



Comparing support vector machine contextual approaches for urban area classification

R. G. Negri, L. V. Dutra & S. J. S. Sant'Anna

To cite this article: R. G. Negri, L. V. Dutra & S. J. S. Sant'Anna (2016) Comparing support vector machine contextual approaches for urban area classification, Remote Sensing Letters, 7:5, 485-494, DOI: [10.1080/2150704X.2016.1154218](https://doi.org/10.1080/2150704X.2016.1154218)

To link to this article: <https://doi.org/10.1080/2150704X.2016.1154218>



Published online: 01 Mar 2016.



Submit your article to this journal [↗](#)



Article views: 194



View Crossmark data [↗](#)



Citing articles: 2 View citing articles [↗](#)

Comparing support vector machine contextual approaches for urban area classification

R. G. Negri ^a, L. V. Dutra ^b and S. J. S. Sant'Anna^b

^aInstituto de Ciência e Tecnologia, UNESP - Univ. Estadual Paulista, São José dos Campos, Brazil; ^bDivisão de Processamento de Imagens, INPE - Inst. Nacional de Pesquisas Espaciais, São José dos Campos, Brazil

ABSTRACT

Support vector machine (SVM) has been receiving a great deal of attention for remote sensing data classification. Although the original formulation of this method does not incorporate contextual information, lately different formulations have been proposed to incorporate such information, with the aim of improving the mapping accuracy. In general, these proposals modify the SVM training phase or integrate the SVM classifications in stochastic models. Recently, two new contextual versions of SVM, context adaptive and competitive translative SVM (CaSVM and CtSVM, respectively), were proposed in literature. In this work, two case studies of urban area classification, using IKONOS-II and hyperspectral digital imagery collection experiment (HYDICE) data sets were conducted to compare SVM, SVM integrated with the iterated conditional modes (ICM) stochastic algorithm, SVM smoothed using the mode filter and the recent approaches CaSVM and CtSVM. The results indicated that although it possesses a high computational cost, the CaSVM method was able to produce classification results with similar accuracy (using kappa coefficient) to those obtained using SVM integrated with ICM (SVM + ICM) and the mode filter (SVM + Mode), all of them found statistically superior to the SVM result at 95% confidence level for the IKONOS-II image. For HYDICE image, all results were found statistically insignificant at 95% confidence level. Investigation of what happens at transition regions between classes, however, showed that some methods can present superior performance. To this objective, a new performance measure, called epsilon coefficient, was introduced in this work, which measures the impact that the smoothing effect, typical of contextual methods, can have in distorting the edges between regions. With this new measure was found that CaSVM is the one which has better performance followed with SVM + ICM.

ARTICLE HISTORY

Received 30 April 2015
Accepted 8 February 2016

1. Introduction

Support vector machine (SVM) is a pattern recognition method that has become popular since its introduction in Boser, Guyon, and Vapnik (1992). The primary objective of this method is to distinguish patterns using hyperplanes with a maximum separation margin in the feature space. According to Mountrakis, Im, and Ogole (2011), the majority of scientific

investigations in the field of remote sensing that use SVM for image classification adopt the pixel-based approach. One way of possibly improving the image classification accuracy is using the neighbourhood information of a pixel, what is called contextual classification.

Integrating SVM with Markov random fields (MRFs) (Bovolo and Bruzzone 2005), implementing changes in the training core of the method (Bovolo, Bruzzone, and Marconcini 2006; Gurram and Kwon 2013) or using smoothing techniques are some approaches that have been reported in the literature for producing contextual classifications from SVM. Furthermore, Negri, Sant'Anna, and Dutra (2013) and Negri, Dutra, and Sant'Anna (2014) verified the possibility of producing contextual classifications using the SVM method by adaptations (displacements) of separating hyperplanes for each pixel of the image as a function of its context information.

This paper presents a comparative analysis of the different contextual versions of SVM, particularly those previously reported in Bovolo and Bruzzone (2005), Negri, Sant'Anna, and Dutra (2013) and Negri, Dutra, and Sant'Anna (2014) and by the simple application of the mode filter as a smoothing technique, as discussed in Cortijo and Blanca (1998). In this way, methods with distinct contextual classification principles are analysed.

This comparative analysis is based on two case studies of urban area classification using an IKONOS-II satellite image and an extract of a scene acquired by the hyperspectral digital imagery collection experiment (HYDICE) over the city of Washington, DC. The classification results are compared based on accuracy values measured using the kappa agreement coefficient, computed from test samples located inside well-known ground truth targets, as well as by adopting a new measure introduced in this work to assess the classification accuracy on the edges of mapped targets.

2. Contextual versions of SVM for image classification

Formally, a classifier is represented by a function $f: X \mapsto \Omega$ that assigns elements from an attribute space X to a class ω_i that belongs to the set of classes $\Omega = \{\omega_1, \omega_2, \dots, \omega_c\}$. Image classification consists of applying f to the pixels that compose an image I , defined on a support $S \subset \mathbb{N}^2$. Considering the image where the classification process is conducted, $I(s) = \mathbf{x}$ denotes that the coordinate $s \in S$ in I is assigned to an attribute vector $\mathbf{x} \in X$. When f is applied to each x of I , the classification C is produced, which is defined on the same support S of I . The neighbour positions of s are elements of $V_\rho(s) = \{t \in S : 0 \leq md(s, t) \leq \rho\}$, where ρ is the *neighbourhood influence radius* and $md(\cdot, \cdot)$ is the *maximum distance*, defined as $md(s, t) = \max\{|s_1 - t_1|, |s_2 - t_2|\}$, where $s, t \in \mathbb{N}^2$ such that $s = (s_1, s_2)$ and $t = (t_1, t_2)$.

SVM is a method of pattern recognition with supervised learning. This method has recently received considerable attention from the scientific community. The classification process using SVM consists of distinguishing patterns (attribute vectors) using hyperplanes for which the separating margin is maximal. A separating hyperplane is equal to the geometric location where the value of the function $f_{\text{SVM}}: X \rightarrow \mathbb{R}$ becomes null. A complete discussion of this method can be found in Theodoridis and Koutroumbas (2008).

The original SVM formulation is characterized as a supervised learning method. Moreover, in the majority of remote sensing studies in which SVM is employed for image classification, the pixel-based approach is used (Mountrakis, Im, and Ogole 2011). However, there are different proposals for adapting this method to unsupervised and semi-supervised learning paradigms (Hamel 2006; Bruzzone and Persello 2009), as well as to the region-based approach (Tzotsos 2006; Negri, Dutra, and Sant'Anna 2012).

The derivation of contextual versions from SVM is not an exception. A pioneering proposal, not only with respect to integration with MRF but also as a contextual version of SVM, is presented in Bovolo and Bruzzone (2005). For this integration, the SVM decision rule, which expresses the distance relationship between the separating hyperplane and the patterns, is converted into probability distributions according to the method presented in Platt (2000). Subsequently, these probability distributions are used in algorithms such as iterated conditional modes (ICM) (Besag 1986).

The use of smoothing techniques is another approach for producing contextual classifications based on SVM. These techniques include applying filters on the classification results and/or incorporating additional bands, such as textures, prior to the pixel-based classification. The use of the mode filter is a common choice. The mode filter smoothing process takes into account the frequency at which the pixels of a given neighbourhood $V_\rho(s)$ are classified in the class set Ω for each $s \in S$. The most frequent class inside this neighbourhood defines the contextual classification of pixel s . Note that the influence of the neighbourhood ρ is a parameter in this contextual classification process. An increase in ρ implies a more homogeneous contextual classification. The homogenization of the classification results and low computational cost are two main characteristics of this contextual classification process.

New contextual versions of SVM for image classification are presented in Negri, Sant'Anna, and Dutra (2013) and Negri, Dutra, and Sant'Anna (2014). In contrast to the aforementioned proposals, neither stochastic nor filtering techniques are adopted. Rather, a local adjustment of the separating hyperplanes based on contextual information is performed. Although the local adjustments based on the contextual information in Negri, Sant'Anna, and Dutra (2013) only consider translation movements (i.e., hyperplane shifting), with the contextual version formulated in Negri, Dutra, and Sant'Anna (2014), the hyperplane translates and rotates. The latter method is called 'competitive translative support vector machine' (CtSVM), and the former is called 'context adaptive support vector machine' (CaSVM). In general, both versions follow a similar framework, which starts from an initial pixel-based classification generated by SVM followed by local reclassification using only the information present in the neighbourhood of each pixel.

For the CaSVM method, to define the local decision rule that contextually classifies the pixel s such that $I(s) = \mathbf{x}$, it is necessary to construct a new hyperplane $f_{\text{Local}}(\mathbf{x})$ from elements in the neighbourhood $V_\rho(s)$. In the case of CtSVM, the decision rule is determined from a translation in $f_{\text{SVM}}(\mathbf{x})$. The amplitude of this translation is determined from $V_\rho(s)$. CtSVM may be considered a specific case of CaSVM. Further details regarding CaSVM and CtSVM are presented in Negri, Sant'Anna, and Dutra (2013) and Negri, Dutra, and Sant'Anna (2014), respectively.

3. A measure for edge classification assessment

Regardless of the methods and data used, the accuracy of the classification results should be analysed to determine whether they are reliable. Over the years, different accuracy measures have been developed, and each is relevant for a specific goal (Liu, Frazier, and Kumar 2007). Among the several agreement measures, the most used is the kappa coefficient.

In remote sensing applications, it is common to compute the accuracy measures using reference samples located inside the regions regarding the mapped thematic classes. A complementary approach for evaluating classification results is to quantify the performance of classifiers on the edges of objects (regions).

The term ‘edges’ is defined as the transition between two thematic classes. Pixels located on edges correspond to pixels that are members of a region of a specific class and that have some neighbour pixel that is part of another region of a different class. Based on this definition, an *edge pixel set* E is composed of pixels supposedly known associated with a minimum of two different classes and that are located at the edges between these classes. Formally, let $E = \{s_i, s_j : C(s_i) = \omega_a, C(s_j) = \omega_{\bar{a}}; i = 1, \dots, z_a; j = 1, \dots, z_{\bar{a}}\}$.

This paper introduces a new measure for quantifying the classifier performance on edges of regions. The motivation for this measure, called the *upsilon* (Y) coefficient, is to quantify not only the number of pixels correctly classified at the edges of some class but also the fidelity with respect to the position of the edges on the classification results.

The formalization of *upsilon* coefficient starts by considering a classification result C and an edge pixel set E_k with z_a pixels assigned to a given class ω_a and $z_{\bar{a}}$ pixels assigned to any other class different from ω_a , denoted by $\omega_{\bar{a}}$. The number of elements from the edge pixel set correctly classified in terms of the classes ω_a and $\omega_{\bar{a}}$ are denoted by u_a and $u_{\bar{a}}$, respectively. In turn, \bar{u}_a and $\bar{u}_{\bar{a}}$ represent the number of elements incorrectly classified with respect to the classes ω_a and $\omega_{\bar{a}}$, respectively, such that $\bar{u}_a = z_a - u_a$ and $\bar{u}_{\bar{a}} = z_{\bar{a}} - u_{\bar{a}}$. Using these quantities, we can formulate the following expression for measuring the accuracy over the edge pixel set E_k :

$$Y_k = \frac{\left(1 - \frac{\bar{u}_a}{z_a}\right) \left(1 - \frac{\bar{u}_{\bar{a}}}{z_{\bar{a}}}\right) (u_a + u_{\bar{a}})}{z_a + z_{\bar{a}}} = \frac{u_a u_{\bar{a}} (u_a + u_{\bar{a}})}{z_a z_{\bar{a}} (z_a + z_{\bar{a}})}. \quad (1)$$

In expression (1), the factors $\left(1 - \frac{\bar{u}_a}{z_a}\right)$ and $\left(1 - \frac{\bar{u}_{\bar{a}}}{z_{\bar{a}}}\right)$ serve as weights in the measure, where increasing the number of incorrect classifications with respect to either ω_a or $\omega_{\bar{a}}$ results in a lower value of Y_k . In practical terms, these factors prevent the value of Y_k from depending only on the percentage of correct classifications in ω_a or $\omega_{\bar{a}}$. After simple algebraic manipulations, the right-side equivalent expression is obtained only as a function of z_a , $z_{\bar{a}}$, u_a and $u_{\bar{a}}$. It is possible to verify in Equation (1) that $Y_k \rightarrow 0$ when $u_a \rightarrow 0$ or $u_{\bar{a}} \rightarrow 0$; conversely, $Y_k \rightarrow 1$ when $u_a \rightarrow z_a$ and $u_{\bar{a}} \rightarrow z_{\bar{a}}$ simultaneously. To formulate a global expression, a simple average from the considered ‘per-edges’ measures is adopted, that is, $Y = \frac{1}{c'} \sum_{k=1}^{c'} Y_k$, where c' is the number of edge pixel sets considered in the assessment. In general, the ω_a adopted to define any E_k and E_l are usually different in each set, for $k, l \in \{1, \dots, c'\}$ with $k \neq l$.

4. Experiments and results

In this section, two case studies are performed to analyse the performance of different contextual versions of SVM. These studies address multiclass urban area classification using high-resolution images. The contextual methods considered in these analyses were as follows: SVM integrated with the ICM algorithm (SVM + ICM), SVM results smoothed using a mode filter (SVM + Mode), CaSVM and CtSVM. SVM following a pixel-based approach was included as a reference.

The parameter p was considered to be equal to 1, 2 and 3, which determines spatial windows of 3×3 , 5×5 and 7×7 pixels, respectively. To denote these different neighbourhood radii, the respective window sizes were appended to the name of the method. The polynomial kernel function and the one-against-all multiclass strategy were adopted by all methods.

To conduct these experiments, a computer with an Intel Core i5 processor and 3 GB of RAM running the Ubuntu Linux version 14.04 operating system was used. The implementations were conducted using the Interactive Data Language version 7.1 programming language. The algorithm SVM^{Light} (Thorsten 1999) was used to train and estimate f_{SVM} and f_{LOCAL} .

In the first case study, an IKONOS-II high-resolution image, available in Santos (2013), corresponding to part of São José dos Campos City, state of São Paulo, Brazil, was used in the experiments. This image has a 4 m resolution, 600 × 600 pixels and four bands that correspond to blue, green, red and near-infrared spectral regions. The second case study uses an extract of 305 × 266 pixels of an image over the mall area in Washington, DC from the HYDICE multispectral data set, available in <https://engineering.purdue.edu/~biehl/MultiSpec/hyperspectral.html>. This image has a spatial resolution of approximately 2.8 m and 191 bands.

In the IKONOS-II image, the following classes were considered: high vegetation, low vegetation, ceramic roof, concrete/metal roof, bare soil and asphalt. For the HYDICE image, only three classes were considered: concrete/metal roof, green area (low and high vegetation as a single class) and asphalt. High vegetation refers to locations in which the surface is covered by trees, whereas areas covered by grass or even small bushes are considered to be low vegetation. Although there are other classes in the HYDICE study area, the classes that were considered allow to select well-defined edges which are distant to each other.

As previously mentioned, the methods analysed in this study were assessed in terms of their accuracy inside well-known regions through the kappa coefficient. The global upsilon coefficient formalized in Section 3 was adopted for assessing the classifiers' performance on edges. Additionally, the kappa coefficient was used to measure the accuracy on edge pixels. Such measurement was referred to as 'kappa coefficient on edges', where just the internal side of the edge was considered in this case.

Figures 1(a) and 2(a) present the images used in the case studies. Samples defined to train the analysed methods and to test the classification results using the kappa coefficient are identified in Figures 1(b) and 2(b). Figures 1(c) and 2(c) present the selected edge pixel sets for assessing the methods with respect to the classification performance on edges. The marked areas (red dotted rectangle) in Figures 1(a) and 2(a) are further discussed in Figure 5. Table 1 shows the quantity of pixels used for training the methods, for testing the results of each case study and to assess the accuracy on the edges through upsilon coefficient.

A *grid search* procedure associated with multiple cross-validation (10-fold) was performed to tune the penalty and the degree of the polynomial kernel of the SVM method as well as the parameters *maximum number of iterations* and *minimum percentage change* of the ICM algorithm.

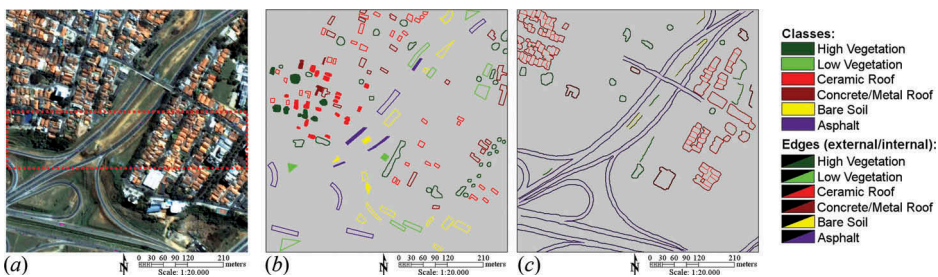


Figure 1. (a) IKONOS-II image, (b) training (solid) and testing (void) sample polygons and (c) edge pixel sets. Coordinates of the centre of the image: 23° 12'31" S, 45°53'38" W.

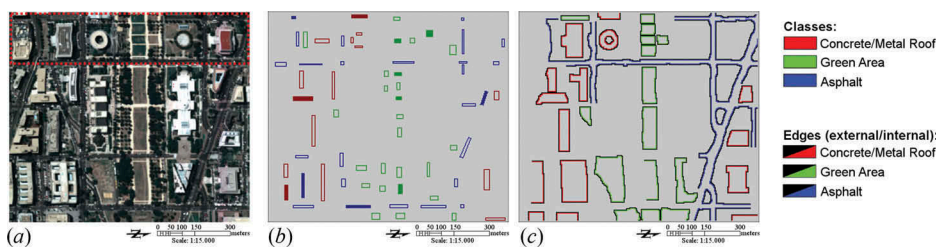


Figure 2. (a) HYDICE image, (b) training (solid) and testing (void) sample polygons and (c) edge pixel sets. Coordinates of the centre of the image: 38°53'21" N, 77° 1'9" W.

Table 1. Training, testing and edge (internal/external) samples.

IKONOS-II	Number of pixels		HYDICE	Number of pixels	
	Samples Training/testing	Edges Internal/external		Samples Training/testing	Edges Internal/external
High vegetation	1723/5828	737/828	Green area	330/1119	1310/1421
Low vegetation	439/4119	378/394	Asphalt	265/1186	2171/2108
Concrete/metal roof	469/4261	656/649	Concrete/metal roof	380/1044	1404/1504
Ceramic roof	1933/4026	5123/2792			
Bare soil	470/4299	375/395			
Asphalt	1047/3648	8700/6067			

Using the training/testing samples and the selected parameters, the SVM, CaSVM, CtSVM, SVM + Mode and SVM + ICM methods were applied for the classification of the IKONOS-II and HYDICE images. Context windows of 3×3 , 5×5 and 7×7 pixels were considered. Clearly, context windows were not used by SVM and explicitly by SVM + ICM. The graphs presented in Figures 3 and 4 show the kappa coefficient values computed from the testing samples (usual way – dashed bars) and considering the edge pixels sets (kappa coefficient on edges – dotted bars), as well as the global upsilon coefficient (solid bars) obtained from each method when applied in the study images. The error bars on the normal kappa coefficients mark one standard deviation.

Bilateral hypothesis tests were performed to compare the kappa coefficient values, computed in the usual way, from testing samples. Table 2 presents the p -values related to the performed tests. In the following discussions, decisions about differences on kappa coefficient

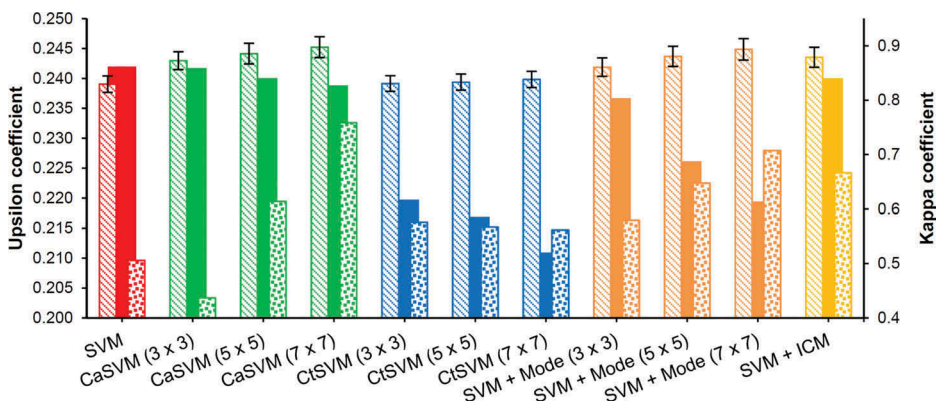


Figure 3. Kappa and upsilon coefficients related to the analysed methods for IKONOS data set. Solid bars denote the upsilon coefficient, dashed the kappa coefficient and dotted the kappa coefficient on edges.

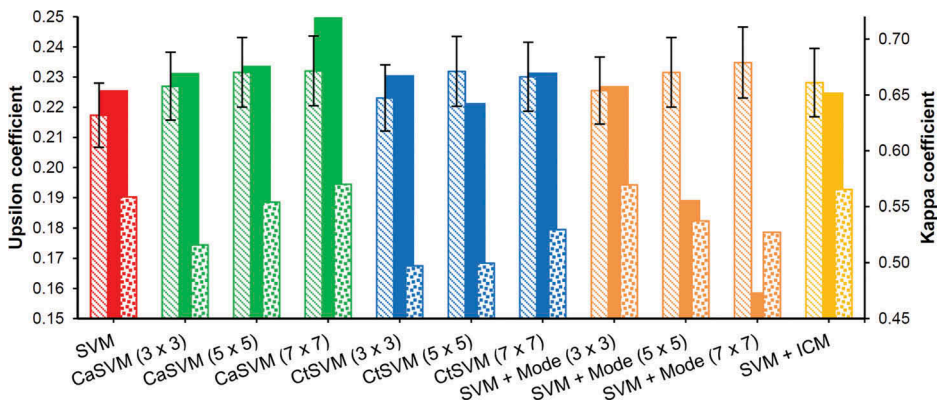


Figure 4. Kappa and upslon coefficients related to the analysed methods for HYDICE data set. Solid bars denote the upslon coefficient, dashed the kappa coefficient and dotted the kappa coefficient on edges.

Table 2. *p*-Values from hypothesis tests to verify the statistical equality between kappa agreement coefficients.

		[A]	[B]	[C]	[D]	[E]	[F]	[G]	[H]	[I]	[J]	[K]
[A]	- SVM		0.30	<u>0.02</u>	<u>0.02</u>	0.96	0.86	0.68	0.17	<u>0.03</u>	<u>0.01</u>	0.04
[B]	- CaSVM(3 × 3)	0.62		0.19	0.16	0.31	0.38	0.51	0.74	<u>0.26</u>	<u>0.11</u>	0.29
[C]	- CaSVM(5 × 5)	0.30	0.59		0.91	<u>0.02</u>	<u>0.03</u>	0.05	0.33	0.85	0.77	0.80
[D]	- CaSVM(7 × 7)	0.28	0.56	0.96		<u>0.02</u>	<u>0.02</u>	<u>0.04</u>	0.28	0.76	0.86	0.72
[E]	- CtSVM(3 × 3)	0.71	0.90	0.51	0.48		<u>0.90</u>	<u>0.71</u>	0.18	<u>0.03</u>	<u>0.01</u>	0.04
[F]	- CtSVM(5 × 5)	0.35	0.67	0.91	0.87	0.58		0.81	0.22	<u>0.05</u>	<u>0.01</u>	0.05
[G]	- CtSVM(7 × 7)	0.41	0.75	0.83	0.79	0.66	0.91		0.32	0.08	<u>0.03</u>	0.09
[H]	- SVM + Mode(3 × 3)	0.59	0.97	0.61	0.58	0.87	0.69	0.77		0.43	<u>0.21</u>	0.46
[I]	- SVM + Mode(5 × 5)	0.36	0.68	0.90	0.86	0.59	0.98	0.93	0.71		0.63	0.95
[J]	- SVM + Mode(7 × 7)	0.27	0.54	0.95	0.98	0.47	0.86	0.77	0.57	0.84		0.59
[K]	- SVM + ICM	0.49	0.84	0.73	0.70	0.75	0.82	0.90	0.87	0.83	0.68	

Values above the diagonal line correspond to the IKONOS-II data set and values below the diagonal line correspond to HYDICE. Numbers in parentheses indicate window sizes (in pixels). Underline values indicate that the difference between the two compared coefficients is statistically significant at a confidence level of 95%.

are taken based on a confidence level of 95%. Values lower than 0.05 are underlined, indicating that the difference between the two compared coefficients is statistically significant.

According to the performed hypothesis tests, it appears that the majority of the analysed results are statistically equivalent. For the classifications performed on the IKONOS-II image, the results that are statistically superior to those obtained using SVM are provided by the CaSVM and SVM + Mode methods with context windows of 5 × 5 or 7 × 7 pixels. The SVM + ICM method also provides results that are statistically superior to those of SVM. It also appears that the CtSVM method provides similar results to SVM. In general, note that when the neighbourhood influence radius is increased, the kappa coefficient values assigned to the CaSVM and SVM + Mode methods are higher. Regarding the HYDICE data set, due to the very close kappa coefficient values and the large deviations, the performed hypothesis tests allow us to conclude that there is no statistical difference, at 95% confidence, between any of the obtained classification results.

Concerning the assessment on edges, the kappa coefficient on edges and upslon coefficient were computed, showing different behaviours, such as follows: (i) for the IKONOS-II image, while the kappa coefficient on edges increases with the neighbourhood radius of SVM + Mode, the upslon coefficient value decreases; (ii) for the HYDICE image,

while SVM is more accurate than CaSVM(3×3) when kappa coefficient on edges is considered, the opposite is observed when epsilon coefficient is used.

Another remark relates to an expected behaviour with respect to SVM + Mode, which tends to produce results around target edges exceedingly smoothed with increasing neighbourhood radius. Consequently, such smoothing implies that the classification of one side of the edge (i.e., internal or external) occurs erroneously. Keeping it in mind, the epsilon coefficient is again consistent with the results unlike the values shown by the kappa coefficient on edges in Figures 3 and 4.

When comparing the classification results obtained from the IKONOS-II image, shown in Figure 5(a), it was possible to observe that CaSVM(5×5) is similar to SVM + ICM as it happens with epsilon coefficient values in Figure 3. Although SVM + Mode(5×5) produces a smoother classification result compared to CaSVM(5×5), SVM + ICM and SVM, the computed epsilon coefficient is smaller.

Analogously, with respect to the results shown in Figure 5(b) from the HYDICE image, the CaSVM(5×5) method is obviously more accurate on edges than SVM + Mode(5×5), as can be seen by the epsilon coefficient values in Figure 4. Moreover, the kappa coefficient on edges indicates that the CaSVM(5×5) and SVM + Mode(5×5) are very similar. Hence, it is possible to conclude that it is convenient to use epsilon coefficient to assess the classification on edges.

Considering IKONOS-II classifications, the results of SVM(5×5), SVM(7×7), SVM + ICM and SVM + Mode(7×7) are statistically superior when compared with SVM. However, observing the epsilon coefficient is possible to conclude that CaSVM result is the best one, followed by SVM + ICM result.

Observing HYDICE results, all classifications are statistically insignificant at 95%, but the epsilon coefficient allows one to conclude that CaSVM has the best result. This conclusion agrees with the visual inspection of Figure 5(b).

CaSVM is more adaptable to the context than CtSVM, so CtSVM has comparatively worse results where the transitions are noisier; this is the case of the irregular edges inside the city portrayed in the IKONOS image. Unlike the results for the IKONOS-II image, CtSVM behaves similarly to SVM, for the HYDICE image.

Regarding the computational cost, 24.7 and 4.7 s were expended by SVM to perform the classifications on the IKONOS-II and HYDICE images, respectively. SVM + Mode and

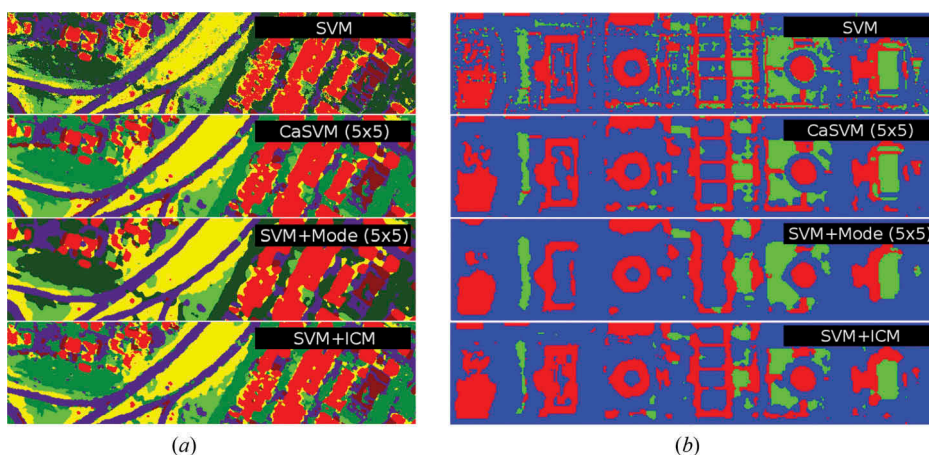


Figure 5. Extract, previously identified in Figures 1(a) and 2(a), of some relevant classification results obtained by the analysed methods in (a) IKONOS-II and (b) HYDICE imagery.

SVM + ICM are approximately twofold more expensive than SVM. On the other hand, CaSVM/CtSVM are 200/61 and 60/8 times more expensive than SVM for classifying the IKONOS-II and HYDICE data sets, respectively.

5. Conclusions

This paper presented an analysis among different contextual versions of SVM, including integrations of SVM with the ICM algorithm and mode filter, denoted as SVM + ICM and SVM + Mode, respectively, as well as the recently reported approaches CaSVM and CtSVM. Kappa coefficient was used to compare the results. As a complementary method to evaluate the performance of the analysed methods, kappa coefficient has been calculated near the transition regions (kappa coefficient on edges) and a new measure called upsilon coefficient was introduced to take into account what is happening on both sides of transition regions.

In general, the accuracy of the analysed contextual methods are similar when the assessment is performed in the usual way through the kappa coefficient as shown in [Figures 3](#) and [4](#) and [Table 2](#). However, CaSVM and SVM + ICM achieved better performance on edges between regions of different classes, as measured by upsilon coefficient, when compared to the other analysed contextual methods as shown in [Figures 3–5](#).

In addition, the computational time of CaSVM is considerably higher than any other evaluated method.

Moreover, in most cases, a direct relationship between the neighbourhood size and the accuracy of the results was found. When increasing the neighbourhood size, more homogeneous classifications are obtained. For SVM + Mode, an increase in the neighbourhood size implies a decrease in the performance on the edges of regions.

Although CaSVM has a higher computational cost, a better balance between homogenized classifications and preservation of edges can be obtained. In face of these facts, the use of CaSVM is more indicated when the accuracy of results is crucial. If computational time is an important factor, SVM + ICM appears to be preferable.

Furthermore, this research demonstrated that usual kappa coefficient alone is not able to perform a more complete assessment of classification results because it is not sensitive for what is happening at transition regions. It was shown that proposed upsilon coefficient is a consistent complementary indicator for choosing, among equivalent classification results, the ones that have better performance in terms of preserving the edge resolution while maintaining the classification quality.

Future work should consider the following: (a) investigate parallel implementations to reduce the computational time of CaSVM, (b) test other multiclass strategies and (c) develop a deviation measure for the upsilon coefficient.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The authors thank FAPESP [grant 2014/14830-8], UNESP/PROPe [grant 2015/857], CAPES and CNPq [grant 307666/2011-5, 401528/2012-0 and 151571/2013-9] for funding this research.

ORCID

R.G. Negri  <http://orcid.org/0000-0002-4808-2362>

L. V. Dutra  <http://orcid.org/0000-0002-7757-039X>

References

- Besag, J. 1986. "On the Statistical Analysis of Dirty Pictures." *Journal of the Royal Statistical Society* B-48: 259–302.
- Boser, B. E., I. M. Guyon, and V. N. Vapnik. 1992. "A Training Algorithm for Optimal Margin Classifiers." In *Proceedings...*, 144–152. Annual Workshop on Computational Learning Theory. Pittsburgh: ACM Press.
- Bovolo, F., and L. Bruzzone. 2005. "A Context-Sensitive Technique Based on Support Vector Machines for Image Classification Lecture Notes in Computer Science." *Pattern Recognition and Machine Intelligence* 3776: 260–265.
- Bovolo, F., L. Bruzzone, and M. Marconcini. 2006. "A Novel Context-Sensitive SVM for Classification of Remote Sensing Images." In *Proceedings...*, 2498–2501. International Geoscience and Remote Sensing Symposium. Denver: IEEE.
- Bruzzone, L., and C. Persello. 2009. "A Novel Context-Sensitive Semisupervised SVM Classifier Robust to Mislabeled Training Samples." *IEEE Transactions on Geoscience and Remote Sensing* 47 (7): 2142–2154. doi:10.1109/TGRS.2008.2011983.
- Cortijo, F. J., and N. P. Blanca. 1998. "Improving Classical Contextual Classifications." *International Journal of Remote Sensing* 19 (8): 1591–1613. doi:10.1080/014311698215379.
- Gurram, P., and H. Kwon. 2013. "Contextual SVM Using Hilbert Space Embedding for Hyperspectral Classification." *IEEE Geoscience and Remote Sensing Letters* 10 (5): 1031–1035. doi:10.1109/LGRS.2012.2227934.
- Hamel, L. 2006. "Visualization of Support Vector Machines with Unsupervised Learning." In *Proceedings...*, 1–8. Symposium on Computational Intelligence in Bioinformatics and Computational Biology. Toronto: IEEE.
- Liu, C., P. Frazier, and L. Kumar. 2007. "Comparative Assessment of the Measures of Thematic Classification Accuracy." *Remote Sensing of Environment* 107 (4): 606–616. doi:10.1016/j.rse.2006.10.010.
- Mountrakis, G., J. Im, and C. Ogole. 2011. "Support Vector Machines in Remote Sensing: A review." *ISPRS Journal of Photogrammetry and Remote Sensing Society* 66 (3): 247–259. doi:10.1016/j.isprsjprs.2010.11.001.
- Negri, R. G., L. V. Dutra, and S. J. S. Sant'Anna. 2012. "Support Vector Machine and Bhattacharyya Kernel Function for Region Based Classification." In *Proceedings...*, 5422–5425. International Geoscience and Remote Sensing Symposium. Munich: IEEE.
- Negri, R. G., L. V. Dutra, and S. J. S. Sant'Anna. 2014. "An Innovative Support Vector Machine Based Method for Contextual Image Classification." *ISPRS Journal of Photogrammetry and Remote Sensing* 87: 241–248. doi:10.1016/j.isprsjprs.2013.11.004.
- Negri, R. G., S. J. S. Sant'Anna, and L. V. Dutra. 2013. "A New Contextual Version of Support Vector Machine Based on Hyperplane Translation." IEEE international geoscience and remote sensing symposium, 3116–3119, Melbourne, July.
- Platt, J. C. 2000. *Probabilistic Outputs for Support Vector Machines and Comparisons to Regularized Likelihood Methods*, 61–74. Cambridge: MIT Press. Advances in Large Margin Classifiers.
- Santos, R. D. C. 2013. "Java Image Processing Cookbook." Acesso em 2 mai. 2013. <http://www.lac.inpe.br/JIPCookbook/>.
- Theodoridis, S., and K. Koutroumbas. 2008. *Pattern Recognition*. 4th ed. San Diego: Academic Press.
- Thorsten, J. 1999. *Making Large-Scale Support Vector Machine Learning Practical*, 169–184. Cambridge: MIT Press.
- Tzotsos, A. 2006. *A Support Vector Machine Approach for Object Based Image Analysis*. Salzburg: ISPRS International Conference on Object-based Image Analysis.