

Chapter 13

Support Vector Machines for Land Cover Mapping from Remote Sensor Imagery

Dee Shi and Xiaojun Yang

Abstract Land cover mapping is an important activity leading to the generation of various thematic products essential for numerous environmental monitoring and resources management applications at local, regional, and global levels. Over the years, various pattern recognition techniques have been developed to automate this process from remote sensor imagery. Support vector machines (SVM) as a group of relatively novel statistical learning algorithms have demonstrated their robustness in classifying homogeneous and heterogeneous land cover types. In this chapter, we review the status and potential challenges in the SVM implementation for land cover classification. The chapter is organized into two major parts. The first part reviews the research status of using SVM for land cover classification, focusing on some comparative studies that demonstrated the algorithm effectiveness over other conventional classifiers. We identify several areas for additional work, which are mostly related to appropriate treatments of some parametric and non-parametric factors in order to achieve improved mapping accuracies particularly for working over heterogeneous landscapes. Then, we implement the support vector machine technique to map various land cover types from a satellite image covering an urban area, and demonstrate the robustness of this pattern recognition technique for mapping heterogeneous landscapes.

Keywords Land cover • Image classification • Support vector machines • Heterogeneous landscapes • Thematic accuracy assessment

13.1 Introduction

Land cover is the pattern of ecological resources and human activities dominating different areas of Earth's surface (Turner and Meyer 1994). It is a critical type of data source essential for many environmental monitoring and natural resources management applications at local, regional, and global scales (Foley et al. 2005;

D. Shi (✉) • X. Yang

Department of Geography, Florida State University, Tallahassee, FL 32306, USA

e-mail: ds10f@my.fsu.edu; xyang@fsu.edu

© Springer Science+Business Media Dordrecht 2015

J. Li, X. Yang (eds.), *Monitoring and Modeling of Global Changes: A Geomatics Perspective*, Springer Remote Sensing/Photogrammetry, DOI 10.1007/978-94-017-9813-6_13

265

Alberti 2008). Land cover patterns are observable and therefore can be mapped by ground surveys or remote sensing. While ground surveys are largely limited by logistical constraints, remote sensing makes direct observations across large areas of the land surface, thus allowing land cover patterns to be mapped in a timely and cost-effective mode. Both visual interpretation and computer-based digital classification can be used to extract information on land cover from a variety of remotely sensed data varying in spatial, spectral, radiometric, and temporal resolutions. Digital pattern classification is generally preferred over visual interpretation for mapping land cover in large areas (Jensen 2005).

While conventional pattern classifiers (e.g., maximum likelihood) have been widely used, they generally work well with medium-resolution images and in relatively homogeneous areas rather than highly heterogeneous areas (Yang 2002). Over the years, substantial research efforts have been directed to improve the performance of land cover mapping in heterogeneous areas (e.g. Hoffer 1978; Richards et al. 1982; Skidmore et al. 1997; Duda et al. 2001; Yang and Lo 2002; Schmidt et al. 2004; Del Frate et al. 2007; Foody 2008; Heikkinen et al. 2010; Zhou and Yang 2011; Liu and Yang 2013).

This study targets support vector machines (SVM), a group of relatively novel machine learning algorithms based on statistical learning theory that have not been extensively exploited in the remote sensing community. They are found to outperform most of the conventional classifiers (Huang et al. 2002; Keuchel et al. 2003; Kavzoglu and Colkesen 2009; Su and Huang 2009). Moreover, SVM were found to even outperform some novel pattern recognition methods, such as neural networks (Huang et al. 2002; Foody and Mathur 2004a, b). Nevertheless, there are some parametric and non-parametric factors that can affect the performance of SVM, and there is a need to investigate them so that SVM could be used with improved performance (Yang 2011).

In this chapter, we examine the utilities of support vector machines (SVM) as a pattern recognition technique for landscape mapping particular for heterogeneous areas. It is organized into two major parts, beginning with a brief introduction of some basic knowledge on SVM and a review on the research status and possible challenges of using SVM for land cover mapping. The review focuses on some comparative studies that demonstrated the effectiveness of SVM over other conventional classifiers. Based on the review, we further discuss several areas that need additional research in order to improve SVM classification accuracies and reduce computational burdens, which are mostly related to appropriate treatments of some parametric and non-parametric factors. The second part of the paper discusses our implementation of SVM to map various land cover types from a remote sensor image covering an urban area, demonstrating the robustness of this type of pattern recognition technique for mapping heterogeneous landscapes.

13.2 Support Vector Machines

13.2.1 Basics

The basic idea behind the support vector machines (SVM) is to construct separating hyperplanes between classes in feature space through the use of support vectors which are lying at the edges of class domains; SVM seek the optimal hyperplane that can separate classes from each other with the maximum margin (Vapnik 1995).

SVM were originally designed as a binary linear classifier, which assumes two linearly separable classes to be partitioned. In most cases, the best separable hyperplane may not be located exactly between two classes. To account for this, an error item is introduced to manipulate the tradeoff between maximizing the separation margin and minimizing the count of training samples that locates on the wrong side. SVM are further extended to deal with non-linear classification by using a non-linear kernel function to replace the inner product of optimal hyperplane. Several commonly used kernel functions include linear kernel, polynomial kernel, radial basis function (RBF), and sigmoid kernel (Haykin 1999). Each of these kernel functions is constructed with multiple parameters, and the parameter settings can influence the performance of a specific support vector machine (Yang 2011).

Moreover, SVM have been used for multi-class mapping through reducing the multi-class problem into a set of binary problems so that the basic SVM principles can be still applied. Two commonly used strategies for this purpose include one-against-one and one-against-all (Foody and Mathur 2004b; Kavzoglu and Colkesen 2009). The former is generally preferred because of its less computational intensity and comparable accuracy to the later. The one-against-all method can result in unclassified instances (Huang et al. 2002; Hsu and Lin 2002; Pal and Mather 2005; Mountrakis et al. 2011), which is not suitable for land cover mapping.

13.2.2 SVM for Land Cover Classification

The performance of SVM has been examined through some comparative studies with other pattern classifiers for various land cover types (e.g., Huang et al. 2002; Foody and Mathur 2006; Keramitsoglou et al. 2006; Su and Huang 2009). Huang et al. (2002) found that SVM substantially outperformed maximum likelihood (MLC) or decision tree (DC) in terms of classification accuracy and even surpassed multilayer perceptron neural networks (MLP). Su and Huang (2009) implemented SVM and MLC on a Multi-angle Imaging SpectroRadiometer (MISR) image to differentiate eight semi-arid vegetation types, and found that SVM significantly outperformed MLC. Keramitsoglou et al. (2006) mapped various vegetation types using IKONOS data, and compared the performance of SVM with radial basis (RBF) neural networks. They found that SVM had strengths in terms of

classification accuracy and training time. Foody and Mathur (2006) also found that SVM can produce a more accurate classification of cultivated landscape types. Dixon and Candade (2008) compared SVM, MLC, and backpropagation neural networks (NN) for classifying a Landsat scene, and found that SVM and NN performed identically in the classification accuracy but SVM was more efficient in the training phase. They also noted that SVM can be quite attractive when working with high-dimensional data. This seems to be in line with an earlier work conducted by Huang et al. (2002) who found that SVM performed better for an image with seven bands than with three bands. The effectiveness of SVM for working with high-dimensional data classification was also confirmed by several other studies (e.g., Bazi and Melgani 2006; Camps-Valls et al. 2007), indicating that they could provide a solution to dealing with the problem of “curse-of-dimensionality” (Hughes 1968). Although SVM have demonstrated strengths when comparing with other classifiers, their performance can vary across different land cover types (Foody and Mathur 2004a, b; Keramitsoglou et al. 2006; Su and Huang 2009).

The performance of SVM can be affected by both parametric and non-parametric factors (Foody and Mathur 2006; Yang 2011). Existing studies on SVM classification have largely concentrated on either improving classification accuracy on specific land cover types or reducing computational burdens, both of which can be manipulated at the SVM configuration stage and at the training stage. The inner-product kernel between the support vectors in feature space and in input space largely determines the separability of optimal separable hyperplane (Haykin 1999). While introducing non-linear kernel functions could help deal with complex, non-linear classification, it can also lead to the difficulty in choosing the most appropriate kernel type and in the subsequent kernel parameterization (Huang et al. 2002; Kavzoglu and Colkesen 2009; Yang 2011). Yang (2011) conducted an empirical study assessing the performance of several most commonly used kernel types, along with their internal parameterization, and found that the kernel type and error penalty can substantially affect image classification accuracy. Some customized kernels, particularly those incorporating both spatial and spectral information, were found to be quite promising when comparing with spectral-based kernel types (Camps-Valls et al. 2006, 2007; Plaza et al. 2009).

Since the SVM is a supervised classifier by nature, both the size and quality of training sample can affect the classification accuracy (Foody and Mathur 2006). For land cover mapping from remote sensor imagery, training samples should consist of relatively pure pixels, and should be identified from homogeneous areas in large fields, which can be applicable for a variety of classifiers (Foody and Arora 1997). SVM performance can be sensitive to the noise in training samples due to the use of support vectors at the edges of class domains in feature space (Rodriguez-Galiano et al. 2012). A minimum of 10–30 pixels per class per waveband should be used to meet the assumption of normal distribution and be representative of the subclass (Foody and Mathur 2004a, b, 2006). Like other non-parametric classifiers, there is no need to maintain normal distributions in training samples for a SVM classification. Since only the support vectors are actually needed in constructing separate

hyperplanes for SVM, it may be highly possible to reduce training sample size to a small number of the most informative samples that are used to fit the decision hyperplanes. Several studies have been conducted to identify these critical samples. For example, Foody and Marthur (2004a, b, 2006) incorporated ancillary information of soil types and geographical boundary pixels of mixed spectral characteristics of two crop types in the selection of useful training samples, which dramatically reduced training samples before being applied to classification. They also examined the usefulness of applying other ancillary information (e.g., landform, moisture, and spatial texture) in targeting support vectors. Various techniques have been identified to automatically reduce the training sample size and hence help reduce the computational burden for SVM. For example, clustering-based algorithms are applied in training pattern selection to remove samples locating at the high density regions or to detect support vectors at the clustering centers (Demir and Ertürk 2009; Su 2009). With these support vectors obtained from clustering preprocessing, the computational load has been substantially reduced, while the classification accuracy was much higher than using the full training samples.

13.3 Implementation of SVM for Land Cover Mapping

In order to demonstrate the effectiveness of SVM for heterogeneous land cover mapping, we implemented SVM to map land cover types in an urban area. In this section, we will discuss the specific procedures, including the study site and data acquisition, classification scheme design, SVM configuration, and classification and accuracy assessment (Fig. 13.1).

13.3.1 Study Site and Data Acquisition

The study site covers the entire Gwinnett County, a suburban county located at northeastern Atlanta metropolitan area, Georgia, USA (Fig. 13.2). The county has an area of about 1,122 km² and its population was 805,321 according to the 2010 census survey. The majority of topography is relatively flat and has primarily a humid subtropical climate. Gwinnett has been one of America's fastest-growing counties and the second most populated county in Georgia. Its landscape is characterized by a mosaic of complex land use and land cover types, and therefore Gwinnett is an ideal site to examine the effectiveness of SVM for heterogeneous landscape mapping.

A cloud-free Landsat-5 Thematic Mapper (TM) image dated on 19 May 2007 was acquired from USGS EROS Data Center, and a subset of this scene covering the entire Gwinnett County was actually used in our study (Fig. 13.3). The image has been geometrically corrected at the EROS data center, and no further preprocessing was conducted. The spatial resolution of this image is 30 m for all

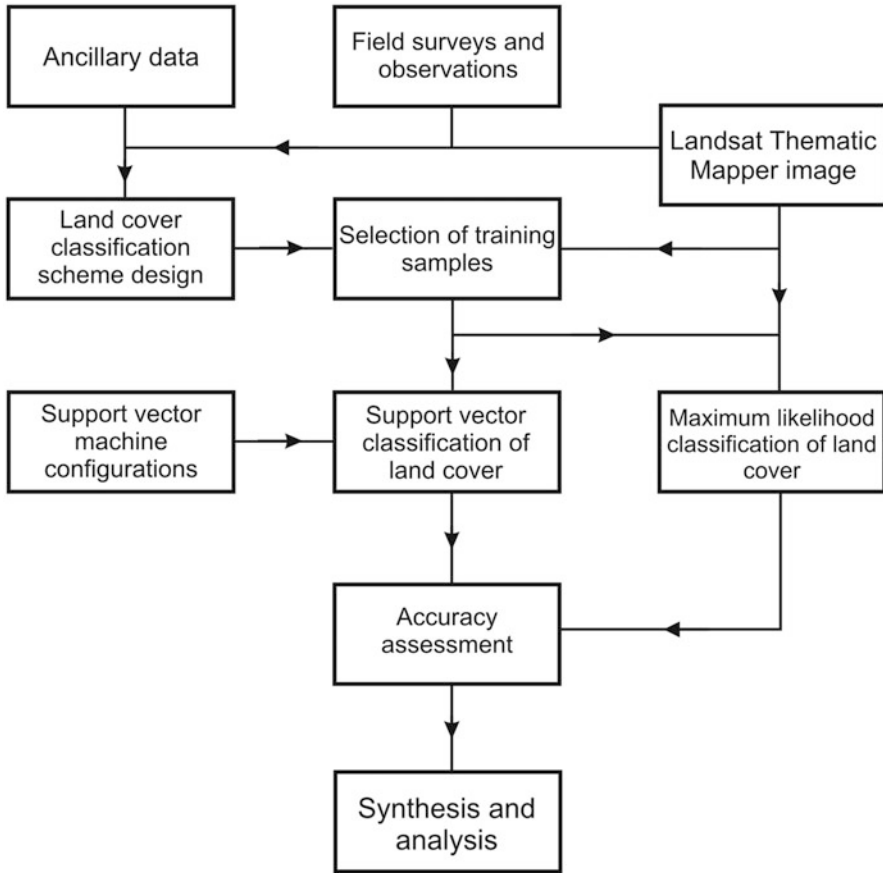


Fig. 13.1 Flowchart of the working procedural route used in this study

six non-thermal infrared bands, and 120 m for the thermal band. It was projected into the Universal Transverse Mercator Zone 16N with NAD 83 as the horizontal datum. Only six non-thermal infrared bands were used for land cover classification.

13.3.2 Classification Scheme and Training Samples

We designed a land use/cover classification scheme based on the Anderson scheme (Anderson et al. 1976) and our field surveys across the Atlanta metropolitan area. The study area covers a mosaic of different land use cover types, and our classification system includes ten major categories: high-density urban, low-density urban, barren or fallow land, pasture and cropland, grassland, shrub and scrub, evergreen forest, deciduous forest, mixed forest, and water (Table 13.1 and Fig. 13.4).

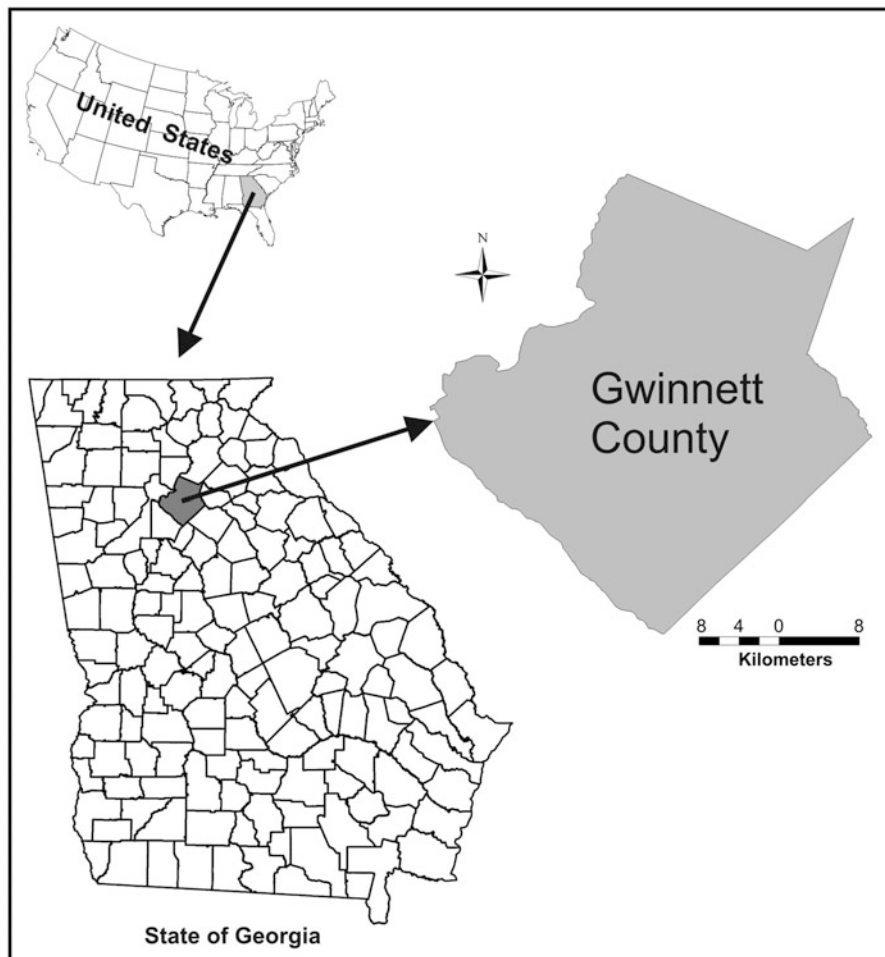


Fig. 13.2 Location of the study site. It covers the entire Gwinnet County in the State of Georgia, USA

After the classification scheme was adopted, we carefully selected training samples for each of the ten major categories by using several reference sources such as the high-resolution images from Google Earth and the 2006 National Land Cover Data (NLCD). Note that each information class listed in Table 13.1 may include multiple spectral classes. For the information classes with multiple spectral classes, we collected at least one training set with 25–35 pixels for each spectral class. Specifically, eight information classes, namely, high-density urban, low-density urban, barren or fallow land, pasture and cropland, grassland, evergreen forest, mixed forest, and water, are comprised of training data from multiple spectral classes. For the high density urban class, training samples were collected for three spectral classes with one for large roofs and the other two for parking lots

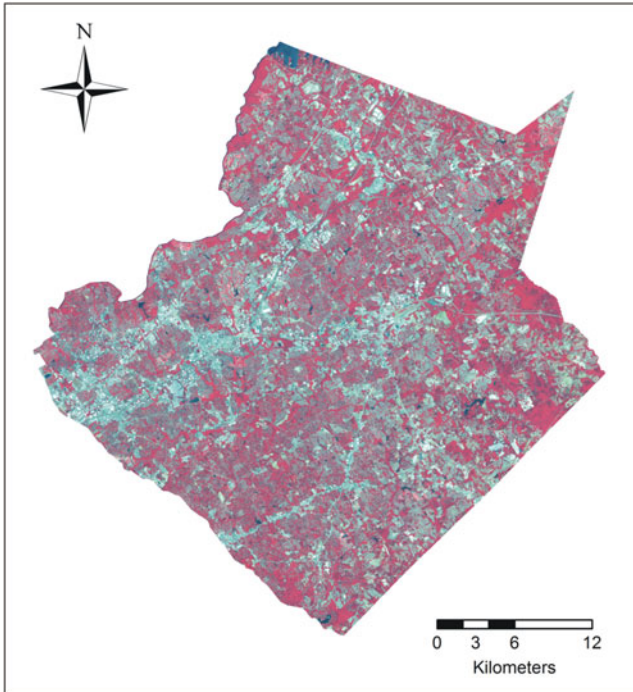


Fig. 13.3 The Landsat Thematic Mapper (TM) image used in this study. It was clipped to match the geographic coverage of Gwinnett County, Georgia. Note that the image is displayed in false color composite

with various pavement materials. For grassland, training samples were collected for two spectral classes with one for golf course with a bright color and the other for urban green spaces with low woody cover. Two spectral classes were defined for evergreen forest with one for highland evergreen forest and the other for wetland evergreen forest. For mixed forest, training samples were collected for two spectral classes that vary due to soil types. We calculated the spectral separability for each pair of the spectral classes, and finally selected 20 classes for use in the training phase of the SVM classification that will be discussed later.

13.3.3 SVM Configuration and Classification

As discussed before, SVM parameter settings can affect the classification performance (Huang et al. 2002; Kavzoglu and Colkesen 2009). Among them, the kernel type, error penalty, and Gamma term are the three most critical parameters (Yang 2011). We configured a support vector machine with radial basis function as the kernel type, a moderate error penalty value ($C = 100$), and a Gamma term equaling

Table 13.1 Land cover classification system, training sample size and reference data size

Class name	Description	Training sample size (# pixels)	Reference sample size (# pixels)
High-density urban	More than two-thirds impervious surfaces, mainly commercial, industrial, institutional facilities with large roofs, and public retail buildings, large transportation facilities	60	52
Low-density urban	Residential areas with impervious surfaces account for lower than two-thirds of total cover, including residential developments, smaller urban service buildings, such as detached stores and restaurants, state highways	54	84
Barren or fallow land	Urban areas with low percentages of constructed materials, vegetation, and low level of impervious surfaces, including bare soil lands, small amount fallow lands, exposed rock, mines and quarries	71	48
Grassland	Herbaceous cover, trees and shrub less than 10 %. Parks, lawns and golf courses	55	86
Pasture and cropland	Grazing area, field crops, horticulture, and vegetable	41	52
Shrub and scrub	Residential and agricultural shrub, scrub, orchards, groves, and transitional vegetation areas	27	47
Evergreen forest	Trees remain green throughout the year, wetland evergreen forests included, mainly cedar and pine trees	47	55
Deciduous forest	Trees lose their leaves when the dry or cold season, wetland deciduous forests included, mainly oak, maple, elm, and hickory	31	50
Mixed forest	Either evergreen or deciduous trees also mixed with shrub and scrub less than 10 %	49	114
Water	Rivers, streams, lakes, reservoirs	125	54

to 0.143 (Yang 2011). We used this SVM configuration to classify the Gwinnett subset of the 7-band TM image with the training samples described above. For comparison purpose, we also used the same training samples to classify the same image by using the maximum likelihood classifier (MLC) that has been widely used. After the implementation of SVM and MLC, we combined the 20 spectral classes into 10 information classes prior to the thematic accuracy assessment (Fig. 13.5).

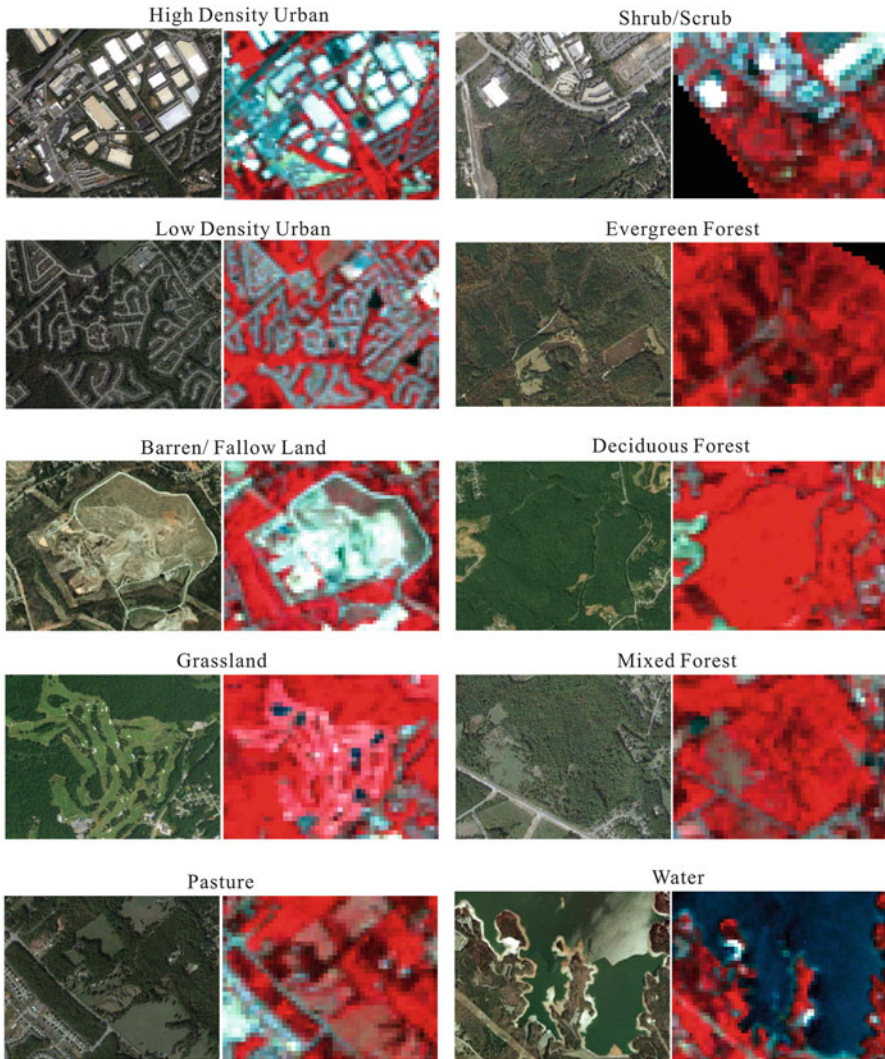


Fig. 13.4 Major land cover types shown in the very high resolution image (Source: Google Earth) and the corresponding Landsat Thematic Mapper (TM) image used in this study. For each image pair, the *left* is a very high resolution image displayed in natural color composite and the *right* is a TM image subset in false color composite

13.3.4 Accuracy Assessment

The accuracy assessment was conducted by using visual comparison and the error matrix approach. The visual comparison is qualitative by nature, while the error matrix approach is a quantitative method that compares the classification map with the ground reference information (Congalton 1991). A total of 498 reference

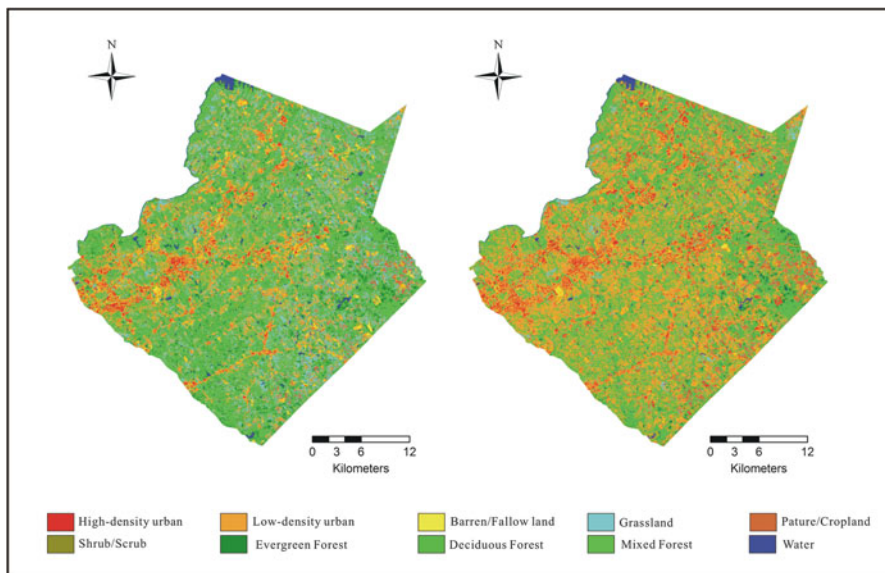


Fig. 13.5 Land cover maps produced by using support vector machines (SVM) (*Left*) and maximum likelihood classifier (*Right*)

samples were generated through the stratified random sampling method (Table 13.2). The identity of each sample was determined by the combined use of high spatial resolution data from Google Earth, USGS 2006 National Land Cover Data, and our field survey data. Kappa coefficients were calculated to quantify the overall and categorical accuracies (Congalton 1991).

13.3.5 Results and Analyses

The classification maps from SVM and MLC are displayed in Fig. 13.5. Both maps were geographically linked with the original remote sensor image, and specific land cover categories were further checked. In general, both maps show an overall correct land cover classification but misclassified areas or pixels can be clearly observed. While the two maps do not show much different large landscape patches, the one from SVM shows many scattered, isolated patches being correctly classified. In terms of specific classes, grassland and low density urban are classified differently, as shown on the two maps. Some grassland patches on the map from SVM were misclassified as low density urban class on the other map. And some mixed forest patches were classified as low density area, and some small patches of evergreen forests and shrubs were classified as mixed forest. Thus, if the spectral characteristics of a class are similar to other classes or if a class is dominated by mixed pixels, SVM clearly performed better than MLC.

Table 13.2 Summary of the thematic accuracy assessment for the two land cover maps produced by support vector machines (SVM) and maximum likelihood classifier (MLC), respectively

Class name	Conditional kappa coefficient (K)		100×(KSVM-KMLC)/KMLC
	Support vector machines (SVM)	Maximum likelihood classifier (MLC)	
High density urban	0.80	0.57	40 %
Low density urban	0.69	0.39	77 %
Barren/fallow land	0.71	0.80	-11 %
Grassland	0.70	0.55	27 %
Pasture	0.81	0.56	45 %
Shrub/scrub	0.76	0.69	10 %
Evergreen forest	0.94	0.94	0 %
Deciduous forest	0.95	0.88	8 %
Mixed forest	0.77	0.55	40 %
Water	1.00	1.00	0 %
Overall kappa coefficient	0.80	0.58	38 %

To further assess the performance of SVM when separating spectrally complex landscape categories, several sites were selected for a closer look. Figure 13.6 illustrates the original TM image, high resolution image from Google Earth, the two classified maps from SVM and MLC, for each of the three sites. For the two spectrally complex categories, namely, low density urban and mixed forest, MLC tended to include more neighboring pixels into these classes. MLC also misclassified some evergreen forest patches into water, barren land patches into high density urban, and grassland patches into low density urban and cropland. Contrastingly, SVM seemed to have done a better job in mapping spatially scattered patches. And SVM had correctly classified the residential patches on all the three sites and the pasture patches on Site 2.

For quantitative accuracy assessment, Kappa coefficient and conditional Kappa coefficients were calculated and summarized in Table 13.2. If judging by the overall Kappa coefficient, SVM significantly outperformed MLC. As for specific classes, SVM significantly surpassed MLC in terms of classification accuracy for most classes, except evergreen forest and water. And the largest improvements were with the categories of high density urban, low density urban, pasture, and mixed forest, of which the second and last classes are most spectrally complex. SVM also showed a moderate improvement for grassland. However, SVM and MLC had almost identical classification accuracies for several relatively homogeneous classes, such as evergreen forest and water.

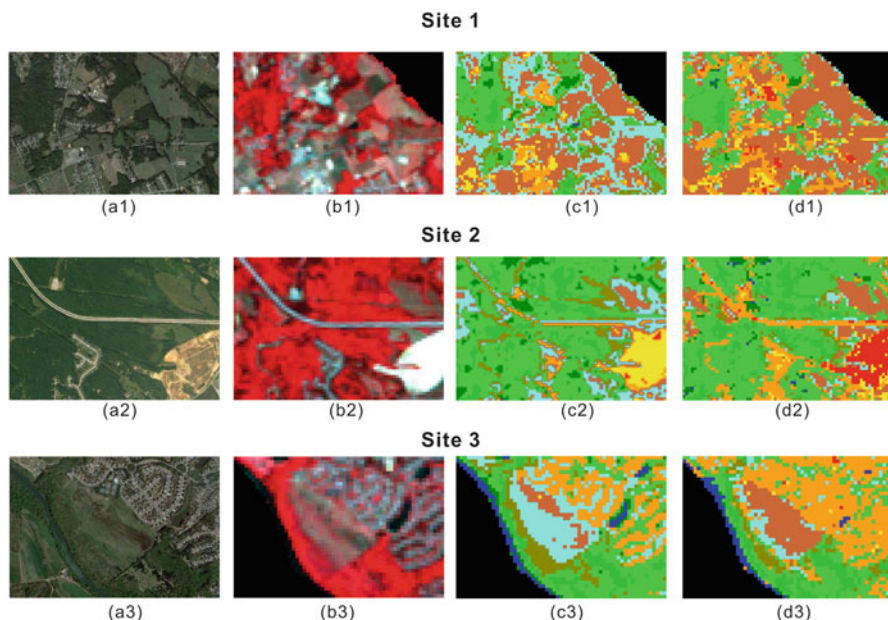


Fig. 13.6 Visual comparison of the land cover classification by support vector machines (SVM) and maximum likelihood classifier (MLC) at the three selected sites. Note that a1, a2, and a3 are natural color composites of very high resolution satellite images from Google Earth; b1, b2, and b3 are false color composites of the Landsat TM image used in this study; c1, c2, and c3 are subsets of the land cover classification by SVM; and d1, d2, and d3 are subsets of the classification by MLC. See Fig. 13.5 for specific legends for the land cover maps

13.4 Conclusion

In this chapter, we have reviewed the research status of using support vector machines (SVM) for land cover mapping with special attention on heterogeneous landscape types. Then, we have implemented this technique to map various land cover types in an urban area from a satellite remote sensor image. Our studies further confirm that SVM can significantly outperform the maximum likelihood classifier (MLC), the most widely used pattern recognition method in the remote sensing community. We found that SVM can significantly improve mapping accuracy, particularly for spectrally and spatially complex land cover categories.

Acknowledgements The authors like to thank the Florida State University for the time release in conducting this work. The research was partially supported by the Florida State University Council on Research and Creativity, CAS/SAFEA International Partnership Program for Creative Research Teams of “Ecosystem Processes and Services”, the Natural Science Foundation of China through the grant “A Study on Environmental Impacts of Urban Landscape Changes and Optimized Ecological Modeling” (ID 41230633).

References

- Alberti M (2008) *Advances in urban ecology: integrating humans and ecological processes in urban ecosystems*. Springer, New York
- Bazi Y, Melgani F (2006) Toward an optimal SVM classification system for hyperspectral remote sensing images. *Geosci Remote Sens IEEE Trans* 44(11):3374–3385
- Camps-Valls G, Gomez-Chova L, Muñoz-Mari J, Vila-Frances J, Calpe-Maravilla J (2006) Composite kernels for hyperspectral image classification. *IEEE Geosci Remote Sens Lett* 3(1):93–97
- Camps-Valls G, Bandos T, Zhou D (2007) Semi-supervised graph-based hyperspectral image classification. *IEEE Trans Geosci Remote Sens* 45(10):3044–3054
- Congalton RG (1991) A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sens Environ* 37(1):35–46
- Del Frate F, Pacifici F, Schiavon G, Solimini C (2007) Use of neural networks for automatic classification from high-resolution images. *IEEE Trans Geosci Remote Sens* 45(4):800–809
- Demir B, Ertürk S (2009) Clustering based extraction of border training patterns for accurate SVM classification of hyperspectral images. *IEEE Geosci Remote Sens Lett* 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? *Int J Remote Sens* 29(4):1185–1206
- Duda RO, Hart PE, Stork DG (2001) *Pattern classification*. Wiley, New York
- Foley JA, DeFries R, Asner GP, Barford C, Bonan G, Carpenter SR, Chapin FS, Coe MT, Daily GC, Gibbs HK, Helkowski JH, Holloway T, Howard TEA, Kucharik CJ, Monfreda C, Patz JA, Prentice IC, Ramankutty N, Snyder PK (2005) Global consequences of land use. *Science* 309:570–574
- Foody GM (2008) RVM-based multi-class classification of remotely sensed data. *Int J Remote Sens* 29(6):1817–1823
- Foody GM, Arora MK (1997) An evaluation of some factors affecting the accuracy of classification by an artificial neural network. *Int J Remote Sens* 18:799–810
- Foody GM, Mathur A (2004a) Toward intelligent training of supervised image classifications: directing training data acquisition for SVM classification. *Remote Sens Environ* 93(1–2):107–117
- Foody GM, Mathur A (2004b) A relative evaluation of multiclass image classification by support vector machines. *IEEE Trans Geosci Remote Sens* 42(6):1335–1343
- Foody GM, 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 Sens Environ* 103(2):179–189
- Haykin S (1999) *Neural networks: a comprehensive foundations*, 2nd edn. Prentice Hall, Upper Saddle River
- Heikkinen V, Tokola T, Parkkinen J, Korpela I, Jaaskelainen T (2010) Simulated multispectral imagery for tree species classification using support vector machines. *IEEE Trans Geosci Remote Sens* 48(3):1355–1364
- Hoffer RM (1978) Biological and physical considerations in applying computer aided analysis techniques to the remote sensor data. In: Swain PH, Davis SM (eds) *Remote sensing: the quantitative approach*. McGraw-Hill, New York, pp 227–289
- Hsu CW, Lin CJ (2002) A comparison of methods for multiclass support vector machines. *Neural Netw IEEE Trans* 13(2):415–425
- Huang C, Davis LS, Townshend JRG (2002) An assessment of support vector machines for land cover classification. *Int J Remote Sens* 23(4):725–749
- Hughes GF (1968) On the mean accuracy of statistical pattern recognizers. *IEEE Trans Inf Theory* 14:55–63
- Jensen JR (2005) *Introductory digital image processing: a remote sensing perspective*, 5th edn. Prentice Hall, Upper Saddle River
- Kavzoglu T, Colkesen I (2009) A kernel functions analysis for support vector machines for land cover classification. *Int J Appl Earth Obs Geoinfr* 11(5):352–359

- Keramitsoglou I, Sarimveis H, Kiranoudis CT, Kontoes C, Sifakis N, Fitoka E (2006) The performance of pixel window algorithms in the classification of habitats using VHSR imagery. *ISPRS J Photogramm Remote Sens* 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 Sens Environ* 86(4):530–541
- Liu T, Yang X (2013) Mapping urban vegetation using layered classification and multiple endmember spectral mixture analysis. *Remote Sens Environ* 133:251–264
- Mountrakis G, Im J, Ogole C (2011) Support vector machines in remote sensing: a review. *ISPRS J Photogramm Remote Sens* 66(3):247–259
- Pal M, Mather PM (2005) Support vector machines for classification in remote sensing. *Int J Remote Sens* 26(5):1007–1011
- Plaza A, Benediktsson JA, Boardman JW, Brazile J, Bruzzone L, Camps-valls G, Chanussot J, Fauvel M, Gamba P, Gualtieri A, Marconcini M, Tilton JC, Trianni G (2009) Recent advances in techniques for hyperspectral image processing. *Remote Sens Environ* 113(1):110–122
- Richards JA, Landgrebe DA, Swain PH (1982) A means for utilizing ancillary information in multispectral classifications. *Remote Sens Environ* 12(6):463–477
- Rodriguez-Galiano VF, Ghimire B, Rogan J, Chica-Olmo M, Rigol-Sanchez JP (2012) An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS J Photogramm Remote Sens* 67:93–104
- Schmidt KS, Skidmore AK, Kloosterman EH, Van Oosten H, Kumar L, Janssen JAM (2004) Mapping coastal vegetation using an expert system and hyperspectral imagery. *Photogramm Eng Remote Sens* 70(7):703–715
- Skidmore AK, Turner BJ, Brinkhof W, Knowles E (1997) Performance of a neural network: mapping forests using GIS and remotely sensed data. *Photogramm Eng Remote Sens* 63(5):501–514
- Su L (2009) Optimizing support vector machine learning for semi-arid vegetation mapping by using clustering analysis. *ISPRS J Photogramm Remote Sens* 64(4):407–413
- Su L, Huang X (2009) Support vector machine (svm) classification: comparison of linkage techniques using a clustering-based method for training data selection. *GISci Remote Sens* 46(4):411–423
- Turner BL, Meyer WB (eds) (1994) *Changes in land use and land cover: a global perspective*. Cambridge University Press, Cambridge, UK
- Vapnik VN (1995) *The nature of statistical learning theory*. Springer, New York
- Yang X (2002) Satellite monitoring of urban spatial growth in the Atlanta metropolitan region. *Photogramm Eng Remote Sens* 68(7):725–734
- Yang X (2011) Parameterizing support vector machines for land cover classification. *Photogramm Eng Remote Sens* 77(1):27–37
- Yang X, Lo CP (2002) Using a time series of normalized satellite imagery to detect land use/cover change in the Atlanta, Georgia metropolitan area. *Int J Remote Sens* 23(9):1775–1798
- Zhou L, Yang X (2011) An assessment of internal neural network parameters affecting image classification accuracy. *Photogramm Eng Remote Sens* 77(12):1233–1240