hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing* 47 (8), 2957–2965.
- Tang, S., Chen, C., Zhan, H., Zhang, T., 2008. Determination of ocean primary productivity using support vector machines. *International Journal of Remote Sensing* 29 (21), 6227–6236.
- Tipping, M.E., 2000. The relevance vector machine. In: Solla, S.A., Leen, T.K., Muller, K.R. (Eds.), *Advances in Neural Information Processing Systems*, vol. 12. MIT Press, Cambridge, MA.
- Tipping, M.E., 2001. Sparse Bayesian learning and the relevance vector machine. *Journal of Machine Learning Research* 1, 211–244.
- Tan, C.P., Koay, J.Y., Lim, K.S., Ewe, H.T., Chuah, H.T., 2007. Classification of multi-temporal sar images for rice crops using combined entropy decomposition and support vector machine technique. *Progress in Electromagnetics Research* 71, 19–39.
- Tarabalka, Y., Benediktsson, J.A., Chanussot, J., 2009. Spectral-spatial classification of hyperspectral imagery based on partitioned clustering techniques. *IEEE Transactions on Geoscience and Remote Sensing* 47 (8), 2973–2987.
- Tso, B., Mather, P., 2009. *Classification Methods for Remotely Sensed Data*, 2nd ed. CRC Press, 376 p.
- Tuia, D., Camps-Valls, G., 2009. Semisupervised remote sensing image classification with cluster kernels. *IEEE Geoscience and Remote Sensing Letters* 6 (2), 224–228.
- Tuia, D., Pacifici, F., Kanevski, M., Emery, W.J., 2009. Classification of very high spatial resolution imagery using mathematical morphology and support vector machines. *IEEE Transactions on Geoscience and Remote Sensing* 47 (11), 3866–3879.
- Vapnik, V., 1979. *Estimation of Dependences Based on Empirical Data*. Nauka, Moscow, pp. 5165–5184, 27 (in Russian) (English translation: Springer Verlag, New York, 1982).
- Walton, J.T., 2008. Subpixel urban land cover estimation: comparing cubist, random forests, and support vector regression. *Photogrammetric Engineering & Remote Sensing* 74 (10), 1213–1222.
- Wang, L., Jia, X., 2009. Integration of soft and hard classifications using extended support vector machines. *IEEE Geoscience and Remote Sensing Letters* 6 (3), 543–547.
- Warner, T.A., Nerry, F., 2009. Does single broadband or multispectral thermal data add information for classification of visible, near- and shortwave infrared imagery of urban areas? *International Journal of Remote Sensing* 30 (9), 2155–2171.
- Waske, B., Benediktsson, J.A., 2007. Fusion of support vector machines for classification of multisensor data. *IEEE Transactions on Geoscience and Remote Sensing* 45 (12), 3858–3866.
- Waske, B., van der Linden, S., 2008. Classifying multilevel imagery from SAR and optical sensors by decision fusion. *IEEE Transactions on Geoscience and Remote Sensing* 46 (5), 1457–1466.
- Watanachaturaporn, P., Arora, M.K., Varshney, P.K., 2008. Multisource classification using support vector machines: an empirical comparison with decision tree and neural network classifiers. *Photogrammetric Engineering & Remote Sensing* 74 (2), 239–246.
- Wilson, M.D., Ustin, S.L., Rocke, D.M., 2004. Classification of contamination in salt marsh plants using hyperspectral reflectance. *IEEE Transactions on Geoscience and Remote Sensing* 42 (5), 1088–1095.
- Xie, X., Liu, T., Tang, B., 2008. Spacebased estimation of moisture transport in marine atmosphere using support vector regression. *Remote Sensing of Environment* 112 (4), 1846–1855.
- Yang, F., Ichii, K., White, M.A., Hashimoto, H., Michaelis, A.R., Votava, P., Zhu, A., Huete, A., Running, S.W., Nemani, R.R., 2007. Developing a continental-scale measure of gross primary production by combining MODIS and AmeriFlux data through support vector machine approach. *Remote Sensing of Environment* 110 (1), 109–122.
- Yang, F., White, M.A., Michaelis, A.R., Ichii, K., Hashimoto, H., Votava, P., Zhu, A., Nemani, R.R., 2006. Prediction of continental-scale evapotranspiration by combining MODIS and AmeriFlux data through support vector machine. *IEEE Transactions on Geoscience and Remote Sensing* 44 (11), 3452–3461.
- Zebedin, L., Klaus, A., Gruber-Geymayer, B., Karner, K., 2006. Towards 3D map generation from digital aerial images. *ISPRS Journal of Photogrammetry and Remote Sensing* 60 (6), 413–427.
- Zhang, L., Huang, X., Huang, B., Li, P., 2006. A pixel shape index coupled with spectral information for classification of high spatial resolution remotely sensed imagery. *IEEE Transactions on Geoscience and Remote Sensing* 44 (10), 2950–2961.
- Zhang, R., Ma, J., 2008. An improved SVM method P-SVM for classification of remotely sensed data. *International Journal of Remote Sensing* 29 (20), 6029–6036.
- Zhang, R., Ma, J., 2009. Feature selection for hyperspectral data based on recursive support vector machines. *International Journal of Remote Sensing* 30 (14), 3669–3677.
- Zheng, S., Shi, W., Liu, J., Tian, J., 2008. Remote sensing image fusion using multiscale mapped LS-SVM. *IEEE Transactions on Geoscience and Remote Sensing* 46 (5), 1313–1322.
- Zhu, G., Blumberg, D.G., 2002. Classification using ASTER data and SVM algorithms: The case study of Beer Sheva, Israel. *Remote Sensing of Environment* 80 (2), 233–240.