

Máquina de Vetores Suporte Adaptativa ao Contexto

Rogério Galante Negri¹, Luciano Vieira Dutra², Sidnei João Siqueira Sant'Anna²

¹Programa de Doutorado em Computação Aplicada – CAP
Instituto Nacional de Pesquisas Espaciais – INPE

²Divisão de Processamento de Imagens – DPI
Instituto Nacional de Pesquisas Espaciais – INPE

{rogerio, dutra, sidnei}@dpi.inpe.br

Abstract. *Support Vector Machine is a Pattern Recognition method that has overcome many systems in different applications. The original formulation of this method does not incorporate contextual information in image classification applications. This work presents a new approach on the development of contextual Support Vector Machine. A study case involving a comparison with other classification methods shows superiority of the new proposal.*

Resumo. *Máquina de Vetores Suporte é um método de Reconhecimento de Padrões que tem superado muitos sistemas em diferentes aplicações. A formulação original deste método não provê a incorporação da informação contextual nas aplicações em classificação de imagens. Este trabalho traz uma nova vertente sobre Máquina de Vetores Suporte contextual. Um estudo de caso envolvendo a comparação com outros métodos mostra superioridade da nova proposta.*

Palavras-chave: *Classificação contextual, Máquina de Vetores Suporte.*

1. Introdução

Reconhecimento de Padrões pode ser entendido como uma área da Ciência da Computação onde são desenvolvidos métodos e técnicas voltadas a emulação de tarefas inteligentes, capazes de auxiliar a realização de atividades manualmente inviáveis. O vasto conjunto de ferramentas que compõe o Reconhecimento de Padrões fornece suporte ao desenvolvimento de inúmeras aplicações em diversas áreas de atividade, que variam desde o controle de qualidade industrial à medicina, navegação de robôs, aplicações militares, exploração geofísica e sensoriamento remoto.

Uma das mais importantes aplicações de Reconhecimento de Padrões em imagens de sensoriamento remoto da Terra é a classificação de imagens [Khedam et al. 2003]. A classificação de imagens de sensoriamento remoto consiste no procedimento de extração de informações dos *pixels* para identificação automática dos alvos, dentre um número definido de classes temáticas. Tradicionalmente, o processo de classificação tem sido conduzido a partir da informação espectral dos *pixels*, pelos denominados “classificadores pontuais”. Esta abordagem pode ser insatisfatória em alguns casos, por exemplo, na classificação de imagens com altas resoluções espaciais e espectrais, onde a heterogeneidade das informações é maior. Este problema tem estimulado o desenvolvimento dos “classificadores contextuais”, os quais exploram as relações espaciais entre os pixels como fonte adicional de informação.

Introduzido por Vladimir Vapnik, a Máquina de Vetores Suporte (*Support Vector Machine* - SVM) é um método de Reconhecimento de Padrões que em pouco tempo têm superado muitos sistemas em uma ampla variedade de aplicações [Cristianini and Shawe-Taylor 2000]. No entanto, o método SVM é incapaz de incorporar a informação contextual ao processo de classificação. Em [Bovolo and Bruzzone 2005] é proposta uma maneira de produzir classificações contextuais derivadas do método SVM a partir da integração com modelos estocásticos. Outra maneira de derivar classificações contextuais do método SVM consiste no pós-processamento por técnicas de suavização. Um tipo de suavização resume-se a aplicação de filtros, usualmente o filtro de maioria (moda), sobre os rótulos da imagem classificada.

Este trabalho traz uma nova vertente sobre o desenvolvimento de SVM contextual. Diferentes das propostas mencionadas, não são adotadas técnicas estatísticas e de suavização, mas sim um processo de modelagem local da informação contextual, formalizado sobre os conceitos originais do método SVM.

2. Classificação de Imagem

2.1. Visão formal do processo de classificação de imagem

Formalmente, um classificador é representado por uma função $f : \mathcal{X} \mapsto \Omega$, que associa elementos do espaço de atributos \mathcal{X} a uma dada classe de $\Omega = \{\omega_1, \omega_2, \dots, \omega_c\}$. A classificação de imagens consiste na aplicação de f sobre os *pixels* que compõe uma imagem \mathcal{I} , definida sobre um reticulado $\mathcal{S} \subset \mathbb{N}^2$. Com relação a imagem em que é conduzido o processo de classificação, $\mathcal{I}(s) = \mathbf{x}$ denota que $\mathbf{x} \in \mathcal{X}$ ocupa a posição $s \in \mathcal{S}$ em \mathcal{I} . Ainda, as posições ocupadas pelos vizinhos de s é representado pelo conjunto $\mathcal{V}_\rho(s) = \{t \in \mathcal{S} : 0 \leq dm(s, t) \leq \rho\}$, sendo ρ denominado por *raio de influência da vizinhança* e $dm(\cdot, \cdot)$ a *distância do máximo*.

A estimação de f é definida de acordo com o paradigma de aprendizado do método de classificação. Os classificadores de aprendizado supervisionado realizam a estimação de f a partir de informações extraídas do conjunto de treinamento $\mathcal{D} = \{(\mathbf{x}_i, y_i) \in \mathcal{X} \times \mathcal{Y} : i = 1, \dots, m\}$, sendo $\mathcal{Y} = \{1, 2, \dots, c\}$ denominado por conjunto dos indicadores de classe. Nesta notação, (\mathbf{x}_i, y_i) indica que o padrão \mathbf{x}_i esta associado a classe ω_j quando $y_i = j$.

2.2. Máquina de Vetores Suporte

O método SVM consiste em distinguir padrões a partir de hiperplanos cuja margem de separação é máxima. Um hiperplano de separação equivale ao lugar geométrico onde a seguinte função discriminante torna-se nula:

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

sendo \mathbf{w} o vetor ortogonal ao hiperplano de separação $f_{SVM}(\mathbf{x}) = 0$ e b é um escalar real tal que $|b|/\|\mathbf{w}\|$ representa a distância entre o hiperplano e a origem do espaço de atributos. Os parâmetros \mathbf{w} e b que determinam o hiperplano de margem máxima são obtidos com a resolução do seguinte problema de otimização quadrática, com base nos padrões do conjunto de treinamento \mathcal{D} [Theodoridis and Koutroumbas 2008]:

$$\begin{aligned} & \max_{\gamma} \sum_{i=1}^m \gamma_i - \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m \gamma_i \gamma_j y_i y_j \langle \mathbf{x}_i, \mathbf{x}_j \rangle \\ & \text{sujeito a: } \begin{cases} 0 \leq \gamma_i \leq C, i = 1, \dots, m \\ \sum_{i=1}^m \gamma_i y_i = 0 \end{cases} \end{aligned} \quad (2)$$

onde γ_i são multiplicadores de Lagrange, $\mathcal{Y} = \{-1, +1\}$, e o parâmetro C , introduzido para o tratamento de casos não-linearmente separáveis, atua como um limitante superior para γ_i . Sendo $SV = \{\mathbf{x}_i : \gamma_i \neq 0; i = 1, \dots, l\}$, o conjunto de vetores suporte, o parâmetro \mathbf{w} é computado por $\sum_{\mathbf{x}_i \in SV} \gamma_i y_i \mathbf{x}_i$, enquanto b equivale a $1 - \mathbf{w}^T \mathbf{x}_i$ para qualquer $\mathbf{x}_i \in SV$ tal que $f(\mathbf{x}_i) = 1$.

Com a determinação de $f_{SVM}(\mathbf{x})$, um dado *pixel* \mathbf{x}_i pertence à classe ω_1 desde que $f_{SVM}(\mathbf{x}_i) \geq 0$, ou à classe ω_2 quando $f_{SVM}(\mathbf{x}_i) < 0$. Cabe ressaltar que, a medida que $|f_{SVM}(\mathbf{x}_i)|$ aumenta, maior é o *nível de confiança* na classificação de \mathbf{x}_i , e que a construção do hiperplano $f_{SVM}(\mathbf{x}) = 0$ é baseada em *pixels* de baixa confiança. Nota-se ainda que, o método SVM é capaz de realizar a classificação de padrões entre apenas duas classes. A aplicação deste método em problemas que abrangem mais que duas classes faz necessário o emprego de estratégias multiclasse. Um exemplo típico é a estratégia multiclasse Um-Contra-Um (*One-Against-One* - OAO) [Webb 2002].

O produto interno $\langle \mathbf{x}_i, \mathbf{x}_j \rangle$ presente em (2) pode ser substituído por funções simétricas $k(\mathbf{x}_i, \mathbf{x}_j)$, denominadas funções *kernel*. Implicitamente, estas funções computam o produto interno entre os padrões mapeados em um espaço de maior dimensão, isto é, $k(\mathbf{x}_i, \mathbf{x}_j) = \langle \phi(\mathbf{x}_i), \phi(\mathbf{x}_j) \rangle$. Geralmente a função $\phi(\cdot)$, responsável pelo mapeamento dos padrões, é desconhecida. Alguns exemplos típicos de *kernel*, são as funções Linear, $k(\mathbf{x}_i, \mathbf{x}_j) = \langle \mathbf{x}_i, \mathbf{x}_j \rangle$, Polinomial, $k(\mathbf{x}_i, \mathbf{x}_j) = (\langle \mathbf{x}_i, \mathbf{x}_j \rangle + 1)^q$, e a Função de Base Radial (RBF), $k(\mathbf{x}_i, \mathbf{x}_j) = e^{-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{\sigma}}$, sendo $q \in \mathbb{N}$ e $\sigma \in \mathbb{R}$. No caso do *kernel* Polinomial, é possível determinar analiticamente $\phi(\cdot)$.

3. Uma nova proposta contextual de Máquina de Vetores Suporte

Nesta seção é introduzido um novo método de classificação contextual de imagens baseado em conceitos de SVM. Este método está fundamentado nas seguintes hipóteses:

- **Hipótese do hiperplano local:** a modelagem de um contexto local pode ser realizado por hiperplanos de separação.
- **Hipótese da influência do contexto:** a influência exercida pelos *pixels* em uma dada vizinhança está relacionada ao *nível de confiança* com que estes *pixels* são classificados.

Baseado nestas hipóteses, o método proposto consiste em incorporar a informação contextual em uma classificação inicial, gerada pelo método SVM, a partir de uma modelagem local do contexto de cada *pixel*. A modelagem de cada *pixel* é realizada por um hiperplano de separação, construído a partir das informações da vizinhança e sob a influência do nível de confiança com que os *pixels* desta vizinhança foram classificados inicialmente. Este hiperplano é utilizado na reclassificação do *pixel* que define a vizinhança em questão.

Formalmente, seja \mathcal{I} uma imagem inicialmente classificada por um hiperplano de separação $f_{SVM}(\mathbf{x}) = 0$, obtido pelo método SVM. Considerando $\bar{\mathbf{x}}$ um *pixel* de \mathcal{I} tal que $\mathcal{I}(s) = \bar{\mathbf{x}}$, determina-se $\mathcal{D}_s = \{(\mathbf{x}_t, y_t) \in \mathcal{X} \times \mathcal{Y} : t \in \mathcal{V}_\rho(s)\}$, um conjunto formado pelos pares de informação sobre os atributos e identificadores de classe, atribuídos por $f_{SVM}(\mathbf{x})$, localizados na vizinhança de $\bar{\mathbf{x}}$. A partir de \mathcal{D}_s tem-se como objetivo definir um novo hiperplano de separação, que modela o comportamento do contexto de $\bar{\mathbf{x}}$, e que aplicado sobre o mesmo, pode acarretar na alteração de sua classificação inicial.

No entanto, ao recapitular a discussão realizada ao fim da Subseção 2.2, sobre o nível de confiança dos vetores que definem o hiperplano de separação, nota-se que a **hipótese do hiperplano local** é conflitante a **hipótese da influência do contexto**, pois a modelagem dos contextos locais será realizada com forte influência de *pixels* com baixa confiança, segundo sua classificação inicial. Este comportamento introduz o conceito de **reprojeção**, que consiste em inverter, com relação aos limites da margem de separação, a posição dos *pixels* no espaço de atributos. Com isso, os *pixels* classificados inicialmente com alto nível de confiança são reposicionados (reprojetados) para posições do espaço de atributos localizadas próximas ao hiperplano $f_{SVM}(\mathbf{x}) = 0$. De forma análoga, os *pixels* classificados com baixa confiança, inicialmente localizados próximos de $f_{SVM}(\mathbf{x}) = 0$, são reprojetados para posições mais distantes deste hiperplano. A reprojeção é realizada pela seguinte função:

$$h(\mathbf{x}_i) = \mathbf{x}_i + \frac{\alpha(\mathbf{x}_i)}{\|\mathbf{w}\|} \cdot \frac{\mathbf{w}}{\|\mathbf{w}\|} \cdot \text{sgn}(f_{SVM}(\mathbf{x}_i)) \quad (3)$$

onde $\alpha(\mathbf{x}_i) = \frac{1}{|f_{SVM}(\mathbf{x}_i)|} - |f_{SVM}(\mathbf{x}_i)|$ é denominado fator de reprojeção. Em (3) o membro $\frac{\alpha(\mathbf{x}_i)}{\|\mathbf{w}\|}$ representa a quantidade, em termos de \mathbf{w} , do deslocamento de \mathbf{x}_i no espaço de atributos, $\frac{\mathbf{w}}{\|\mathbf{w}\|}$ é o vetor unitário do deslocamento realizado e $\text{sgn}(f_{SVM}(\mathbf{x}_i))$ é a função sinal, necessária para generalizar a reprojeção independente do valor $f_{SVM}(\mathbf{x}_i)$ ser positivo ou negativo.

Nota-se em (3) que para os *pixels* localizados nos limites da margem de separação (i.e. $|f_{SVM}(\mathbf{x})| = 1$), não ocorre reprojeção. Com $f_{SVM}(\mathbf{x}_i) \rightarrow 0$, o nível de confiança da classificação inicial de \mathbf{x}_i diminui, e logo, \mathbf{x}_i é reprojetado para distante do hiperplano $f_{SVM}(\mathbf{x}) = 0$. Um comportamento inverso acontece quando $|f_{SVM}(\mathbf{x}_i)| \rightarrow \infty$.

A aplicação de $h(\cdot)$ sobre a primeira coordenada dos elementos de \mathcal{D}_s resulta em um novo conjunto, denotado por \mathcal{H}_s . Para distinção de notação entre os elementos de \mathcal{D}_s e \mathcal{H}_s , denota-se $h(\mathbf{x}_i) = \mathbf{x}'_i$. Partindo das informações de \mathcal{H}_s , é possível definir um hiperplano local que modela o contexto de $\bar{\mathbf{x}}$ sob influência dos *pixels* de maior confiança. No entanto, de acordo com a **hipótese da influência do contexto**, os níveis de confiança das classificações na vizinhança devem influenciar na determinação deste hiperplano local. Para isso, é introduzido o **Modelo Repulsivo**. Este modelo, inspirado no fenômeno da repulsão magnética, consiste em utilizar a distância dos *pixels* reprojetados, com relação ao limite da margem $|f_{SVM}(\mathbf{x})| = +1$ da respectiva classe, como quantidade de força. Esta força é aplicada na repulsão dos *pixels* reprojetados, da outra classe, e *vice-versa*. Tal modelagem é dada pela seguinte função:

$$r(\mathbf{x}'_i) = \mathbf{x}'_i + \lambda \cdot \frac{\mathbf{w}}{\|\mathbf{w}\|} \cdot \frac{F_u}{E_v} \cdot \frac{1}{\|\mathbf{w}\|} \cdot \text{sgn}(f_{SVM}(\mathbf{x}'_i)) \quad (4)$$

sendo:

$$E_v = \min \{1, \#S_v\} \quad (5)$$

$$F_u = \sum_{\forall \mathbf{x}'_i \in S_u} (1 - |f_{SVM}(\mathbf{x}'_i)|) \quad (6)$$

$$S_p = \{(\mathbf{x}'_j, y_j) \in \mathcal{H}_s : 0 \leq f_{SVM}(\mathbf{x}'_j) \leq 1\} \quad (7)$$

$$S_n = \{(\mathbf{x}'_j, y_j) \in \mathcal{H}_s : -1 \leq f_{SVM}(\mathbf{x}'_j) < 0\} \quad (8)$$

tal que $(u, v) = (n, p)$ se $f_{SVM}(\mathbf{x}'_i) \geq 0$ ou $(u, v) = (p, n)$ se $f_{SVM}(\mathbf{x}'_i) < 0$, $\#$ é o operador de cardinalidade, e λ é um parâmetro que pondera a influência do contexto. Com a aplicação de $r(\cdot)$ sobre a primeira coordenada dos elementos de \mathcal{H}_s é produzido o conjunto \mathcal{R}_s . Novamente, para distinção entre os elementos de \mathcal{H}_s e \mathcal{R}_s , denota-se $r(\mathbf{x}'_i) = \mathbf{x}''_i$. Baseado nas informações de \mathcal{R}_s é definido o hiperplano local $f_{Local}(\mathbf{x}) = 0$, que modela o contexto de $\bar{\mathbf{x}}$, respeitando a **hipótese da influência do contexto**. Diante a característica de redefinir hiperplanos de separação localmente, em função do comportamento da vizinhança, o método proposto é denominado por Máquina de Vetores Suporte Adaptativa ao Contexto (*Contextual Adaptive Support Vector Machine* - CaSVM).

Com objetivo de complementar o entendimento do método introduzido nesta seção, a Figura 1 ilustra os diferentes processos que contemplam a modelagem do contexto local, admitindo por simplicidade, $\lambda = 1$. Inicialmente é representado o processo de **reprojeção** (passo 1) realizado sobre os elementos da vizinhança de um dado *pixel*, destacado pelo símbolo \blacktriangle . A posição inicial destes *pixels* no espaço de atributos é obtida da imagem, já os rótulos (vermelho para ω_1 e verde para ω_2) são provenientes de uma classificação inicial em função do hiperplano $f_{SVM}(\mathbf{x}) = 0$, definido pelo método SVM. Em seguida é aplicado o **Modelo Repulsivo** (passo 2). Neste caso, apenas a classe ω_2 possui elementos classificados inicialmente com alta confiança na vizinhança observada. Estes elementos devem influenciar a reclassificação de \blacktriangle . Nota-se também que, uma vez que a classe ω_1 não possui nesta vizinhança classificações de alta confiança, tornando $F_p = 0$, a mesma não exerce “repulsão” aos elementos da classe ω_2 . Por fim, o hiperplano $f_{Local}(\mathbf{x}) = 0$, construído por *pixels* “reprojetados” e “repelidos”, e que modela o comportamento do contexto de \blacktriangle , provocando a alteração de sua classificação inicial (passo 3).

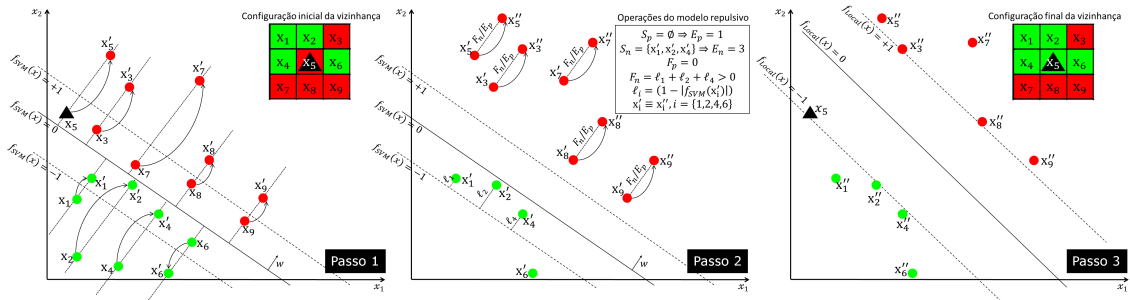


Figura 1. Exemplo hipotético sobre o processo de modelagem do contexto local.

3.1. O processo de classificação binária e sua extensão multiclasse

Na seção anterior foram introduzidos formalmente os principais conceitos e hipóteses do método de classificação contextual proposto. O uso desta formalização para contextualizar uma dada classificação consiste na contextualização de cada um de seus *pixels*, em função das respectivas vizinhanças.

Formalmente, para todo s , tal que $\mathcal{I}(s) = \bar{x}$, é definido o conjunto \mathcal{R}_s a partir de \mathcal{I} e $f_{SVM}(\mathbf{x}) = 0$ com relação aos elementos da vizinhança de s . Este conjunto é determinado pela aplicação de (3) e (4) nos conjuntos \mathcal{D}_s e \mathcal{H}_s , respectivamente. As informações de \mathcal{R}_s são utilizadas na estimação da função de classificação $f_{Local}(\mathbf{x})$, de acordo com as formalizações apresentadas na Subseção 2.2. Determinando $\mathcal{J}(s) = f_{Local}(\bar{x})$, após a contextualização de todos os *pixels* de \mathcal{I} , obtém-se em \mathcal{J} os valores de funções de classificação, modeladas ao contexto, em cada posição de \mathcal{S} . O procedimento descrito é esquematizado no diagrama da Figura 2.

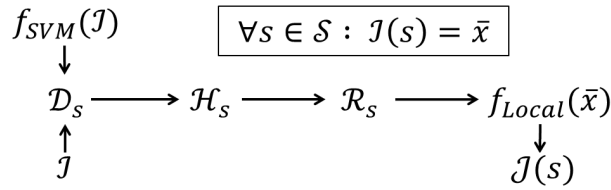


Figura 2. Etapas da classificação contextual binária segundo o método CaSVM.

Esta estrutura discutida formaliza os casos de contextualização de problemas de classificação binária. A extensão do método proposto aos problemas multiclasse está vinculada ao uso de estratégias multiclasse que decompõem o problema original em subproblemas binários. Após tal decomposição, cada subproblema binário é tratado individualmente. Com a contextualização de cada problema binário, os respectivos resultados são analisados de acordo com a regra de classificação multiclasse, proveniente da estratégia adotada inicialmente, para que então seja obtida a classificação contextual multiclasse.

4. Experimentos e Resultados

Esta seção apresenta um estudo de caso sobre a classificação multiclasse de uma imagem real de sensoriamento remoto a partir dos métodos CaSVM e SVM. Ainda, são realizadas comparações com outras duas formas de contextualizar classificações SVM. A primeira consiste na aplicação do filtro da maioria, enquanto a outra forma refere-se a proposta de [Bovolo and Bruzzone 2005], que realiza a contextualização da classificação SVM a partir da integração com o método *Iterated Conditional Modes* (ICM) [Besag 1993]. Os métodos mencionados são denotados respectivamente por SVM+Moda e SVM+ICM.

Um segmento de imagem do sensor ALOS/PALSAR, composta pelas polarizações HH, HV e VV em amplitude, adquirida em 13 de Março de 2009, foi utilizado nesse estudo. Esta imagem referente a uma região da Floresta Nacional do Tapajós - PA, onde foram identificadas quatro classes de cobertura da terra: Floresta, Pastagem, Agricultura e Solo Exposto. O segmento de imagem mencionado e amostras usadas para treinamento e validação dos resultados são ilustrados na Figura 3.

No processo de classificação, foi empregada a função *kernel* Polinomial de grau 1 (i.e. $q = 1$) e penalidade (C) igual a 100. A escolha da função *kernel* e dos respectivos parâmetros foi baseado em um procedimento de busca exaustiva (Grid Search). Adotando $q = 1$ a função *kernel* Polinomial torna-se equivalente à Linear. Para lidar com o problema de classificação multiclasse foi adotada a estratégia OAA. Com relação ao parâmetro λ , foi adotado um procedimento de busca binária visando selecionar o valor que maximize a acurácia na classificação dos dados de treinamento. Nas classificações CaSVM e

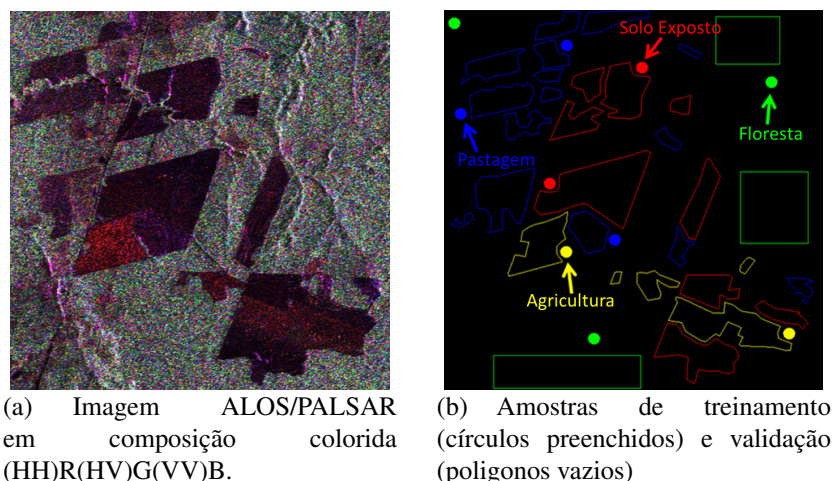


Figura 3. Imagem e amostras usadas no estudo de caso realizado.

Tabela 1. Medidas de acurácia obtidas pelos métodos analisados.

	SVM	CaSVM	SVM+Moda	SVM+ICM
Coefficiente τ	0,029	0,738	0,700	0,691
Desvio padrão do τ	0,002	0,003	0,003	0,002
Tempo de processamento (seg.)	66,93	7076,74	126,44	184,06

SVM+Moda foi adotado $\rho = 2$, enquanto para SVM+ICM foi usado $\rho = 1$, que por sua vez proporcionam janelas de contexto com dimensões 5×5 e 3×3 , respectivamente. Ainda, pelo caráter iterativo do método ICM, admitiu-se que sua convergência acontece após oito iterações ou quando os resultados de duas iterações apresentam menos de 5% de diferença. A acurácia dos resultados de classificação foram quantificadas utilizando o coeficiente de concordância τ [Congalton and Green 1999]. A Figura 4 ilustra os diferentes resultados de classificação obtidos, enquanto os índices de acurácia e o tempo de processamento são apresentados na Tabela 1.

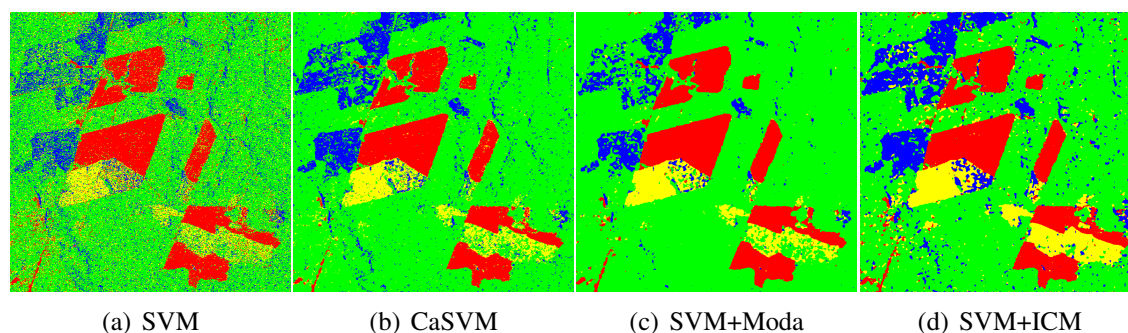


Figura 4. Resultados de classificação obtidos pelos métodos analisados.

Com a análise dos resultados de classificação, é notável a melhora produzida pelo método CaSVM em comparação ao SVM. A eliminação de *pixels* isolados, característica dos “métodos contextuais”, é mais perceptível nas regiões de Floresta. Embora os métodos SVM+Moda e SVM+ICM tenham propiciado a redução de *pixels* isolados, é perceptível a degradação causada nas bordas, principalmente nas regiões de Agricultura. Outro com-

portamento negativo destes métodos foi a redução das áreas de Agricultura e Pastagem, em função da baixa acurácia nestas regiões, na classificação inicial (Figura 4(a)). No entanto, o método CaSVM foi capaz de proporcionar um resultado de classificação contextual com preservação dos detalhes de borda e sem a eliminação de regiões cuja classificação apresenta baixa acurácia.

5. Conclusões

Neste trabalho foi apresentado um novo método de classificação contextual fundamentado em conceitos de SVM. Foi conduzida uma análise comparativa com o método SVM e outras propostas contextuais, também baseadas em SVM. Os resultados obtidos mostram superioridade do método proposto. É ressaltada não só a superioridade no coeficiente de acurácia calculado, mas também com relação a preservação nas bordas e das áreas de baixo contraste.

Como perspectivas para trabalhos futuros, a redução do custo computacional e a comparações com outros métodos contextuais devem ser consideradas. Ainda, devem ser refletidas sobre possíveis simplificações no **Modelo Repulsivo** ou mesmo o desenvolvimento de modelos concorrentes.

Referências

- Besag, J. (1993). Statistical analysis of dirty pictures. *Journal of Applied Statistics*, 20(5):63–87.
- Bovolo, F. and Bruzzone, L. (2005). A context-sensitive technique based on support vector machines for image classification. In *PReMI*, pages 260–265.
- Congalton, R. G. and Green, K. (1999). *Assessing the accuracy of remotely sensed data: principles and practices*. Lewis Publisher, New York.
- Cristianini, N. and Shawe-Taylor, J. (2000). *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge University Press, Cambridge, UK.
- Khedam, R., Belhadj-Aissa, A., and Ranchin, T. (2003). Study of icm parameters influence on images satellite contextual classification. *Geoinformation for European-wide Integration*, pages 79–85.
- Theodoridis, S. and Koutroumbas, K. (2008). *Pattern Recognition, Third Edition*. Academic Press, Inc., Orlando, FL, USA.
- Webb, A. R. (2002). *Statistical Pattern Recognition, 2nd Edition*. John Wiley & Sons.