

Avaliação de métodos de classificação por regiões em imagens de alta resolução do sensor Rapideye

Luiz Felipe Parente Santiago^a, Wagner Barreto da Silva^b

^a4º Centro de Geoinformação (4º CGEO)

Avenida Marechal Bittencourt, 97, 69029-160, Compensa I, Manaus, AM, Brasil.

^bInstituto Militar de Engenharia (IME)

Praça General Tibúrcio, 80, 22290-270, Praia Vermelha, Rio de Janeiro, RJ, Brasil.

*luizparente.felipe@eb.mil.br

RESUMO: O objetivo deste trabalho é avaliar o desempenho dos métodos de classificação digital de imagens por regiões, Máquina de Vetores Suporte (SVM) e Distância Estocástica Mínima (SMD), e compará-los aos métodos de classificação digital de imagens por pixel, SVM e Máxima Verossimilhança (MaxVer), aplicados a imagens de alta resolução do satélite sensor RapidEye. A avaliação será segundo os critérios de matriz de confusão e Coeficiente de Concordância Kappa. As imagens utilizadas são do estado de Rondônia e apresentam resolução espacial de cinco metros. São utilizados também, para fins de aquisição de amostras, alguns resultados, da aquisição manual sobre as imagens, cedidos pela 5ª Divisão de Levantamento (5ª DL). Os testes mostraram que as classificações por regiões foram, em média, 13% superiores à classificação SVM por pixel e 34% superiores à classificação MaxVer por pixel na Área 1, 3% e 11% na Área 2 e 10% e 11% na Área 3.

PALAVRAS-CHAVE: classificação, rapideye, kappa.

ABSTRACT: This work aims in the assessment of region based image classification methods, Support Vector Machine (SVM) and Minimum Stochastic Distance (SMD), and to compare them with pixel based image classification methods, SVM and Maximum Likelihood, applied to high resolution imagery from RapidEye satellite. The assessment will be carried out according to the confusion matrix and the Kappa Concordance Coefficient. Images from Rondônia State were used. Such images have spatial resolution of 5 meters. In order to get some samples from the land cover classes, cartographic features acquired from 5th DL staff were used. From the performed evaluations, one can say that the classifications by regions were on average 13% higher than SVM by pixel and 34% higher than the MaxVer by pixel in Area 1, 3% and 11% in Area 2 and 10% and 11% in Area 3.

KEYWORDS: classification, rapideye, kappa.

1.1. INTRODUÇÃO

No universo das engenharias, as técnicas de Sensoriamento Remoto (SR) surgiram como um incremento na forma de obter informações dos objetos na superfície terrestre. Uma definição pode ser dada por [1]: *Sensoriamento Remoto é a ciência e a arte de obter informações sobre um objeto, área ou fenômeno através da análise de dados adquiridos por um dispositivo que não está em contato com o objeto, área ou fenômeno que está sendo investigado.*

Uma das maneiras de obter informações acerca dos objetos de forma remota é utilizando imagens de alta resolução.

As imagens de alta resolução são muito utilizadas para determinar o mapeamento do uso da terra. Entretanto, a classificação por *pixel* nestas imagens, em áreas urbanas, pode apresentar resultados não satisfatórios, devido à alta heterogeneidade e complexidade destas áreas, conforme indica [2]. Neste sentido, conhecer métodos de classificação digital de imagens que possam melhorar seus resultados é importante, pois, de posse de uma imagem classificada, a extração automática de feições pode ser realizada, utilizando um programa de processamento digital de imagem (PDI) e, assim, economizar tempo e, conseqüentemente, recursos financeiros durante o processo de produção cartográfica e, em especial, na Diretoria de Serviço Geográfico (DSG).

Uma maneira de realizar a classificação digital de imagens é utilizando abordagem por regiões, método que está sendo utilizado em imagens de alta resolução, em áreas urbanas, como uma alternativa ao método de classificação por *pixel*, na tentativa de melhorar os resultados [2]. Foram utilizados neste artigo os métodos máquina de vetores suporte (*support vector machine* – SVM) e distância estocástica mínima (*stochastic minimum distance* – SMD). Os métodos mencionados foram implementados por [3]

em que os autores desenvolveram um programa chamado *Region-Based Classification (RBC)*, utilizando a linguagem de programação *Interactive Data Language (IDL)*.

A presente trabalho tem como objetivo principal avaliar o desempenho dos métodos de classificação digital de imagens por regiões, SVM e SMD, e compará-los aos métodos de classificação digital de imagens por *pixel*, SVM e Máxima Verossimilhança (MaxVer), aplicados a imagens de alta resolução do satélite sensor RapidEye. Para as imagens em questão, os métodos por regiões devem apresentar melhores resultados, inclusive considerando diferentes níveis de segmentações, o que será confirmado ou não segundo os critérios de matriz de confusão (matriz de erro) e Coeficiente de Concordância Kappa

2. FUNDAMENTAÇÃO TEÓRICA

Serão apresentados conceitos básicos pertinentes à pesquisa.

2.1 Classificação de Imagens

A classificação de imagens é um ramo da área de reconhecimento de padrões e tem como principal objetivo rotular cada *pixel* da imagem dentro de uma classe de objetos determinada. Essa classificação dos *pixels* que formam a imagem dentro de classes é realizada a partir de regras de decisão que modelam as classes temáticas, e pode ser realizada classificando cada *pixel* da imagem individualmente ou um conjunto de *pixels* que compartilham de características comuns. As maneiras de classificação citadas recebem o nome de classificação por *pixel* e classificação por regiões, respectivamente. Existe também a classificação contextual que, apesar de não utilizar regiões, faz uso das informações da vizinhança de cada *pixel* para sua classificação, mas esta última não será objeto de análise neste artigo.

2.2 Classificação por pixel

Na classificação por *pixel* ou “*pixel a pixel*”, o *pixel* a ser classificado é rotulado dentro de alguma das classes definidas e nesta classificação é considerada apenas o atributo de cada *pixel*.

2.3 Classificação por regiões

Na classificação por regiões, é necessário que a imagem a ser classificada passe por um processo de segmentação. Os *pixels* que compartilham de características comuns são agrupados dentro do mesmo segmento (sub-região) da imagem, de forma que estas sub-regiões (C_1, C_2, \dots, C_r) apresentem as seguintes propriedades:

- $U_{i=1}^r C_i = I$
- C_i é um conjunto conectado, $i=1, 2, \dots, r$
- $C_i \cap C_j = \emptyset, \forall i, j, i \neq j$
- $Q(C_i) = \text{VERDADEIRA}$ para $i=1, 2, \dots, r$
- $Q(C_i \cup C_j)$ para quaisquer regiões adjacentes C_i e C_j em $Q(C_k)$, que é uma propriedade lógica definida sobre os pontos do conjunto C_k e \emptyset é o conjunto vazio [4].

De posse da imagem segmentada, as sub-regiões de pixels são rotuladas dentro de alguma das classes modeladas, de maneira única; todos os *pixels* da sub-região recebem o rótulo, da classe a que pertencerão, simultaneamente [5].

2.4 Máquina de Vetores Suporte (SVM)

O classificador digital de imagens *SVM* por *pixel* está sendo estudado há pouco mais de 20 (vinte) anos, tendo um de seus primeiros estudos publicados em 1992 [6] e, desde então, a comunidade científica está intensificando cada vez mais a sua utilização.

A crescente utilização do método *SVM* é justificado por suas características: algoritmo de arquitetura simples, complexidade computacional moderada, boa capacidade de generalização e robustez perante o fenômeno de Hughes o qual equivale à degradação da acurácia, de forma gradual, à medida que se aumenta a dimensionalidade dos dados [7] e [8].

Sabe-se que o processo de classificação por regiões consiste em associar cada sub-região (C_1, C_2, \dots, C_r) da imagem (I) de espaço de atributos X a uma classe (w_1, w_2, \dots, w_k). E no método *SVM*, o princípio básico é a rotulação dos *pixels* pertencentes as sub-regiões (C_1, C_2, \dots, C_r), dentro das classes a partir de hiperplanos de separação máxima. Este hiperplano é determinado na fase de treinamento e representa o lugar geométrico onde a função (1) é nula, sendo q o vetor ortogonal ao hiperplano, x é o vetor de atributos tal que $x \in X$ e $|b| / \|w\|$ a distância entre o hiperplano e o espaço de atributos.

$$f(x) = \langle q, x \rangle + b \quad (1)$$

Seja $D_l = \{(x_i, w_j) \in X \times \Omega : i = 1, 2, \dots, r; j = 1, 2\}$ e considerando inicialmente o conjunto de classes $\Omega = \{w_1, w_2\}$, os parâmetros q e b da função (1) são obtidos solucionando o problema de otimização listado em (2) por [9], sujeito às condições: e $\sum_{u=1}^m \lambda_u y_u = 0$ e $0 \leq \lambda_u \leq C, u=1, 2, \dots, m$, onde λ_u são multiplicadores de Lagrange, y_u é o indicador de classe de x_u de forma y_u é $+1$ se (x_u, w_1) e -1 se (x_u, w_2) que é esse

escalar um parâmetro (penalidade) que atua no processo de treinamento do método.

$$\max_{\lambda} \left(\sum_{u=1}^m \lambda_u - \frac{1}{2} \sum_{u,v=1}^m \lambda_u \lambda_v y_u y_v \langle x_u, x_v \rangle \right) \quad (2)$$

A partir da determinação dos parâmetros da função (1), o padrão x é associado a w_1 se $f(x) \geq 0$ ou então a w_2 se $f(x) < 0$, o que caracteriza o *SVM* como um classificador binário.

Utilizando a estratégia um-contratodos (estratégia multi-classes) explicada por [9], é possível utilizar o método *SVM* em problemas com mais de duas classes.

O produto interno $\langle x_u, x_v \rangle$, presente no problema (2), e, por sua vez, nos cálculos de pode ser substituído por funções *kernel* $K: X^2 \rightarrow \mathbb{R}$. Tais funções são usualmente adotadas para realização de classificação entre padrões não linearmente separáveis.

2.5 Distância Estocástica Mínima (SMD)

Para realizar a classificação por regiões utilizando *SMD*, é necessário primeiramente identificar na I , imagem antes da segmentação, as classes presentes (w_1, w_2, \dots, w_k). Para (w_1, w_2, \dots, w_k) são calculados os parâmetros das distribuições de probabilidade (N_1, N_2, \dots, N_k) para cada classe (w_1, w_2, \dots, w_k).

Realiza-se a segmentação da imagem a ser classificada, gerando as sub-regiões (C_1, C_2, \dots, C_r) e são estimados os parâmetros para as distribuições de probabilidade para cada sub-região formada na segmentação ($\bar{N}_1, \bar{N}_2, \dots, \bar{N}_r$). Para cada sub-região $C_i, i=1, 2, \dots, r$ é calculada a distância estocástica entre a sua distribuição de probabilidade $N_i, i=1, 2, \dots, r$ estimada a partir dos dados provenientes das amostras das classes e as distribuições de probabilidade estimadas a partir dos dados das sub-regiões ($\bar{N}_1, \bar{N}_2, \dots, \bar{N}_k$). A classe que tiver a menor distância (distância estocástica mínima) com a sub-região $C_i, i=1, 2, \dots, r$ será atribuída à sub-região em questão [10].

2.6 Matriz de Confusão (ou Matriz de Erro)

É uma matriz que apresenta como os *pixels* da “verdade terrestre” (dados de teste) foram classificados. A matriz indica para cada classe dos dados de teste a quantidade de *pixels* classificados em cada uma das classes presentes na classificação. A matriz de confusão compara, classe por classe, a relação entre os dados de teste e os correspondentes resultados da classificação. O número de linhas e o número de colunas dessa matriz devem ser iguais ao número de classes da classificação.

As linhas da matriz representam a verdade terrestre e as colunas, o resultado da classificação. Para cada classe da verdade terrestre será indicado a quantidade de *pixels* classificados em cada uma das classes do resultado da classificação. Resultados da diagonal principal indicam acerto na classificação e resultados fora da diagonal principal indicam erros de classificação. Os erros de classificação são: erros de omissão (EO) – elementos das linhas classificados em classes diferentes da que pertencem e erros de comissão (EC) – elementos das colunas classificados em classes diferentes da que pertencem [11].

A Tabela 1 mostra um exemplo simulado de matriz de confusão.

Tab 1: Exemplo de Matriz de Confusão.

Classes		Classificação					Acurácia Percentual		
		C1	C2	C3	C4	Total			
Verdade	C1	15	1	1	0	17	88 %	2	5
	C2	0	25	0	1	26	96 %	1	9
	C3	2	1	35	0	38	92 %	3	2
	C4	3	7	1	45	56	80 %	11	1
	Total	20	34	37	46	137	87 % (Global)	17	17

Da Tabela 1 pode-se extrair a acurácia percentual por classe que é o número de *pixels* classificados corretamente pelo total de *pixels* da classe, os erros de omissão e comissão e também a acurácia global da classificação que é representada pelo total de *pixels* classificados corretamente pelo total de *pixels* classificados. Neste exemplo tem-se:

Acurácia da classe C1: $(15/17) \cdot 100 = 88 \%$

Acurácia da classe C2: $(25/26) \cdot 100 = 96 \%$

Acurácia da classe C3: $(35/38) \cdot 100 = 92 \%$

Acurácia da classe C4: $(45/56) \cdot 100 = 80 \%$

Acurácia Global: $(120/137) \cdot 100 = 87 \%$

2.7 Coeficiente de Concordância Kappa (\hat{K})

O Coeficiente de Concordância Kappa (\hat{K}) é outro indicativo da acurácia da classificação e indica a concordância entre os dados de teste e o resultado da classificação.

O (\hat{K}) é calculado através da fórmula (3) descrita em [11]:

$$\hat{K} = \frac{\theta_1 - \theta_2}{1 - \theta_2} \quad (3)$$

em que $\theta_1 = \frac{1}{n} \sum_{i=1}^m x_{ii}$ e $\theta_2 = \frac{1}{n^2} \sum_{i=1}^m [(\sum_{j=1}^m x_{ij})(\sum_{j=1}^m x_{ji})]$, m é o número de classes e n o total de amostras de teste. A Tabela 2, formulada por [12], indica uma relação da força de concordância entre os dados de teste e o resultado da classificação de acordo com o valor do (\hat{K}).

Tabela 2: Exatidão da Classificação

Coefficiente de Concordância Kappa	Força da Concordância
$\hat{K} \leq 0$	Pobre
$0 < \hat{K} \leq 0,2$	Ruim
$0,2 < \hat{K} \leq 0,4$	Razoável
$0,4 < \hat{K} \leq 0,6$	Boa
$0,6 < \hat{K} \leq 0,8$	Muito Boa
$0,8 < \hat{K} \leq 1$	Excelente

Fonte: Adaptado de [12].

3. METODOLOGIA

A metodologia foi dividida em 4 fases. Após a realização das classificações, as fases de avaliação. A Fig. 1 ilustra a divisão da metodologia e avaliação dos métodos de classificação.

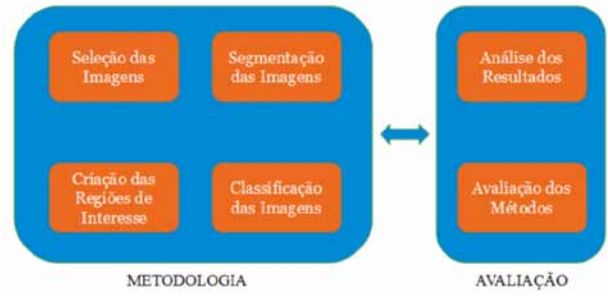


Fig. 1: Fases da Metodologia e Avaliação

3.1 Seleção das Imagens

Na etapa de seleção das imagens, foram escolhidas imagens que apresentavam dados de aquisição manual, realizada pela 5ª DL, de classes de uso e cobertura do solo, formadas por primitiva geométrica do tipo área, de classes de feições presentes nas seguintes categorias da Especificação Técnica para a Estruturação de Dados Geoespaciais Vetoriais (ET-EDGV): sistemas de transporte, hidrografia, vegetação, relevo e localidades.

3.2 Segmentação das Imagens

Na fase de segmentação, foi utilizado o segmentador por crescimento de regiões do programa SPRING 5.1.8. A segmentação por crescimento de regiões tem por princípio agregar regiões adjacentes que satisfaçam a um teste de hipótese dentro de um limiar estabelecido (similaridade) e as regiões precisam também atender ao critério de quantidade mínima de *pixels* por área. Para a realização dos testes dos classificadores por regiões *SVM* e *SMD*, foram realizadas seis segmentações compreendendo seis combinações diferentes de valor de similaridade e área mínima.

3.3 Criação das Regiões de Interesse (ROI)

A criação das *ROI* foi baseada nos dados de aquisição manual da 5ª DL. Foram identificadas amostras das classes de interesse nas imagens, para que o classificador possa realizar a classificação automática da imagem inteira. Nesta etapa foi feito um estudo de separabilidade das classes. A necessidade de realizar este estudo de separabilidade das classes se deve ao fato da 5ª DL não ter atribuído corretamente à vegetação tipo “campo”. Ou seja, observando as imagens, é possível identificar visualmente diferenças na classe “campo” da aquisição manual da 5ª DL, porém os “diferentes tipos de campo” foram considerados iguais. Então, o estudo de separabilidade é uma tentativa de separar os “diferentes tipos de campo” existentes nas imagens.

3.4 Classificação das Imagens

Nesta fase, foram realizadas as classificações utilizando os classificadores por regiões *SVM* e *SMD* do programa *RBC* e também dois classificadores por *pixel* disponíveis no programa *ENVI 4.8*: Máxima Verossimilhança (*MaxVer*) e *SVM*. A utilização destes dois classificadores por *pixel* será para fins de comparação entre os métodos de classificação por regiões e por *pixel*. Foram escolhidos os dois métodos por *pixel* citados, pois estes são consagrados na literatura. Os critérios das classificações serão explicados no capítulo 4 (Parte Experimental).

4. PARTE EXPERIMENTAL

Para a realização desta etapa, foram escolhidas 3 áreas de estudo (recortes de imagens RapidEye) e, para cada área de estudo, foram realizadas 6 segmentações diferentes para realizar as classificações por regiões. As imagens foram classificadas de acordo com os dados de aquisição manual realizados pela 5ª DL. Porém, para a classe “campo”, identificou-se a necessidade de realizar um estudo de separabilidade, pois visualmente era possível identificar diferentes tipos de campo.

Para o estudo de separabilidade da classe “campo”, adotou-se como padrão que existiam 3 campos distintos nas imagens nos casos iniciais. Desta forma as imagens eram classificadas em todos os 4 métodos avaliados e posteriormente os dois campos que apresentavam maior proximidade (mais parecidos) eram agrupados e as imagens novamente classificadas com 2 campos. Depois os três campos eram agrupados e as imagens reclassificadas.

Desta forma, foram gerados 42 resultados de classificação (18 por método de classificação por regiões e 3 por método por *pixel*) em cada área de estudo, totalizando 126 resultados de classificação.

5. ANÁLISE DE RESULTADOS

Foi realizada a análise dos 126 resultados das classificações. Porém, para fins deste artigo, as imagens e todas as tabelas comparativas entre os métodos serão suprimidas e será mostrado apenas um exemplo dos resultados. A Fig. 2 representa a área de estudo 1 e a Fig. 3 ilustra as amostras de treinamento e teste.



Fig. 2: Área de Estudo 1 – Recorte de Imagem RapidEye

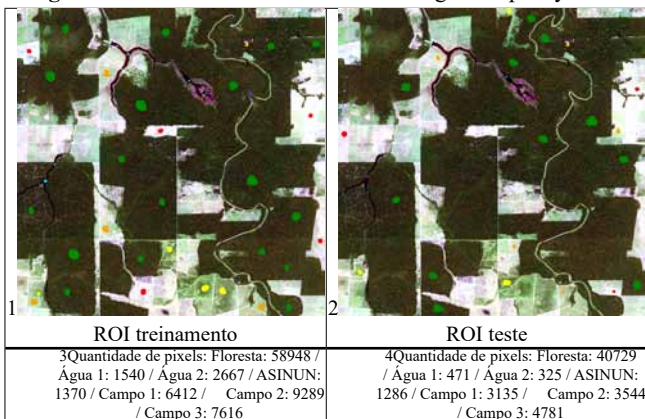


Fig. 3: ROI de Treinamento e Teste

A Tabela 3 ilustra a tabela de cores utilizada nas classificações.

Tabela 3: Legenda de Cores das Classificações

Legenda			
	Floresta		Campo 1
	Água 1		Campo 2
	Água 2		Campo 3
	ASINUN		

A Tabela 4 apresenta o resultado das classificações por regiões e por *pixel*.

Tabela 4: Resultados das Classificações

Classificação		Conceito (Tabela 2)
SVM por regiões	0,9126	Excelente
SMD por regiões	0,9212	Excelente
SVM por pixel	0,7945	Muito Bom
MaxVer por pixel	0,6833	Muito Bom

A Fig. 4 apresenta dois resultados (imagens) de classificações (melhor e pior Kappa) presentes na Tabela 4.

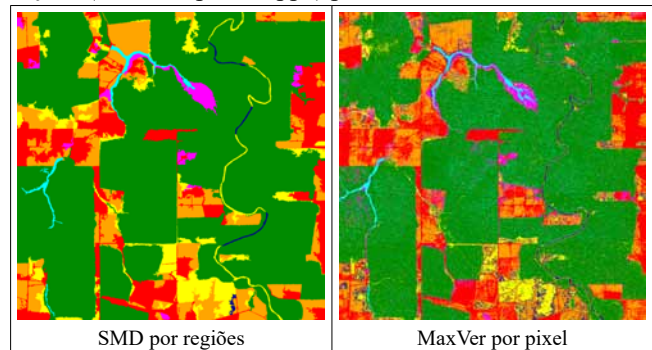


Fig. 4: Classificações (melhor e pior kappa)

Quanto à segmentação, os resultados mostram que a diferença entre a melhor classificação e a pior classificação foram: na Área 1 (8% para SVM e 9% para SMD), Área 2 (6% para SVM e 10% SMD) e na Área 3 (25% SVM e 35% SMD). Não houve uma segmentação que apresentasse melhores resultados em todas as 3 áreas de estudo, porém a segmentação 15_1000 (similaridade 15 e quantidade mínima de *pixels* por área 1000) apresentou os melhores resultados no método SVM por regiões para os casos com 7, 6 e 5 classes na área 1 e a segmentação 5_200 apresentou os melhores resultados no método SMD por regiões para os casos com 7, 6 e 5 classes na área 1. Conclui-se que os resultados das classificações por regiões levando em consideração a segmentação utilizada variou em cada área de estudo, mostrando que a imagem influencia o resultado final.

Quanto à separabilidade entre as classes “campo”, constatou-se que na Área 1 o número de classes mais provável é 6, ou seja, existem dois campos distintos. Na Área 2 o número de classes mais provável é 5, ou seja, existem dois campos distintos. Na Área 3 não foi possível indicar o número de campos, pois os resultados foram semelhantes em cada caso estudado.

Quanto à abordagem da classificação ser baseada em regiões ou por *pixel* os resultados mostram que as classificações por regi-

ões foram em média 13% superiores à classificação *SVM* por *pixel* e 34% superiores que a classificação MaxVer por *pixel* na Área 1, 3% e 11 % na Área 2 e 10% e 11% na Área 3, indicando que para as imagens analisadas os classificadores por regiões foram mais bem-sucedidos, tanto visual quanto quantitativamente, que os classificadores por *pixel*.

6. CONCLUSÃO

Os métodos de classificação por regiões foram em média 9% superiores ao método *SVM* por *pixel* e 19% superiores ao método MaxVer por *pixel*. Entre os métodos de classificação por regiões, o método *SMD* apresentou resultado melhor que o *SVM* em 6 dos 9 testes apresentados.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] LILLESAND, T. M., KIEFER, R. W. Remote Sensing and Image Interpretation. 3. ed. New Jersey: John Wiley & Sons, 1994, 750 p. ISBN 0471 305 758.
- [2] LIU, Desheng, XIA, Fan. Assessing object-based classification: advantages and limitations In: Remote Sensing Letters, 1:4, Dec. 2010, p. 187-194, DOI: 10.1080/01431161003743173. Disponível em: <http://dx.doi.org/10.1080/01431161003743173> [capturado em 19 Fev. 2016].
- [3] NEGRI, R. G. et al. Examining Region-Based Methods for Land Cover Classification Using Stochastic Distances. In: International Journal of Remote Sensing, 2016.
- [4] GONZALEZ, Rafael C., WOODS, Richard E. Digital Image Processing. 3. ed. New Jersey: Pearson Prentice Hall, 2008, 103 p. ISBN 978-0-13-168728-8.
- [5] SILVA, Wagner Barreto da. Classificação de regiões de imagens utilizando testes de hipótese baseados em distâncias estocásticas: aplicações a dados polarimétricos. 143 p. Tese (Doutorado em Sensoriamento Remoto) – Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, 2013.
- [6] BOSER, Bernhard E. et al. A training algorithm for optimal margin classifiers. In: Annual workshop on computational learning theory, 5. Proceedings... Pittsburgh: ACM Press, 1992. p. 144-152, ISBN 0-89791-497-X.
- [7] ZORTEA, M., MARTINO, M. D., SERPICO, S. A SVM ensemble approach for spectral-contextual classification of optical high spatial resolution imagery. In: IEEE International Geoscience and Remote Sensing Symposium. Proceedings... Barcelona: 2007, p. 1489-1492, ISBN 978-1-4244-1211-2.
- [8] BRUZZONE, L., PERSELLO, C. A novel context-sensitive semi-supervised SVM classifier robust to mislabeled training samples. In: IEEE Transactions on Geoscience and Remote Sensing, v. 47, n. 7, p. 2142-2154, 2009.
- [9] THEODORIDIS, Sergios, KOUTROUMBAS, Konstantinos. Pattern Recognition. 4. ed. San Diego: Academic Press, 2009. 961 p. ISBN: 978-1-59749-272-0.
- [10] SILVA, Wagner Barreto da et al. Land cover discrimination at Brazilian Amazon using region based classifier and stochastic distance. In: IEEE International Geoscience and Remote Sensing Symposium, 2011, p. 2900-2903.
- [11] MATHER, P. M. Computer processing of remotely-sensed images: an introduction. 3. ed. West Sussex: John Wiley & Sons, 2004, 324 p. ISBN 0470 849 185.
- [12] LANDIS, J. Richard, KOCH, Gary G. The Measurement of Observer Agreement for Categorical Data. In: Biometrics, Vol. 33, No. 1, Mar. 1977, p. 159-174.