A NEW CONTEXTUAL VERSION OF SUPPORT VECTOR MACHINE BASED ON HYPERPLANE TRANSLATION

Rogério Galante Negri, Sidnei João Siqueira Sant'Anna, Luciano Vieira Dutra

Instituto Nacional de Pesquisas Espaciais - INPE, São José dos Campos - SP, Brazil

ABSTRACT

Support Vector Machine (SVM) is a method widely used for image classification. The original formulation of this method does not incorporate contextual information. This study brings a new perspective regarding contextual SVM. The main idea of the presented proposal consists on translates, individually for each pixel using it contextual information, the separation hyperplane originally designed by SVM. A case study using ALOS PALSAR image shows that the proposed method produces better results than traditional SVM.

***Index Terms*—** Image classification, Support Vector Machine, contextual information, hiperplane translation

1. INTRODUCTION

Image classification is one of the most important applications of pattern recognition in remote sensing. Traditionally, the image classification process has been conducted based only on the pixels features, by the so-called "pixel-based" classifiers. This approach may be unsatisfactory in some cases, e.g., in the classification of images with high spatial and spectral resolutions [1]. This problem has stimulated the development of contextual classifiers, which exploit the spatial relationships among the pixels as an additional source of information.

Introduced by Vladimir Vapnik, SVM is a pattern recognition method that has overcome many systems in a wide variety of applications [2]. However this method is unable to incorporate the contextual information in the classification process. Different proposals to incorporate the contextual information on SVM have been presented in literature. Generally, these incorporations are made using stochastic models [3, 4] or modifying the learning process [5, 6].

This work presents a new perspective on the development of contextual versions of SVM. Different from the aforementioned approaches, statistical techniques or modification on the learning process are not adopted, but displacements on the separation hyperplane according to the pixels contextual information.

Thanks to CAPES, CNPq (Process: 307666/2011-5) and FAPESP (Process: 08/58112-0 and 08/57719-9) agencies for funding.

2. HYPERPLANE TRANSLATION FOR CONTEXTUAL CLASSIFICATION

Formally, a classifier is represented by a function $f : \mathcal{X} \mapsto \Omega$ that assigns elements from the an attribute space $\mathcal{X}$ to a class of $\Omega = \{\omega_1, \omega_2, \dots, \omega_c\}$. Image classification consists on the application of f on the pixels that composes an imagem $\mathcal{I}$, defined on a support $\mathcal{S} \subset \mathbb{N}^2$. Considering the image where is conducted the classification process, $\mathcal{I}(s) = \mathbf{x}$ denotes that $\mathbf{x} \in \mathcal{X}$ has coordinate $s \in \mathcal{S}$ of $\mathcal{I}$. The neighbor positions of s are elements of $\mathcal{V}_\rho(s) = \{t \in \mathcal{S} : 0 \leq md(s,t) \leq \rho\}$, where ρ is called *neighborhood influence radius* and $md(\cdot,\cdot)$ is the *maximum distance*[1].

SVM consists on distinguishing patterns based on a maximum margin hyperplane. A hyperplane is a geometric place where the following function is null:

$$f_{SVM}(\mathbf{x}) = \langle \mathbf{w}, \mathbf{x} \rangle + b, \tag{1}$$

where $\mathbf{w}$ is the orthogonal vector to $f_{SVM}(\mathbf{x}) = 0$, b is a scalar such that $|b|/\|\mathbf{w}\|$ is the distance between the hyperplane and the originof the attribute space, and $\langle \cdot, \cdot \rangle$ represents the inner product operation.

Considering $\mathcal{D} = \{(\mathbf{x}_i, y_i) \in \mathcal{X} \times \mathcal{Y} : i = 1, \dots, m\}$ as training set, with $\mathcal{Y} = \{-1, +1\}$, a set of labels such $\mathbf{x}_i$ is assigned to ω_{c_1} if $y_i = +1$, or to ω_{c_2} if $y_i = -1$, the parameters $\mathbf{w}$ and b that determines the maximum margin hyperplane are obtained from the solution of the following quadratic optimization problem [7]:

$$\begin{array}{l} \max\limits_{\gamma} \left(\sum_{i=1}^{m} \gamma_i - \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \gamma_i \gamma_j y_i y_j \langle \mathbf{x}_i, \mathbf{x}_j \rangle \right) \\ \text{subject to:} \left\{ \begin{array}{l} 0 \leq \gamma_i \leq C, i = 1, \dots, m \\ \sum_{i=1}^{m} \gamma_i y_i = 0 \end{array} \right. \end{array} \tag{2}$$

where γ are Lagrangean multipliers up bounded by C, a real parameter that acts as a penalty to misclassifications. A set of support vector (SV) is defined by pattern $\mathbf{x}_i$ such $\gamma_i \neq 0$. From this optimization, $\mathbf{w} = \sum_{\forall \mathbf{x}_i \in SV} y_i \gamma_i \mathbf{x}_i$ and b is equal to $\frac{1}{\#SV} \left(\sum_{\mathbf{x}_i \in SV} y_i + \sum_{\mathbf{x}_i \in SV} \sum_{\mathbf{x}_j \in SV} \gamma_i \gamma_j y_i y_j \langle \mathbf{x}_i, \mathbf{x}_j \rangle \right)$.

When $f_{SVM}(\mathbf{x})$ is determined, $\mathbf{x}_i$ is assigned to ω_{c_1} if $f_{SVM}(\mathbf{x}_i) \geq 0$, or to ω_{c_2} when $f_{SVM}(\mathbf{x}_i) < 0$. This decision

[1]Lets $s, t \in \mathbb{N}^2$ such that $s = (s_1, s_2)$ and $t = (t_1, t_2)$. The *maximum distance* is defined by $md(s,t) = \max\{|s_1 - t_1|, |s_2 - t_2|\}$.

rule characterizes the SVM as a binary classifier. To apply this method in problems with more than two classes it makes necessary the use of multiclass strategies. One-Against-One (OAO) is typical example of a multiclass strategy. The inner product of (1), as in the quadratic optimization problem (2), can be replaced by kernel functions. Linear, Polynomial and Radial Basis Function are some typical examples of kernel functions [8].

This paper introduces a new version of SVM that consists in translate the hyperplane $f_{SVM}(\mathbf{x}) = 0$ as function of the spatial context of each pixel. Such translation are performed based on the *classification reliability* of the pixels located within a certain neighborhood $\mathcal{V}_\rho(s)$. One way to compute the *classification reliability* according to the hyperplane $f_{SVM}(\mathbf{x}) = 0$ and the separation margin, i.e. $f_{SVM}(\mathbf{x}) = \pm 1$, is given by:

$$r(\mathbf{x}_i) = 1 - |f_{SVM}(h(\mathbf{x}_i))|, \tag{3}$$

where:

$$h(\mathbf{x}_i) = \mathbf{x}_i + \alpha(\mathbf{x}_i) \cdot \frac{1}{\|\mathbf{w}\|} \cdot \frac{\mathbf{w}}{\|\mathbf{w}\|} \cdot \operatorname{sgn}(f_{SVM}(\mathbf{x}_i)), \tag{4}$$

$$\alpha(\mathbf{x}_i) = \left(\frac{1}{|f_{SVM}(\mathbf{x}_i)|} - |f_{SVM}(\mathbf{x}_i)| \right). \tag{5}$$

The equation (3) is used as a *reability metric* for $\mathbf{x}_i$, which return values in $] -\infty, 1\,]$. The equation (4) performs a re-projection of each pattern $\mathbf{x}_i$ taking into account its location relative to $f_{SVM}(\mathbf{x}) = 0$. Function (5) is defined as *re-projection factor*. In the re-projection process patterns near to $f_{SVM}(\mathbf{x}) = 0$ are repositioned far from this hyperplane, unlike the patterns initially far from $f_{SVM}(\mathbf{x}) = 0$.

According to the *reliability* of the patterns located in the neighborhood of $\mathbf{x}_i$, a translations in the hyperplane $f_{SVM}(\mathbf{x}) = 0$ is performed by adding the following term in (1):

$$\Delta(\mathbf{x}_i) = \lambda \cdot \frac{(G_p(s) - G_n(s))}{\|\mathbf{w}\|};\ \mathcal{I}(s) = \mathbf{x}_i, \tag{6}$$

where $\lambda \in \mathbb{R}_+$ is a weigth parameter to the context influence while $G_p(s)$ e $G_n(s)$ are defined by:

$$G_p(s) = \sum_{\mathbf{x}_i \in T_p} r(\mathbf{x}_i);\ \ T_p = \{\mathbf{x}_i : \mathcal{I}(t) = \mathbf{x}_i; t \in \mathcal{V}_\rho(s); \\ 0 \leq f_{SVM}(h(\mathbf{x}_i)) \leq 1\}, \tag{7}$$

$$G_n(s) = \sum_{\mathbf{x}_i \in S_n} r(\mathbf{x}_i);\ \ T_n = \{\mathbf{x}_i : \mathcal{I}(t) = \mathbf{x}_i; t \in \mathcal{V}_\rho(s); \\ -1 \leq f_{SVM}(h(\mathbf{x}_i)) < 0\}. \tag{8}$$

Adding $\Delta(\mathbf{x}_i)$ in (1) is obtained the following modification of $f_{SVM}(\mathbf{x})$:

$$f_{Local}(\mathbf{x}) = \langle \mathbf{w} \cdot \mathbf{x} \rangle + (b + \Delta(\mathbf{x}_i)). \tag{9}$$

The function (9) is redefined for each pixel of the image and is applied only to produce the contextual classification of this pixel. The term $\Delta(\mathbf{x}_i)$ is the amount which the original hyperplane is moved. According to the formulation shown, the hyperplane translation is determined by the *classification reliability* of the patterns that belonging to different classes inside a neighborhood. Because of intrinsic characteristics, the proposed method was named "Competitive Translative Support Vector Machine" (ctSVM).

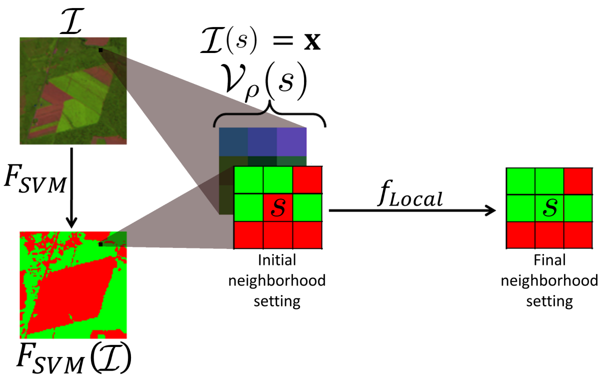
A framework of the contextualization process of the CtSVM method is show in Figure 1(a). A hypothetical example is illustrated in Figure 1(b), where is show a neighborhood behavior of ★ and the re-projection process of its elements ($h(\mathbf{x}_i) = \mathbf{x}'_i$) based on $f_{SVM}(\mathbf{x}) = 0$. From this re-projection is computed $\Delta(\mathbf{x}_i)$ term and the translated hyperplane $f_{Local}(\mathbf{x}) = 0$, as show in Figure 1(c). With this new hyperplane the classification of ★ changes. It is worth note that the local hyperplane does not guarantee maximum separation margin.

The formalization presented in this section deal with contextual classification of binary cases. The extension of ctSVM to multiclass problems is linked to the usage of multiclass strategies that decompose the original problem into binaries sub problems. After this decomposition, each binary sub problem is individually treated. After the contextualization of each binary problem the results are analyzed according to the multiclass classification rule initially adopted to produce the contextual multiclass classification.

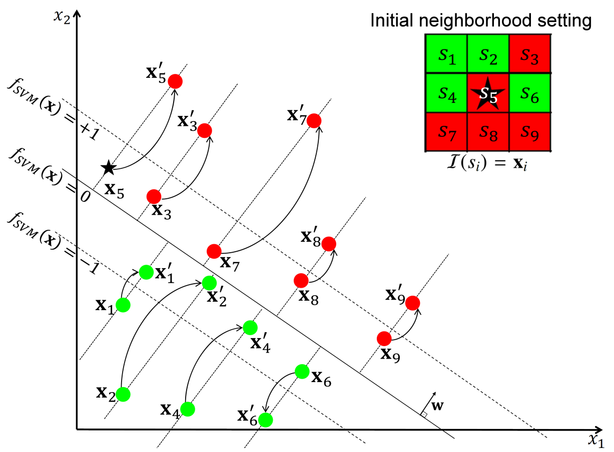
3. EXPERIMENTS AND RESULTS

This section presents a case study where a remote sensing image is classified by the methods SVM and ctSVM. For this purpose was used an image fragment of the ALOS PALSAR sensor, with polarizations HH, HV and VV in amplitude, acquired on March 13, 2009. This image refers to a region of the Tapajós National Forest, Pará State, Brazil, where were identified the following land cover classes: Primary Forest, Pasture, Bare Soil and Agriculture. The mentioned image fragment and the land cover samples used to train the methods and validate the classification results are illustrated in Figure 2.

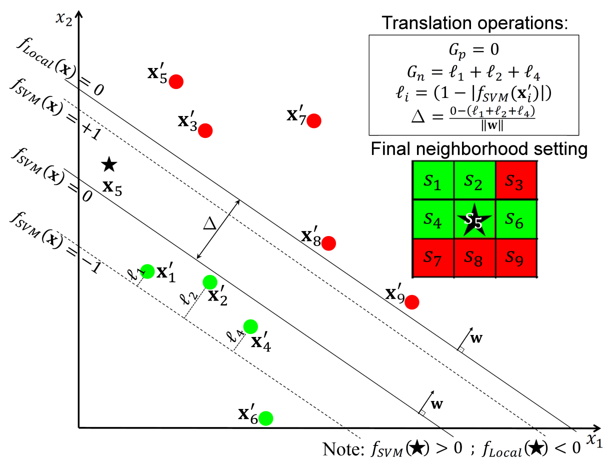
In the classification process, the Linear kernel function and penalty equals to 100 were used. The choice of this kernel function and the penalty value were based on an Grid Search procedure. To deal with the multiclasse problem the OAO strategy was adopted. To set the parameter λ a binary search procedure was conducted in order to select the value which maximizes the classification accuracy on the training set. The *neighborhood influence radius* (ρ), by the method ctSVM, equals to 1, 2 and 3, which provides contextualization windows of 3×3, 5×5 and 7×7 pixels, respectively, were employed. The accuracy of the classification results were quantified using Tau coefficient and the percentage



(a) Steps of local contextualization.



(b) The initial configuration and re-projection of the neighborhood pixels of ★ as a function of $f_{SVM}(\mathbf{x}) = 0$.

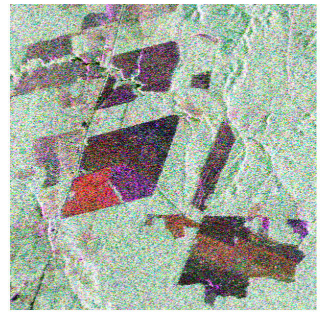


(c) Local hyperplane definition for context classification of ★.

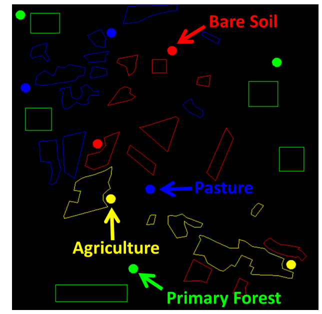
Fig. 1. A framework and an example of contextual classification with ctSVM.

of individual class accuracy [9]. The validation samples illustrated in Figure 2(b) were subsampled.

Figure 3 illustrates the classification results obtained.



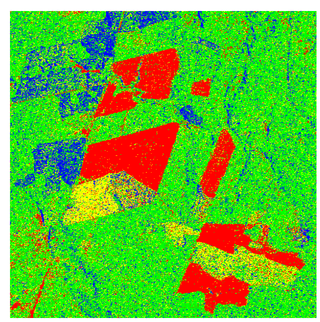
(a) ALOS/PALSAR image in R(HH)G(HV)B(VV) color composition



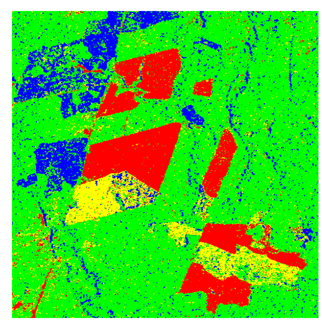
(b) Training samples (solid circles) and validation samples (empty polygons)

Fig. 2. Image and samples adopted in the study case.

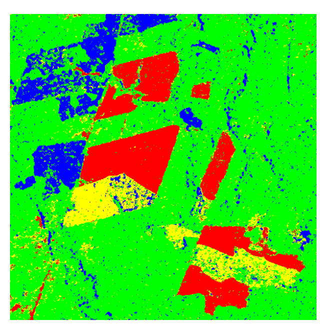
With the analysis of the classification results the improvement produced by ctSVM, compared with SVM, was significant. The elimination of isolated pixels were more notable in Primary Forest areas. A better definition of the classes of Pasture and Agriculture were achieved by the method ctSVM when is adopted contextualization windows of 5×5 and 7×7 pixels, as shows the graph in Figure 4(a). Regarding the accuracy of the results the Figure 4(b) shows that Tau coefficient of the SVM classification result has improved by 20.5%, 25.6% and 27.8% when the ctSVM method is adopted with contextualization windows of 3×3, 5×5, and 7×7.



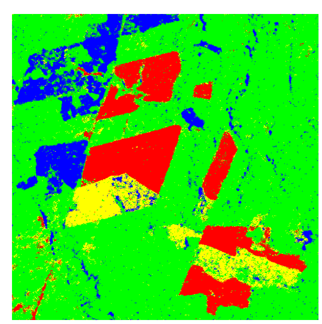
(a) SVM



(b) ctSVM (3×3)

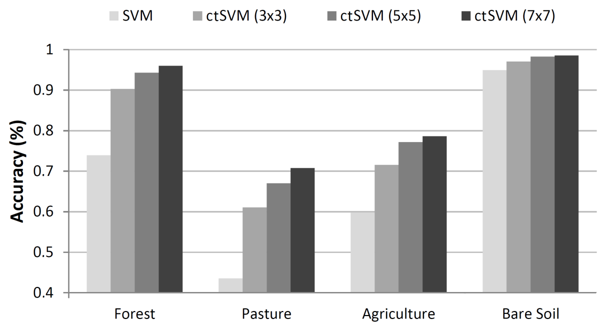


(c) ctSVM (5×5)

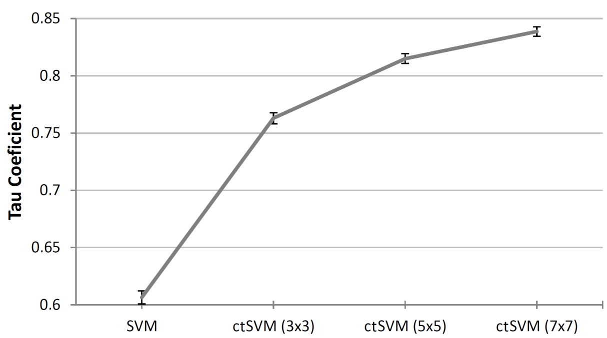


(d) ctSVM (7×7)

Fig. 3. Classification results obtained by the analyzed methods.



(a) Individual class accuracy



(b) Tau coefficient values

Fig. 4. Graphic comparison between the accuracy results.

4. CONCLUSIONS

This work presented a new contextual classification method based on concepts of SVM. A comparative analysis with the SVM method was performed. The results show superiority of the proposed method. As future work, the development of new study cases, analysis on the preservation of edges and punctual regions and comparisons with other contextual methods should be considered.

5. REFERENCES

[1] O. Besbes, N Boujemaa, and Z Belhadj, "Contextual Classification of High-Resolution Satellite Images," in *IEEE Symposium on Computational Intelligence for Image Processing*, Waterloo, Canada, 2009, pp. 41–47.

[2] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*, Cambridge University Press, Cambridge, UK, 2000.

[3] F. Bovolo and L. Bruzzone, "A context-sensitive technique based on support vector machines for image classification," in *Patter Recognition and Machine Intelligence*, 2005, pp. 260–265.

[4] S. Li, B. Zhang, D. Chen, L. Gao, and M. Peng, "Adaptive support vector machine and markov random field model for classifying hyperspectral imagery," *Journal of Applied Remote Sensing*, vol. 5, no. 1, pp. 053538–053538, 2011.

[5] F. Bovolo, L. Bruzzone, and M. Marconcini, "A novel context-sensitive svm for classification of remote sensing images," in *IEEE International Geoscience and Remote Sensing Symposium*, 2006, pp. 2498–2501.

[6] P. Gurram and H. Kwon, "Contextual svm using hilbert space embedding for hyperspectral classification," *IEEE Geoscience and Remote Sensing Letters*, , no. 99, pp. 1–5, 2013.

[7] S. Theodoridis and K. Koutroumbas, *Pattern Recognition, Third Edition*, Academic Press, Inc., Orlando, FL, USA, 2008.

[8] A. R. Webb, *Statistical Pattern Recognition, 2nd Edition*, John Wiley & Sons, 2002.

[9] R. G. Congalton and K. Green, *Assessing the accuracy of remotely sensed data:* principles and practices, Lewis Publisher, New York, 1999.