

# Classificação Supervisionada de Imagens Texturais Utilizando Redes Neurais Artificiais

Emilson Pereira Leite<sup>1</sup>  
Carlos Roberto de Souza Filho<sup>1</sup>

<sup>1</sup> Universidade Estadual de Campinas – UNICAMP/IG  
Caixa Postal 6152 – 13083-970 – Campinas-SP, Brasil  
{emilson, beto}@ige.unicamp.br

**Abstract.** A methodology to perform supervised classification of textural images using Artificial Neural Networks for applications in the Geosciences is presented in this work. Feature vectors are built with textural information composed of semivariogram values, histogram measures of mean, standard deviation and weighted-rank fill ratio. Feed-forward back-propagation Artificial Neural Networks are designed and trained so as to minimize the mean squared error of the differences between feature and target vectors of training sets. At each training iteration, the mean squared error for validation and test sets are also evaluated. Global accuracy and kappa coefficient are calculated for training, validation and test sets, allowing a quantitative appraisal of the predictive power of the Neural Networks. The best model for classification of all pixels in a given textural image is obtained from a  $k$ -fold cross-validation. The methodology was tested using synthetic images and airborne, multi-polarized SAR imagery for geologic mapping, and the overall results are considered quite positive.

**Palavras-chave:** Semivariogramas, Classificação Supervisionada, Redes Neurais de Alimentação Direta, Imagens Texturais, Imagens de RADAR.

## 1. Introdução

Parâmetros texturais podem ser extraídos de imagens de sensoriamento remoto (e.g. radar ou multi-espectral) e utilizados como atributos para classificação de feições desconhecidas. Esta classificação pode ser aplicada à discriminação litológica (e.g. Chica-Olmo and Abarca-Hernandéz, 2000), mapeamento de uso do solo (e.g. Dekker, 2003), interpretação de fácies sísmicas (e.g. Gao 2008), interpretação de imagens de GPR (e.g. Moysey et al., 2006), classificação de tipos de rocha (e.g. Marmo et al., 2005), e assim por diante.

Vários parâmetros locais de textura como entropia, dimensão fractal, energia *wavelet*, medidas de co-ocorrência e semivariogramas têm sido empregados com sucesso em métodos de classificação supervisionada e não-supervisionada. Particularmente, semivariogramas têm sido úteis em problemas de classificação de imagens de radar orbital como reconhecimento de terreno (Carr, 1996; Carr and Miranda, 1998), mapeamento de vegetação (Miranda et al., 1998), mapeamento urbano (Car and Miranda, 1998) e detecção de vazamento de óleo (Miranda et al., 2004). Os resultados bem sucedidos desses métodos de classificação recentemente desenvolvidos, e que utilizam semivariogramas, constituem a principal motivação para a abordagem metodológica deste trabalho. O objetivo é fornecer uma metodologia para descrever texturas quantitativamente, combinando semivariogramas com parâmetros estatísticos locais tais como média, desvio padrão e razão de classificação-ponderada (Novak et al., 1993), e classificá-los aplicando Redes Neurais Artificiais (RNAs) supervisionadas.

RNAs são modelos matemáticos que simulam a estrutura e funcionamento do cérebro humano. Esses modelos consistem de camadas de processamento nas quais cada elemento em uma dada camada representa um neurônio da camada. Tais modelos possuem poderosas características, tais como a habilidade de aproximar qualquer função arbitrária, pelo aprendizado e adaptação ao seu ambiente; e a facilidade em reconhecer novos dados, mesmo quando o conhecimento estatístico não é completo (e.g. Haykin, 1999). Devido ao freqüente caráter não-linear entre textura de imagens e outros dados geocientíficos, as RNAs constituem uma das ferramentas mais apropriadas a serem utilizadas neste contexto.

## 2. Metodologia

### 2.1 Vetores de atributos

Semivariogramas empíricos são obtidos a partir de mapas de variogramas que são calculados no domínio da frequência segundo procedimento descrito em Marcotte (1996). A transformada de Fourier ( $TF$ ) da variância ( $V$ ) de uma função  $f(x,y)$  na distância de passo  $r$ , é expressa como

$$TF(V) = F^*(u,v)G(u,v), \quad (1)$$

onde:  $F(u,v)$  é a  $FT$  de  $f(x,y)$ ;  $F^*(u,v)$  é o complexo conjugado de  $F(u,v)$ ; e  $G(u,v)$  é a  $FT$  de  $f(x+r, y+r)$ . Um mapa de variograma é obtido pela transformada inversa de Fourier  $FT^{-1}$  para cada distância de passo  $r=1 \dots W$ . Um algoritmo de busca é aplicado para extrair semivariogramas em uma direção específica, que é determinada por um ângulo em relação a direção NS do mapa de variograma. Semivariogramas omni-direcionais são obtidos calculando-se a média radial das variâncias em todas as distâncias a partir do centro do mapa. Medidas de histograma como média ( $\mu$ ), desvio padrão ( $\sigma$ ) e razão de classificação ponderada ( $\eta$ ) (Novak et al., 1993) são incluídas nos vetores de atributos para levar em conta a variabilidade radiométrica global da imagem. A razão de classificação ponderada é dada por:

$$\eta = \frac{\text{soma } x\% \text{ pixels mais brilhantes}}{\text{soma todos pixels}}; \quad (2)$$

onde  $x$  é, tipicamente, igual a 15%. A introdução destas medidas ajuda a aumentar a acurácia da classificação. Com isto, um vetor de atributos  $\mathbf{p}$  genérico, para um dado pixel  $(i, j)$ , que alimenta a RNA, é representado por

$$\mathbf{p}_{i,j} = [\gamma_{i,j}^1, \dots, \gamma_{i,j}^W, \mu, \sigma, \eta]^T. \quad (3)$$

### 2.2 Modelo de Rede Neural Artificial

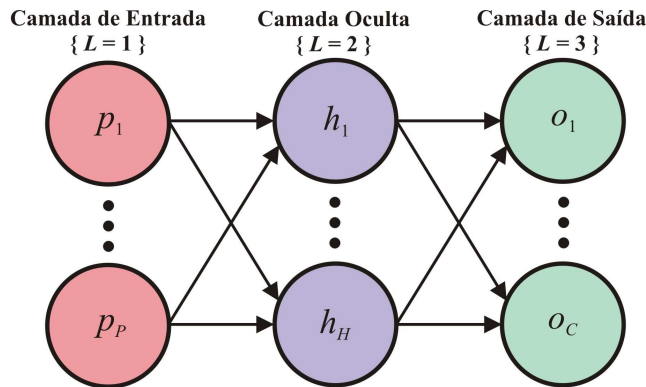
A metodologia proposta emprega RNAs de alimentação direta, treinadas pelo algoritmo de retro-propagação, como mostra o esquema da Figura 1. Cada vetor de atributos de entrada  $\mathbf{p} = [p_1 \dots p_p]$  apresentado à RNA é construído utilizando a Equação 3, sendo completamente conectado à camada oculta. A saída da camada oculta é um vetor  $\mathbf{h} = [h_1 \dots h_H]$ , que é completamente conectado à camada de saída, onde  $H$  é o número de neurônios na camada oculta. A camada de saída fornece um vetor  $\mathbf{o} = [o_1 \dots o_C]$ , onde  $C$  é o número de neurônios desta camada. No contexto deste trabalho,  $C$  é igual ao número de classes texturais na imagem de entrada. Neste tipo de RNA, todas as conexões entre camadas podem ser genericamente expressas por

$$\mathbf{y}^{\{L\}} = \mathbf{W}^{\{L\}} \times \mathbf{x}^{\{L-1\}} + \mathbf{b}^{\{L\}}, \quad L = 2, 3, \quad (4)$$

onde: o  $L$  indica o número da camada à qual os pesos são referidos;

$$\mathbf{W}^{\{L\}} = \begin{pmatrix} w_{1,1} & \cdots & w_{1,R} \\ \vdots & \ddots & \vdots \\ w_{S,1} & \cdots & w_{S,R} \end{pmatrix} \quad (5)$$

é uma matriz  $S \times R$  de pesos;  $S$  é o número de neurônios na camada  $L^{\text{th}}$  e  $R$  é o número de elementos de entrada fornecidos pela camada  $(L-1)^{\text{th}}$  precedente;  $\mathbf{y}^{\{L-1\}}$  é um vetor  $R \times 1$  de entrada; e  $\mathbf{b}^{\{L\}} = [b_1 \cdots b_S]^T$  é um vetor de tendência. Cada camada possui uma função de transferência associada que é avaliada em  $\mathbf{y}^{\{L\}}$  para fornecer uma saída para a camada seguinte. As funções ‘tangente hiperbólica sigmoideal’ e ‘puramente linear’ são empregadas como funções de transferência nas camadas oculta e de saída, respectivamente (Demuth et al., 2008). O índice da linha do maior elemento de  $\mathbf{o}$  é o código da classe textural atribuída a  $\mathbf{p}$ .  $\mathbf{W}^{\{L\}}$  e  $\mathbf{b}^{\{L\}}$  são inicializados de acordo com o algoritmo de Nguyen-Widrow (Nguyen and Widrow, 1990) e são atualizados iterativamente pela aplicação do algoritmo do gradiente descendente (e.g. Snyman, 2005).



**Figura 1.** Rede Neural Artificial com 3 camadas.  $P$ ,  $H$  and  $C$  são as quantidades de neurônios em cada camada. Neste trabalho, o valor máximo de  $P$  é  $W + 3$  (ver Equação 3).

### 2.3 Treinamento, Validação e Teste

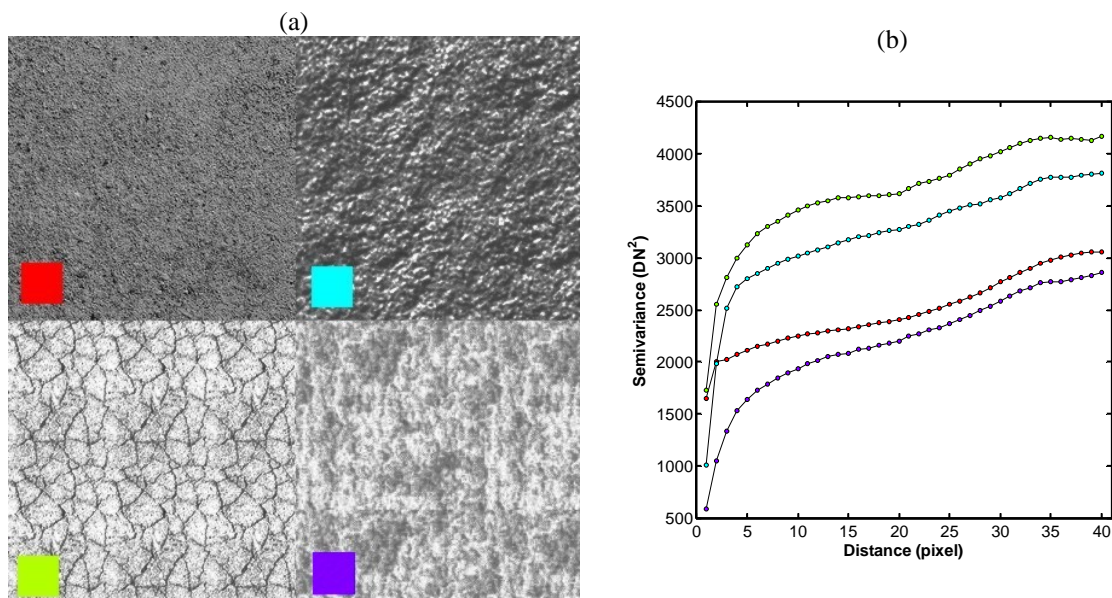
Cada vetor de atributos de entrada  $\mathbf{p}$  é normalizado de forma a possuir média igual a zero e desvio padrão igual a um, através da expressão  $\mathbf{p}' = (\mathbf{p} - \mu_{\mathbf{p}}) / \sigma_{\mathbf{p}}$ .  $N$  amostras são extraídas das classes texturais na imagem e 20% destas amostras são separadas para formar um conjunto de teste. O conjunto de teste é utilizado para estimar o erro de generalização e o poder de predição do modelo de RNA. Aplica-se validação cruzada do tipo  $k$ -fold (e.g. Plutowski et al., 1994) nas  $M$  amostras restantes, onde  $M = N - 0.2 \cdot N$ . Neste conjunto,  $M/k$  amostras são utilizadas para validação em cada um dos  $k$  modelos, enquanto que  $N - M/k$  são utilizadas para treinamento. Para cada vetor  $\mathbf{p}$  nos conjuntos de treinamento, validação e teste, existe um vetor-alvo  $C \times 1$  correspondente associado à  $j^{\text{th}}$  classe textural, onde  $j = 1 \cdots C$ . O  $j^{\text{th}}$  elemento de  $\mathbf{t}$  é igual a um e o resto de seus elementos é igual a zero.

O treinamento da RNA é conduzido até que pelo menos uma das seguintes condições seja atingida, a fim de evitar *overfitting*: (i) o erro quadrático médio (EQM) do conjunto de validação é igual a um erro de minimização global ( $E_g$ ); (ii) o EQM do conjunto de validação não diminui por três vezes consecutivas (*early-stopping*); ou (iii) o número de iterações é igual a um número máximo ( $N_{itr}$ ). A acurácia global e o coeficiente kappa são calculados da

forma definida em Foody (1992). O melhor modelo é selecionado pela comparação dos EQMs, acurácias globais e coeficientes kappa entre os  $k$  modelos de RNAs.

### 3. Estudo de Caso 1 – Imagens Texturais Genéricas

A imagem utilizada neste teste é composta por quatro texturas que representam aquelas comumente observadas em imagens de radar em amplitude (Figure 2a). Esta imagem-teste foi adequada para checar: (i) a habilidade dos modelos de RNAs para reconhecer os limites de classes texturais; (ii) a sensibilidade à variação de escala, que é estabelecida pelo tamanho da janela que limita o cálculo dos semivariogramas; e (iii) a sensibilidade à variação dos atributos dos vetores de entrada. As texturas representam dois tipos de solos (texturas no canto esquerdo superior e inferior); rocha (canto superior direito); e copa de floresta madura (canto inferior direito). Foram selecionadas 20 amostras em cada uma dessas quatro texturas e estas foram submetidas à validação cruzada  $k$ -fold, para  $k = 8$ . Em cada uma das 8 subamostragens no processo de validação cruzada, 12 amostras foram utilizadas para validação e 56 para treinamento.



**Figura 2.** (a) Imagem-teste composta por 471 x 473 pixels mostrando quatro texturas distintas. Os quadrados coloridos representam matrizes de 30 x 30 pixels. (b) Semivariogramas médios calculados a partir de 80 amostras na imagem-teste usando janela de 30 x 30 pixels ( $W = 30$ ).

As distintas texturas são refletidas nos semivariogramas médios apresentados na Figura 4b, obtidos para  $W = 40$ . No primeiro teste, foi utilizado  $H = 30$ , variando-se o tamanho da janela ao longo de quatro treinamentos (Tabela 1). No segundo teste, utilizou-se  $W = 40$ , variando-se  $H$  (Tabela 2). No terceiro teste, foram fixados  $W$  e  $H$  ( $H = 10$  e  $W = 40$ ), variando-se os atributos dos vetores de entradas. Os semivariogramas foram empregados em todos os testes, com  $Eg = 0.001$  e  $Nitr = 2000$ . A melhor combinação de parâmetros foi  $W = 40$ ,  $H = 10$  e  $\mathbf{p} = [\gamma, \mu, \kappa]$ . A Figura 3 mostra os mapas classificados obtidos utilizando-se os quatro valores de  $W$  e  $H$  fixos. As classes verde e ciano são as mais afetadas pela escolha errada de  $W$ . Efeitos de borda são observados no contorno entre as classes violeta/ciano e vermelha/verde. Visualmente, os piores resultados são aqueles obtidos considerando-se como atributos somente as semi-variâncias (Figura 3e).

**Tabela 1** – Performance média da classificação das amostras de teste com o aumento de  $W$ .

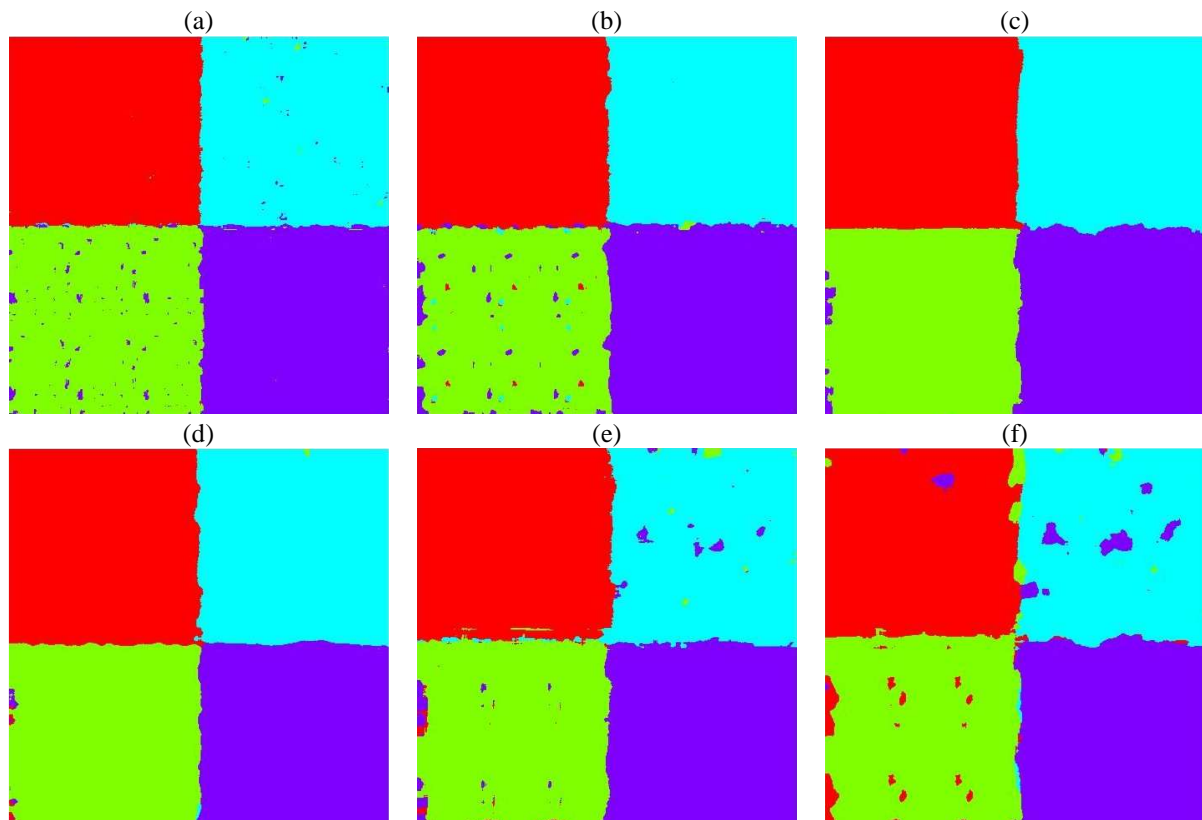
$W$	$\mu_{MSE}$	$\mu_G$	$\mu_\kappa$
10	0.2072	92.1875	0.8958
20	0.1967	93.7500	0.9167
30	0.1303	96.8750	0.9583
40	0.0884	96.8750	0.9583

**Tabela 2.** Performance média da classificação das amostras de teste com o aumento de  $H$ .

$H$	$\mu_{MSE}$	$\mu_G$	$\mu_\kappa$
05	0.1100	96.8750	0.9583
10	0.0956	96.8750	0.9583
20	0.1583	96.8750	0.9583
30	0.1492	96.8750	0.9583
40	0.1562	96.8750	0.9583
60	0.2181	95.3125	0.9375
100	0.2107	92.1875	0.8958

**Tabela 3** Performance média da classificação das amostras de teste para diferentes atributos.

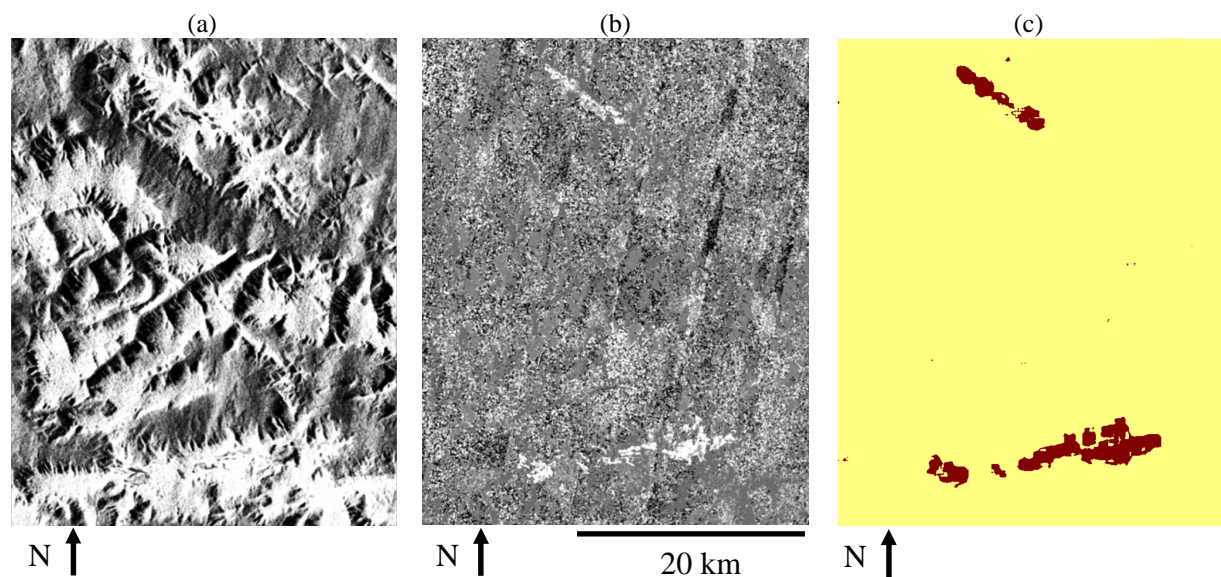
$p$ ( $i = 1 \dots 30$ )	$\mu_{MSE}$	$\mu_G$	$\mu_\kappa$
$[\chi(i)] H$	0.38	93.75	0.92
$[\chi(i), \mu] G$	0.13	96.87	0.96
$[\chi(i), \sigma] F$	0.11	96.87	0.96
$[\chi(i), \eta] E$	0.40	85.94	0.81
$[\chi(i), \mu, \sigma] D$	0.12	98.44	0.98
$[\chi(i), \mu, \eta] C$	0.09	100.00	1.00
$[\chi(i), \sigma, \eta] B$	0.14	96.87	0.96
$[\chi(i), \mu, \sigma, \eta] A$	0.15	96.87	0.96



**Figura 3.** Mapas classificados da imagem-teste ( $H = 10$  para todos os mapas): (a)  $W = 10$  e  $\mathbf{p} = [\gamma, \mu, \eta]$ ; (b)  $W = 20$  e  $\mathbf{p} = [\gamma, \mu, \eta]$ ; (c)  $W = 30$  e  $\mathbf{p} = [\gamma, \mu, \eta]$ ; (d)  $W = 40$  e  $\mathbf{p} = [\gamma, \mu, \eta]$ ; (e)  $W = 40$  e  $\mathbf{p} = [\gamma, \eta]$ ; e (e)  $W = 40$  and  $\mathbf{p} = \gamma$

#### 4.2 Estudo de Caso 2: Imagens R-99B/SAR (SIPAM/SIVAM) na Amazônia

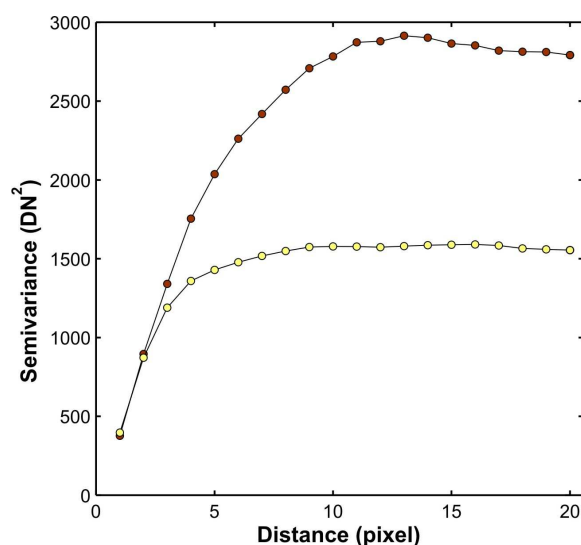
Para mostrar uma aplicação da metodologia no mapeamento geológico, foram utilizadas imagens em amplitude da banda L do sistema aerotransportado R-99B/SAR (SIPAM/SIVAM), particularmente coletadas para simular imagens do sistema orbital chamado MAPSAR (Schröder et al., 2005). Cada imagem compreende 845 x 652 pixels e cobre uma pequena parte da floresta amazônica localizada na Província Mineral de Carajás (PMC). O objetivo deste teste é mapear as Formações Ferríferas Bandadas (FFBs) reveladas como os pixels mais brilhantes e que formam áreas pouco heterogêneas nas porções norte e sul da imagem da segunda Componente Principal do triplete SAR-HH-HV-VV (Figure 4b). A Análise por Componentes Principais foi efetuada utilizando o método da covariância após a subtração do valor médio de cada conjunto de DN's de entrada. Para cada uma das duas classes (FFB e não-FFB), foram extraídas 60 amostras para treinamento, validação e teste das RNAs. Neste caso, foi empregado o semivariograma na direção E-W (Figura 5). Para particionar o conjunto de amostras em conjuntos de treinamento, validação e teste, o mesmo procedimento conduzido no teste apresentado na seção 4.1 foi aqui aplicado.



**Figura 4.** Primeira (a) e Segunda (b) Componente Principal das imagens do R-99B/SAR (banda L; HH-HV-VV) da PMC, região amazônica. (c) Mapa obtido pela classificação de todos os pixels da imagem em (b) pela RNA selecionada. Foram utilizados  $W = 20$ ,  $Nitr = 5000$ ,  $k = 8$ ,  $H = 20$  and  $Eg = 0.001$ .

Cinco dentre oito modelos forneceram acurácia global igual a 100%. O modelo selecionado, utilizado para gerar o mapa da Figura 4c, foi escolhido baseado nos EQMs, sendo estes iguais a 0,0407 e 0,0765 para os conjuntos de validação e teste, respectivamente.

Embora um estudo detalhado, incluindo integração de dados e checagem de campo seja necessário para delinear melhor as FFBs na área de estudo, o método separou satisfatoriamente o ruído (classe não-FFB) do sinal (classe FFB), fornecendo um produto interessante a ser utilizado em trabalhos futuros de maior detalhe. Todas as amostras foram extraídas das FFB ao sul da imagem e, portanto, a RNA detectou as porções ao norte de forma bastante precisa.



**Figura 5.** Semivariogramas médios calculados a partir de 60 semivariogramas E-W extraídos das duas classes (FFB: círculos marrons; não-FFB: círculos amarelos), ilustradas na Figura 4b.

## 5. Conclusões

De forma geral, a metodologia proposta possibilita: (i) extrair informação textural de imagens SAR ou qualquer imagem textural de interesse; (ii) treinar, validar e tester modelos de RNAs; (iv) classificar todos os pixels de uma imagem e obter um mapa a partir do qual podem ser feitas predições de diversas naturezas, tais como geológicas e exploratórias.

Os testes apresentados permitem as seguintes conclusões: (i) a utilização de um menor número de neurônios na camada oculta produz uma maior acurácia na classificação e o menor erro estimado; (ii) a utilização de um tamanho de janela inadequado diminui o poder de predição dos modelos de RNAs; (iii) a técnica de *early-stopping* é eficiente para evitar *overfitting*; (iv) a classificação textural torna-se problemática para pixels próximos aos limites entre classes. Isto pode ser contornado selecionando-se algumas amostras próximas a estes limites, fazendo com que os semivariogramas incorporem informação textural de mais de uma classe; e (v) devido à alta complexidade de cenas naturais, deve-se utilizar o maior número possível de amostras de treinamento, de forma a contemplar pequenas variações nos semivariogramas das classes.

A partir de uma cuidadosa escolha de parâmetros, esta metodologia pode ser utilizada com sucesso na classificação de imagens texturais e os mapas classificados podem ser facilmente integrados em SIGs para uma ampla variedade de aplicações nas Geociências.

## Agradecimentos

E.P. Leite e C.R. Souza Filho agradecem a Fundação de Amparo a Pesquisa do Estado de São Paulo (FAPESP) pela bolsa de pós-doutorado (Proc. N.: 05/04453-3) e pelo suporte financeiro (Proc. N. 03-09916-6), respectivamente.

## Referências Bibliográficas

Carr, J.R., 1996. Spectral and textural classification of single and multiple band digital images. **Computers & Geosciences** 22, 849-865.

Carr, J.R. and Miranda, F.P., 1998. The semivariogram in comparison to the co-occurrence matrix for classification of image texture. **IEEE Trans. Geosci. Remote Sensing** 36, 1945-1952.

Chica-Olmo, M., Arbarca-Hernández, F., 2000. Computing geostatistical image texture for remotely sensed data classification. **Computers & Geosciences** 26 (4), 373-383.

- Dekker, R.J., 2003. Texture analysis and classification of ERS SAR images for map updating of urban areas in the Netherlands. **IEEE Transactions on Geoscience and Remote Sensing** 41(9), 1950-1958.
- Demuth, H., Beale, M., and Hagan, M., 2008. **Neural Network Toolbox™ 6**. User's Guide. The MathWorks™, Natick, MA, 907 pp.
- Foody, G. M., 1992. On the compensation for change agreement in image classification accuracy assessment. **Photogrammetric Engineering and Remote Sensing** 6(10), 1459-1460.
- Gao, D., 2008. Application of seismic texture model regression to seismic facies characterization and interpretation. **The Leading Edge** 27, 394-397.
- Hagan, M.T., Demuth, H.B., and Beale, M.H., 1996. **Neural Network Design**, PWS Publishing Boston, MA, USA, 736 pp. ISBN: 0971732108.
- Haykin, S., 1999. **Neural networks: a comprehensive foundation**. Prentice Hall, Upper Saddle River, NJ, USA, 842 pp. ISBN: 0132733501.
- Novak, L.M., Owirka, G. J., and Netishen, C. M., 1993. Performance of a high-resolution polarimetric SAR automatic target recognition system. **The Lincoln Laboratory Journal** 6(1), 11–23.
- Marcotte, D., 1996. Fast Variogram Computation With FFT. **Computers and Geosciences** 22, 1175–1186.
- Marmo, R., Amodio, S., Tagliaferri, R., Ferreri, V., and Longo, G., 2005. Textural identification of carbonate rocks by image processing and neural network: Methodology proposal and examples. **Computers & Geosciences** 31, 649-659.
- Miranda, F.P., Fonseca, L.E.N., and Carr, J.R., 1998. Semivariogram textural classification of JERS-1 (Fuyo-1) SAR data obtained over a flooded area of the Amazon rainforest. **International Journal of Remote Sensing** 19: 549–556.
- Miranda, F.P., Marmol, A.M.Q., Pedroso, E.C., Beisl, C.H., Welgan, P., and Morales, L.,M., 2004. Analysis of RADARSAT-1 data for offshore monitoring activities in the Cantarell Complex, Gulf of Mexico, using the unsupervised semivariogram textural classifier (USTC). **Canadian Journal of Remote Sensing** 30(3): 424-436.
- Moysey, S., Knight, R.J., and Jol, H.M., 2006. Texture-based classification of ground-penetrating radar images. **Geophysics** 71(6), K111-K118.
- Nguyen, D., and Widrow. B., 1990. Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights. In: **Proceedings of the International Joint Conference on Neural Networks** 3, Washington, DC, USA, pp. 21–26.
- Plutowski, M., Sakata, S. and White H., 1994. Cross-validation estimates IEQM, In: Cowan, J.D., Tesauro, G., and Alspector, J. (Eds.) **Advances in Neural Information Processing Systems** 6, Morgan Kaufman, San Mateo, CA, USA, pp. 391-398.
- Rumelhart, D. E., Hinton, G. E., and Williams, R. J., 1986. Learning Internal Representations by Error Propagation, In: Rumelhart, D.E. and Mclelland, J.L. (Eds) **Parallel Distributed Processing** 1, MIT Press, Cambridge, MA, USA, pp. 318-362.
- Snyman, J., 2005. **Practical Mathematical Optimization: An Introduction to Basic Optimization Theory and Classical and New Gradient-Based Algorithms**. Springer Publishing, New York, USA, 257 pp. ISBN 0387243488.
- Schröder, R., Puls, J., Hanjsek, I., Jochim, F., Neff, T., Kono, J., Paradella, W. R.; Silva, M. M. Q., Valeriano, D. M., Maycira, P. F. C., 2005. MAPSAR: a small L-band SAR mission for land observation. **Acta Astronautica** 56: 35 – 43.
- Tercier, P., R. J.Knight, and H. M. Jol, 2000. Acomparison of the correlation structure in GPR images of deltaic and barrier-spit depositional environments. **Geophysics** 65, 1142–1153.