Korean Journal of Remote Sensing, Vol.24, No.5, 2008, pp.507~515

Support Vector Machine Classification Using Training Sets of Small Mixed Pixels: An Appropriateness Assessment of IKONOS Imagery

Byeong-Hyeok Yu*† and Kwang-Hoon Chi*,**

*Department of Geoinformatic Engineering, School of Science, University of Science and Technology

**Geoscience Information Center, Korea Institute of Geoscience and Mineral Resources

Abstract : Many studies have generally used a large number of pure pixels as an approach to training set design. The training set are used, however, varies between classifiers. In the recent research, it was reported that small mixed pixels between classes are actually more useful than larger pure pixels of each class in Support Vector Machine (SVM) classification. We evaluated a usability of small mixed pixels as a training set for the classification of high-resolution satellite imagery. We presented an advanced approach to obtain a mixed pixel readily, and evaluated the appropriateness with the land cover classification from IKONOS satellite imagery. The results showed that the accuracy of the classification based on small mixed pixels is nearly identical to the accuracy of the classification based on large pure pixels. However, it also showed a limitation that small mixed pixels used may provide insufficient information to separate the classes. Small mixed pixels of the class border region provide cost-effective training sets, but its use with other pixels must be considered in use of high-resolution satellite imagery or relatively complex land cover situations.

Key Words : Support Vector Machine; Mixed Pixels; Training set.

1. Introduction

A supervised image classification is one of the widely used analyzes in remote sensing research. The production of a supervised classification is a thematic map that provides a snapshot representation of spatial distribution of a particular subject. The accuracy of a supervised classification is generally dependent on the suitability of the training set used. Thus, the training stage in a classification is crucial to get high accuracy. However, the nature of an ideal training set is unclear, and may differ according to the type of selected classifier. Many studies have generally promoted a relatively uniform or classifier-independent approach to training set design. Such an approach has used a large number of pure pixels.

Support Vector Machine (SVM) classifier has been shown that it can be more accurate than other widely used classifiers such as neural networks, decision trees and maximum likelihood (Huang *et al.*, 2002).

Received October 1, 2008; Revised October 12, 2008; Accepted October 20, 2008.

† Corresponding Author: Byeong-Hyeok Yu (bynghyk@ust.ac.kr)

SVM based classification only uses a limited training set contribute to the fitting of an optimal separating hyperplane (OSH) between classes. These training samples lie at the edge of the class distributions and between the classes centroids, therefore the acquisition of training samples from the border region can help reduce the training set size without loss of classification accuracy (Foody and Mathur, 2004a).

Foody and Mathur (2006) evaluated the usability of small training sets comprising mixed spectral responses, in which the DN value of the pixel on transact was averaged with the DN value of its neighbor to one side of the boundary in one of the classes. The authors also used a three waveband multispectral SPOT-HRV (pixel size of 20m) image acquired of an agricultural test site, in which was a very simple classification problem.

In this paper, we evaluated the appropriateness of the use of small mixed pixels as a training set for the classification of high-resolution satellite imagery. In a training stage, we presented an advanced approach no making artificially mixed pixels to the training of SVM. We also used a high-resolution IKONOS imagery to classify 4m-resolution multispectral bands with 1m-resolution panchromatic band. The usability of small mixed pixels was evaluated in terms of cost-effective training and accuracy comparison.

2. Methodology

1) SVM classification

The SVM classification is based on the fitting an OSH between classes using the training samples that lie at the edge of the class distributions, called support vectors. All of the other training samples are ignored because they do not contribute to the estimation of hyperplane location.

The simplest means of SVM classification is for the situation that the two classes are linearly separable in q dimensional space (Foody and Mathur, 2004a). The training set is defined as: $\{x_i, y_i\}$, $i = 1, ..., r$, $y \in \{1, -1\}$.

A hyperplane in feature space is defined by the equation: $wx + b = 0$, where x is a point lying on the hyperplane, w is normal to the hyperplane and b is the bias (Fig. 1). A separating hyperplane for the two classes is defined as: $wx_i + b \geq +1$(for the class $y_i = +1$) and (for the class $y_i = -1$). These two equations may be combined as:

$$y_i(wx + b) - 1 \geq 0 \qquad (1)$$

The support vectors of the two classes lie on two hyperplanes are defined as: $wx_i + b = \pm 1$. The margin between these planes is $2/\|w\|$ and the analysis aims to maximize this margin through the constrained optimization problem,

$$\min\left\{\frac{1}{2}\|w\|^2\right\} \qquad (2)$$

under the inequality constraints of Eq. (1).

If the classes are not linearly separable, slack variables, $\{\xi\}_{i=1}^{r}$, that indicate the distance the sample is from the hyperplane P1 or P2 passing through the support vectors of the class to which the sample belongs, and so the amount of violation of the

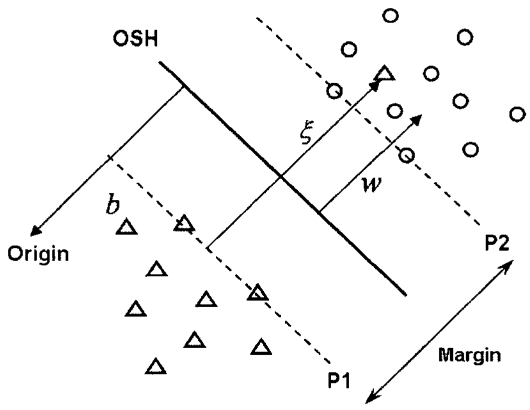

Fig. 1. The case of outlier exists in SVM classification.

constraints allowed, is introduced (Fig. 1). Thus, Eq. (1) may then be rewritten as:

$$y_i(wx_i + b) > 1 - \xi_i \quad (3)$$

If outliers exist in the training set, Eq. (3) can always be satisfied by making ξ_i very large, thus a penalty term $C\sum_{i=1}^{r} \xi_i$ is added to penalize solutions. The constant C controls the trade-off between allowing training errors and forcing rigid margins. With the addition of the penalty term to the analysis, the optimization problem is re-written as,

$$\min\left[\frac{\|w\|^2}{2} + C\sum_{i=1}^{r} \xi_i\right] \quad (4)$$

under the constraints of Eq. (3).

SVM classification may be extended to allow for non-linear decision surfaces. For this, the input data are mapped into a high dimensional space through some non-linear mapping, which has the effect of spreading the distribution of the data points in a way that facilitates the fitting of a linear hyperplane. With this, the classification decision function becomes

$$f(x) = \mathrm{sgn}\left(\sum_{i=1}^{r} \alpha_i y_i k(x, x_i) + b\right) \quad (5)$$

where α_i are Lagrange multipliers and $k(x, x_i)$ is a kernel function. The magnitude of α_i is determined by the parameter C and has a range of 0-C. The kernel used must meet Mercer's condition (Vapnik, 1995) and one such kernel that is widely used is the radial basis function,

$$k(x, x_i) = e^{-\gamma\|(x - x_i)\|^2} \quad (6)$$

where γ is the parameter controlling the width of the Gaussian kernel.

The accuracy of a SVM classification is dependent on the magnitude of the parameters C and γ. With a large value of C and/or γ, there is a tendency for the SVM to over-fit to the training set, yielding a classifier that may generalize poorly. Thus, the magnitude of C and γ must be determined carefully.

2) An advanced approach to the training of SVM

The design of the training stage of a classification has two types of purpose. One is to accurately describe the classes; so a large number of pure pixels are required. The other is to provide the critical information needed to separate the classes; the nature of the classifier used should be informed, and a number of pixels used can be effectively reduced.

With a SVM, the hyperplanes to separate the classes in feature space are presented as Fig. 2. In Fig.

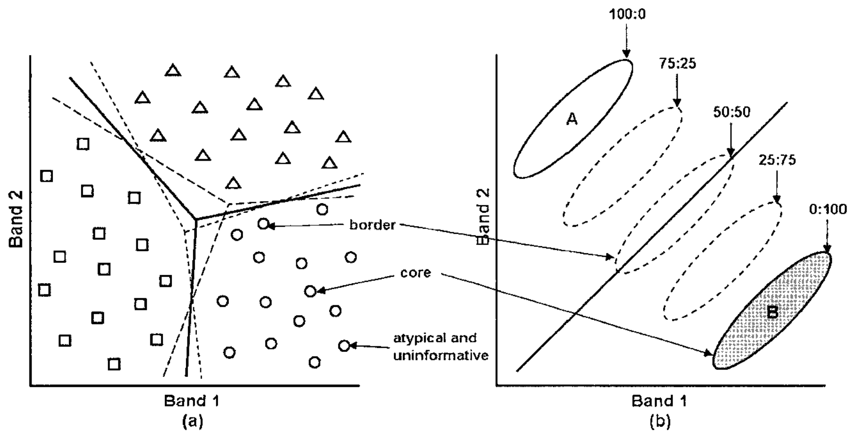

Fig. 2. The location of pure and mixed pixels in feature space.
(a) The spatial location of the training sets, (b) The proportion of the training sets.

2(a), the three sets of lines plotted indicate candidate hyperplanes and the solid line indicates the hyperplane with the greatest generalization ability. A border training samples are the most critical for the SVM classification, while a core and an atypical training samples are uninformative in the SVM classification.

In Fig. 2(b), the proportions of pure and mixed spectral responses in feature space are presented. In transect that the hyperplane lies, a mixed pixel is comprised of equal proportions of the two classes. The mixed pixel to be used as training set should be biased towards a class such as 25:75 or 75:25.

Foody and Mathur (2006) presented the use of artificially mixed pixels that were averaged with its neighbor pixels to one side of the boundary in one of the classes. The approach presented may be effectively used in cases if it is difficult to acquire information on the exact nature of the mixing. However, the presented approach has difficulty in its application, since the boundary between classes to be used should be parallel or diagonal in the pixel grid.

Thus, we presented an advanced approach no making artificially mixed pixels to the training of SVM. The approach supposes that a panchromatic band referenced is mainly composed of pure pixels, and multi-spectral bands classified are more mixed. Between the comparisons of two spatial resolutions, the mixed pixels for the training can be obtained readily and accurately. In this research, we used a high-resolution IKONOS imagery to classify 4m-resolution multispectral bands with 1m-resolution panchromatic band.

3) Study area and data set

The portion image of high-resolution IKONOS satellite imagery, belonging to Daejeon area acquired on May 9, 2002, was used (Fig. 3). The study area showed general land cover properties that may be classified into Built-up (B), Land (L), Asphalt (A), Forest (F), and Shadow (S). The B and L classes might have more segmentations as they have multiple colors in the image. The commonly used land cover classification may not separate them. Thus, we used only 5 classes for the classification.

4) Training stage using different approaches

In general, many studies have showed that classification accuracy tends to be strongly related to training set size (Arora and Foody, 1997; Foody and Mathur, 2004b; Foody *et al.*, 1995; Pal and Mather, 2003; Zhuang *et al.*, 1994). It was also found that the smallness of the training set represents a major problem and may derive a low accuracy in a classification. Alternatively, simple heuristics are often used to determine training sample size (Mather,

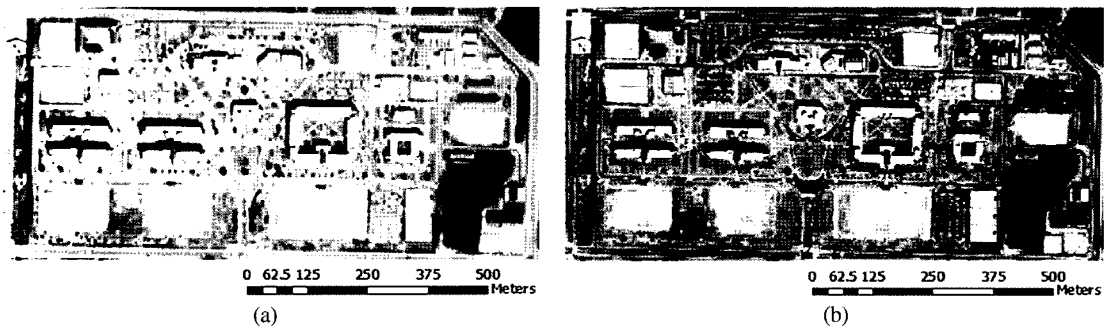

Fig. 3. The portion of IKONOS imagery for the study area.
(a) RGB composites, (b) panchromatic band.

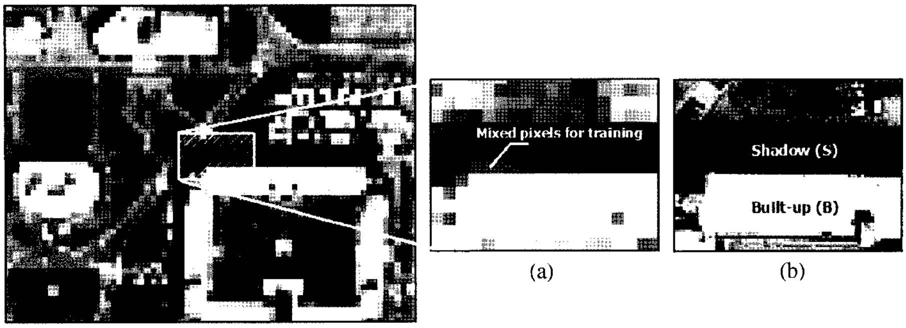

Fig. 4. An example of mixed pixels for classification SVM.
(a) RGB composites, (b) panchromatic band.

2004; Van Niel *et al.*, 2005).

Three types of training dataset were used for comparison. The first training set was acquired using conventional approach. The samples of each class were collected from the centre location of the classes that was expected to be pure pixels. The number of sampling was determined by the heuristic 30 p method, where p means the number of wavebands used. We used the multi-spectral datasets consisting of 4 bands, so the required number of training set for each class was 120. Therefore, a total of 600 training sets were used to evaluate the classification.

The second training set was acquired from the small mixed pixels that were based on each class in the decision boundary; it was theoretically laid closely on an optimal separating hyperplane. Fig. 4 showed an example of a training set acquisition. In the panchromatic band, the two classes (shadow (S) and Built-up (B)) were identified apparently. But the multi-spectral bands showed a mixed distribution of the two classes because of low spatial resolution. The edge between the classes indicated the biased values to the shadow (S), which might be used as training set for the S class. In many cases, the sampling of mixed pixels was not difficult and it had an advantage of getting the small accounts readily. In this experiment, only 138 training pixels were acquired from the image. In which, 29 pixels for B, 27 pixels for Land (L), 35 pixels for Asphalt (A), 18 pixels for Forest (F) and 29 pixels for S. The total number was barely 1/5 compared to the conventional training sets. The total number is barely 1/5 comparing the conventional training sets.

The final training set was collected using the conventional method, but the numbers of pixels for each class were constrained to same as small mixed training set. The purpose of the final training set was to compare the classification ability between conventional and small mixed training sets. The spatial location of the training pixels was considered closely to the one used for small mixed training set to improve the reliability of the experiment.

3. Results and Discussion

To avoid optimistic bias (Hammond and Verbyla, 1996), accuracy assessment was typically based on a sample of pixels not used in training the classifier.

In this experiment, a complete classification map was created manually using visual interpretation of the panchromatic band. It was used as a ground truth data to evaluate the accuracy of the three land cover maps derived from SVM classification based on three training sets (Fig. 5). The land cover map was also derived from the SVM classification derived with the

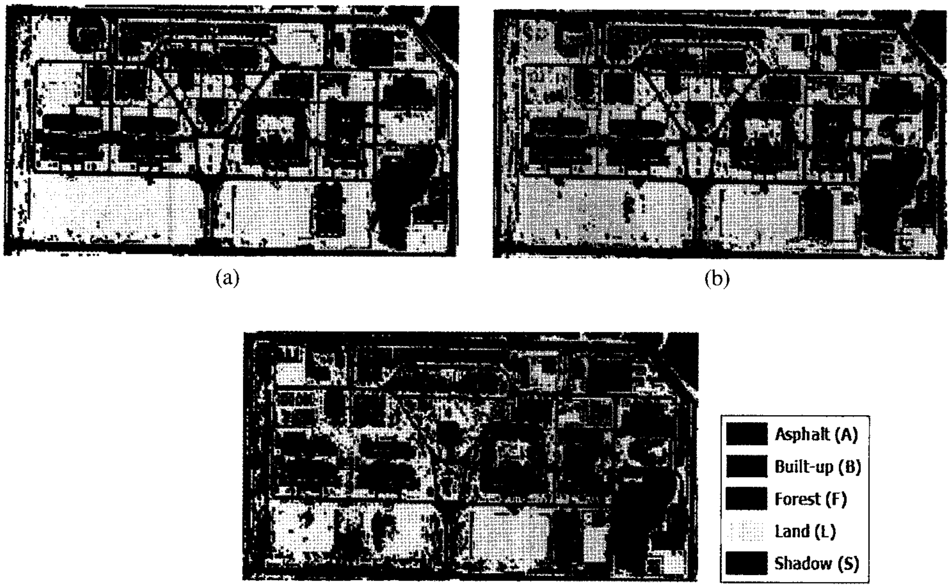

Fig. 5. The land cover maps derived from the use of different training set.
(a) Conventional (b) small mixed (c) conventional, same size.

use of each training set approaches (Fig. 5).

The classification using conventional training set, 30p heuristic, classified the image with overall accuracy of 86.25% and the Kappa coefficient was 0.7993. In the confusion matrixes (Table 1) and Fig. 6, it was identified that the land and forest classes are classified more accurately than the others.

The small mixed pixels based training sets also classified the image with 84.52% having Kappa coefficient of 0.7777. The pattern of confusion matrixes in the two results were nearly identical (Tables 1 and 2, Fig. 6). Besides, the small mixed training set used only 138 ground truths on total fields, whereas the conventional training set used 120 ground truths on each field (5 fields).

The overall accuracy of the classification based on the third training sets was relatively lesser (70.64%) with 0.5929 Kappa coefficient. The pattern of accuracies also showed significant differences compared to other methods.

The results showed that the accuracy of the classification based on small mixed pixels is nearly identical to the accuracy of the classification based on large pure pixels. It showed that a small mixed pixel

Table 1. Confusion matrixes derived from the use of the conventional training set

Class	Ground Truth (Pixels)					
	Built-up (B)	Land (L)	Asphalt (A)	Forest (F)	Shadow (S)	Total
Built-up (B)	2624	66	0	0	287	2977
Land (L)	362	16214	489	4	65	17134
Asphalt (A)	3125	97	9208	0	58	12488
Forest (F)	50	262	0	2764	70	3146
Shadow (S)	9	0	0	177	1309	1495
Total	6170	16639	9697	2945	1789	37240

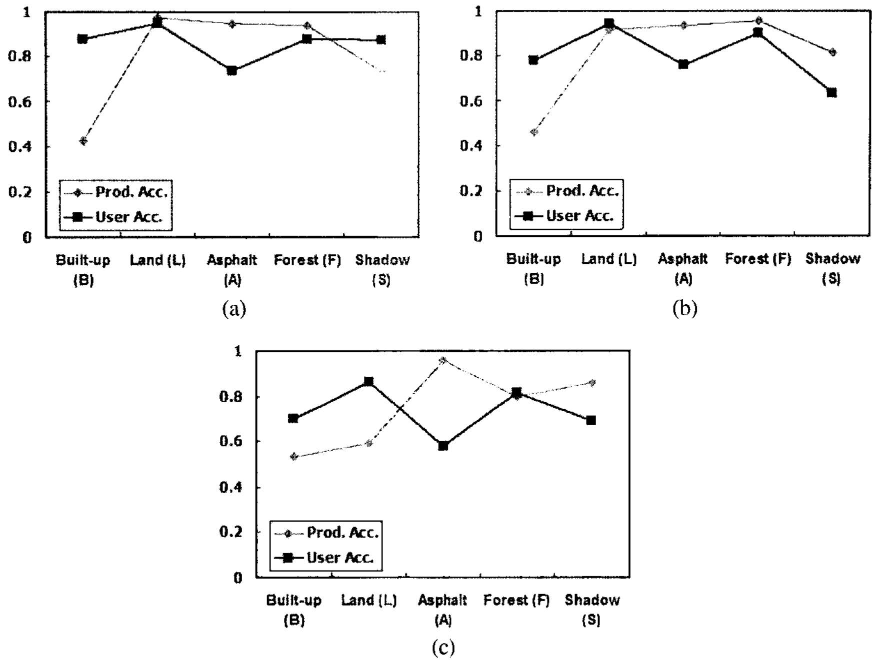

Fig. 6. The producer's and user's accuracies distribution of all classes.
(a) Conventional (b) small mixed (c) conventional, same size.

Table 2. Confusion matrixes derived from the use of the small mixed training set

Class	Ground Truth (Pixels)					
	Built-up (B)	Land (L)	Asphalt (A)	Forest (F)	Shadow (S)	Total
Built-up (B)	2842	652	57	0	96	3647
Land (L)	633	15292	167	90	0	16182
Asphalt (A)	2251	577	9066	5	84	11983
Forest (F)	41	118	0	2818	150	3127
Shadow (S)	403	0	407	32	1459	2301
Total	6170	16639	9697	2945	1789	37240

Table 3. Confusion matrixes derived from the use of the conventional training set, same size with Table 2

Class	Ground Truth (Pixels)					
	Built-up (B)	Land (L)	Asphalt (A)	Forest (F)	Shadow (S)	Total
Built-up (B)	3278	1281	128	0	0	4687
Land (L)	1270	9863	289	0	0	11422
Asphalt (A)	1468	5027	9280	2	240	16017
Forest (F)	54	468	0	2354	17	2893
Shadow (S)	100	0	0	589	1532	2221
Total	6170	16639	9697	2945	1789	37240

is more favorable than a pure pixel in terms of a cost-effective training stage of SVM. The results of the classification accuracy based on pure pixels, containing the same number based on small mixed pixels, proved that the pixels of the border region are more favorable than the pixels of the class centroid in the training.

It is not clear, however, which is better training set between small mixed pixels and large pure pixels in terms of its classification accuracies. The use of high-resolution satellite imagery such as IKONOS showed the distribution of various DN values according to the color or nature in a single class. For example, L class can separate out the portions (wet/dry) according to the soil moisture condition. The pixels of the class centroid containing these portions can be used as pure pixels, but it can be neglected to recognize small mixed pixels because of its spatial location. Experiment results of this study also showed that some pixels of the dry part in L class were misclassified as B class when small mixed pixels were used as training set of SVM (Fig. 5 (b)).

It carries a limitation that small mixed pixels of the class border region may provide insufficient information to separate the classes in the classification of high-resolution satellite. Relatively complex situations in the class distributions also present a similar limitation like this.

4. Conclusions

The previous research presented that the mixed pixels between classes are actually more useful than the pure pixels of each class. The usability of small mixed pixels as a training set for the classification of high-resolution satellite imagery, however, was not yet evaluated.

In this research, we presented an advanced approach to obtain a mixed pixel, and evaluated the accuracy of a small number of mixed pixels with land cover classification of high-resolution satellite imagery. In terms of cost-effective training of SVM, results showed that a mixed pixel of the class border region is more favorable than a pure pixel of the class centroid, similarly to results of a previous research.

In terms of classification accuracy, however, results showed that a mixed pixel may provide a limitation because it may neglect the distribution of various DN values in a single class in the use of high-resolution satellite imagery. When relatively complex situations in the class distributions are used, it may also present a similar limitation.

Finally, small mixed pixels of the class border region provide cost-effective training sets, but its use with other pixels may be more appropriate in use of high-resolution satellite imagery or relatively complex situations in the class distributions.

References

Arora, M. K. and G. M. Foody, 1997. Log-linear modeling for the evaluation of the variables affecting the accuracy of probabilistic, fuzzy and neural network classifications. *International Journal of Remote Sensing*, 18: 785-798.

Foody, G. M., M., M. B. McCulloch, and W. B. Yates, 1995. The effect of training set size and composition on artificial neural network classification. *International Journal of remote sensing*, 16: 1707-1723.

Foody, G. M. and A. Mathur, 2004a. A relative evaluation of multiclass image classification by support vector machines, *IEEE Transactions on Geoscience and Remote Sensing*, 42: 1335-1343.

Giles M. Foody and Ajay Mathur, 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: 179-189.

Giles M. Foody and Ajay Mathur, 2004b. Toward intelligent training of supervised image classifications: directing training data acquisition for SVM classification, *Remote Sensing of Environment*, 93: 107-117.

Hammond, T. O. and D. L. Verbyla, 1996. Optimistic bias in classification accuracy assessment, *International Journal of Remote Sensing*, 17: 1261-1266.

Huang, C., L. S. Davis, and J. R. G. Townshend, 2002. An assessment of support vector machines for land cover classification, *International Journal of remote sensing*, 23: 725-749.

Mather, P. M., 2004. Computer Processing of Remotely-Sensed Images, 3rd edition, Chichester: Wiley.

Pal, M. and P. M. Mather, 2003. An assessment of the effectiveness of decision tree methods for land cover classification, *Remote Sensing of Environment*, 86: 554-565.

Van Niel, T. G., T. R. McVicar, and B. Datt, 2005. On the relationship between training sample size and data dimensionality of broadband multi-temporal classification, *Remote Sensing of Environment*, 98: 468-480.

Vapnik, F., 1995. The Nature of Statistical Learning Theory, NewYork:Springer-Verlag.

Zhuang, X., B. A. Engel, D. F. Lozanogarcia, R. N. Fernandez, and C. J. Johannsen, 1994. Optimization of training data required for neuro-classification, *International Journal of Remote Sensing*, 15: 3271-3277.