

## Análise do potencial de classificação do uso e cobertura do solo por meio de rede neural

Claudio Gelelete<sup>1</sup>  
Carlos Frederico de Sá Volotão<sup>2</sup>

<sup>1</sup> Instituto Militar de Engenharia - IME  
Pça Gen Tibúrcio, 80 – Rio de Janeiro - RJ, Brasil  
cgelelete@yahoo.com.br

<sup>2</sup> 5ª Divisão de Levantamento – Diretoria de Serviço Geográfico do Exército  
Rua Major Daemon, 81 – Rio de Janeiro - RJ, Brasil  
volotao@hotmail.com

**Abstract.** This work use an artificial neural network to distinguish land cover, land use, clouds, shadows and water in a multispectral 7-band Landsat image, compares the results with a maximum likelihood classifier result from the same training set. The neural network gave the best results in generalization and accuracy aspects by visual investigation and uses contextual information.

**Palavras-chave:** redes neurais backpropagation, classificação de imagens, informação contextual.

### 1. Introdução

Nos trabalhos de classificação supervisionada de imagens de sensoriamento remoto tradicionalmente são usados métodos estatísticos e dentre estes, o classificador por máxima verossimilhança é o mais utilizado.

É necessário fazer uma inspeção visual e realizar correções no resultado fornecido pelo classificador. Utilizando editores matriciais, as correções visam fechar polígonos que deveriam ter todos os seus pixels pertencentes a uma mesma classe; definir bordas de classes de forma mais precisa; ou mesmo mudar a classe de todo um polígono, quando este for totalmente confundido com classes cujas amostras de treinamento não estejam claramente separadas das amostras de outras classes.

“Uma nova área de processamento de dados envolve o uso das chamadas técnicas de computação, incluindo as redes neurais artificiais, os algoritmos genéticos e a tecnologia fuzzy. Embora projetadas para tarefas de processamento de dados do mundo real, essas técnicas não penetraram de modo pleno no mundo do sensoriamento remoto (Backer, 2002).”

A utilização de redes neurais feed-forward backpropagation torna-se uma alternativa na busca de métodos mais precisos para a classificação da cobertura do solo.

O desenvolvimento de uma rede única que possa ser utilizada para qualquer sensor e para qualquer tipo de valores radiométricos não parece ser viável, porém redes mais específicas, desenvolvidas para um determinado sensor e para uma determinada distribuição dos dados radiométricos podem ser desenvolvidas no intuito de se alcançar resultados melhores do que os métodos estatísticos.

Em alguns trabalhos, como em Todt et al. (2003) e Salu e Tilton (1993), os autores exemplificam a utilização de redes neurais em trabalhos de classificação de imagens de sensoriamento remoto, mostrando resultados que incentivam a consecução das pesquisas nesta linha.

Heermann e Khazenie (1992) utilizaram três bandas multiespectrais de imagens Landsat para obter 24 entradas binárias correspondentes à concatenação dos dados das três bandas em cada localização de pixels. A saída foi mapeada em cinco classes, correspondentes a: água, nuvem, área urbana, floresta e agricultura. Foram treinadas as seguintes redes: 24/24/3,

24/24/5, 24/30/5 e 24/37/5. Com a verificação visual dos dados concluiu-se que o principal erro ocorreu nos perímetros das nuvens, que foram classificados erroneamente como área urbana.

Bischof et al. (1992) compararam os resultados da aplicação de redes neurais backpropagation de três camadas para a classificação pixel-a-pixel de imagens Landsat com a classificação gaussiana de máxima verossimilhança. É mostrada a capacidade da rede em obter melhor resultado que o classificador de máxima verossimilhança e que informações de textura podem ser integradas à rede sem a definição explícita de uma medida de textura. O trabalho apresentado produz quatro classes de saída: terreno construído, floresta, água e agricultura. A escolha dos dados para treinamento estava voltada para as necessidades da classificação de máxima verossimilhança, o que significa que não estava otimizado para o treinamento da rede.

Para acrescentar informações de textura, em Bischof et al. (1992) foram utilizadas duas janelas na banda 5 de dimensões 5x5 e 7x7, codificando a ativação das 25 ou 49 unidades respectivas, proporcionalmente aos seus níveis de cinza, elevando a camada escondida de 5 para 8 neurônios, o que afetou positivamente o desempenho geral da rede.

Em Queiroz et al. (2004) apresenta-se um breve estudo comparativo entre máxima verossimilhança e rede neural com o uso de imagens do satélite CBERS. A entrada consistia de valores de pixels de três bandas. A saída mapeava as seguintes classes: floresta, água, “não-floresta” e desmatamento. A rede identificou melhor a floresta e a água.

Berberoglu et al. (2000) utiliza a variância e o variograma da geoestatística como medidas de textura e os autores afirmam que informações texturais podem ajudar na discriminação de tipos de cobertura do solo, onde diferenças entre assinaturas espectrais são pequenas. A textura pode ser obtida por tamanhos de janelas diferentes e a abordagem textural resulta em exatidão geral superior Ashish (2002).

O objetivo central deste trabalho foi testar o potencial de uma rede neural, comparando seus resultados com os de um classificador gaussiano de máxima verossimilhança, no que diz respeito a sua capacidade de diferenciar alvos, e habilidade de generalização de alvos que apresentem respostas espectrais diferentes das contidas nas áreas de treinamento, apesar de pertencerem a uma das classes relacionadas com a verdade de campo.

## **2. Classes Propostas**

Para o trabalho foi utilizada uma cena Landsat (227-62) do estado do Pará – Brasil, cuja área já havia sido trabalhada em campo e extraídas verdades terrestres de vários pontos. Apesar destas verdades não serem necessárias para a escolha das áreas de treinamento, foram úteis para a análise dos resultados.

Decidiu-se classificar a cena segundo seis classes genéricas a saber: classe 1 – floresta, abrangendo floresta primária, floresta intermediária e reflorestamento; classe 2 – solo, abrangendo solo arado, desflorestamentos e clareiras; classe 3 – água, abrangendo rios e lagos; classe 4 – cultura, abrangendo culturas em suas diversas fases, campo, capoeira e pasto; classe 5 – nuvens e classe 6 – sombra, basicamente aquelas originadas por nuvens.

## **3. Áreas de Treinamento**

As áreas escolhidas para comporem as amostras de treinamento apresentam aspectos bem característicos das classes que representam e seus pixels possuem respostas bastante uniformes dentro dos padrões esperados como respostas radiométricas dos respectivos alvos. Como a área do trabalho está limitada a uma mesma cena, não havendo interesse de extrapolar o treinamento da rede para outras regiões, foram utilizados os valores de DN

(*digital number*), sem qualquer tipo de correção atmosférica ou reamostragem devido a georreferenciamentos.

Para que a comparação entre resultados fosse isenta de qualquer tendência, as amostras utilizadas para treinamento, teste e validação da rede neural foram exatamente as mesmas amostras utilizadas para treinamento do classificador estatístico.

Todas as áreas retiradas como amostras de treinamento possuem formatos retangulares e tamanhos variados. Classes como floresta e água possuem amostras grandes (2500 pixels) na imagem, no entanto sombra, solo e cultura não se apresentam em áreas extensas, sendo necessário a captura de várias pequenas áreas (entre 200 e 1000 pixels).

As áreas foram escolhidas e armazenadas em arquivos distintos utilizando a ferramenta ROI do software ENVI, valendo-se para isso das respectivas coordenadas de tela. O classificador por máxima verossimilhança utilizado pertence ao mesmo software, bastando então indicar as áreas de treinamento para executar a classificação.

#### **4. Arquivo de Treinamento**

Para a rede neural foi necessário preparar os dados de entrada segundo um arquivo do tipo matriz onde cada linha continha os dados de referentes a um determinado pixel. Desta forma a quantidade de linhas corresponde à quantidade de pixels apresentados à rede para treinamento.

O conjunto de dados referentes a cada pixel foi constituído de 35 valores assim distribuídos: os valores de 1 a 7 correspondem aos valores radiométricos das bandas de 1 a 7 do sensor ETM+. A banda 6 relativa ao infravermelho termal, que possui resolução espacial de 120m, foi reamostrada para 30m, igualando-se a quantidade de pixels das demais bandas, utilizando-se o método do vizinho mais próximo. Os valores relativos às posições de 8 a 14 foram preenchidos com a média do pixel central e seus oito vizinhos que formam uma janela 3x3, para cada banda. Os valores de 15 a 21 são relativos ao desvio padrão para a mesma janela 3x3. Os valores de 22 a 28 são relativos às médias da janela 7x7 e os valores de 29 a 35 relativos ao desvio-padrão da janela 7x7. Devido à estas janelas, os pixels das bordas das áreas de treinamento foram desprezados, sendo considerados apenas para os cálculos de média e desvios-padrão.

#### **5. Dimensionamento da Rede**

A rede neural foi treinada utilizando-se o *software* MATLAB. A primeira camada possui 35 entradas relativas aos valores das bandas, médias e desvios-padrão, conforme descritos anteriormente.

A única camada oculta possui 11 neurônios do tipo tangente hiperbólica. Este número foi encontrado testando-se redes com quantidades variadas de neurônios na camada oculta e, para cada configuração de rede, diferentes valores iniciais para as sinapses, o que configura pontos iniciais diferentes na superfície de erro; objetivando-se com isso fugir de mínimos locais que poderiam influenciar de forma negativa o treinamento. Esta quantidade foi definida por ser a menor onde foi possível obter 100% de acerto com as áreas de treinamento. A semente (valores iniciais das sinapses) foi guardada para que fosse possível repetir o mesmo treinamento sempre que necessário.

A camada de saída possui 6 neurônios do tipo tangente hiperbólica relativos às 6 classes. Cada pixel recebe na saída 6 valores, sendo, teoricamente, cinco valores iguais a “-1” e um valor igual “1”. Desta forma um pixel classificado como solo, receberia “1” na saída relativa à classe solo e “-1” nas saídas relativas às outras classes. Nos pixels onde os cálculos da rede não foram tão conclusivos (obtenção de valores diferentes de “1” e “-1”), o referido pixel foi incluído na classe que apresentou o maior valor. O arquivo de treinamento possui 3483 pixels e os arquivos de teste e validação possuem 1200 pixels cada.

Uma vez a rede treinada (**Figura 2**) e os valores iniciais do treinamento armazenados para que este fosse repetido, foi desenvolvido um programa no ambiente utilizado que calcula, organiza e normaliza os valores do arquivo de entrada, treina a rede segundo a semente guardada e com os pesos sinápticos definidos aplicava-os à imagem original, gerando assim uma classificação de toda a imagem.

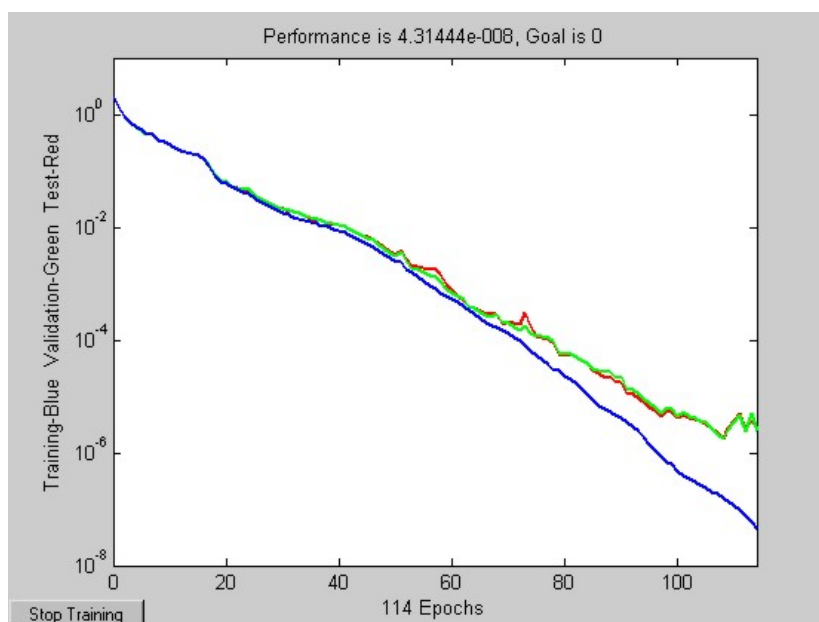


Figura 2. Resultado do Treinamento da Rede.

## 6 . Resultados

Os resultados obtidos foram analisados visualmente, ou seja, de forma qualitativa. As duas classificações foram comparadas com composições coloridas da imagem original, onde foi possível observar as seguintes diferenças entre os resultados descritos neste item.

A rede neural conseguiu generalizar e identificar melhor os rios da região. A existência de dois rios de águas diferentes (um de água mais límpida e outro de água barrenta) confundiu o classificador estatístico, **Figura 3**. As amostras de treinamento foram pegadas apenas no rio de águas límpidas, não sendo apresentadas aos classificadores amostras da água barrenta. O classificador de máxima verossimilhança confundiu o rio de águas barrentas com sombras, mas a rede neural o identificou perfeitamente na classe água.

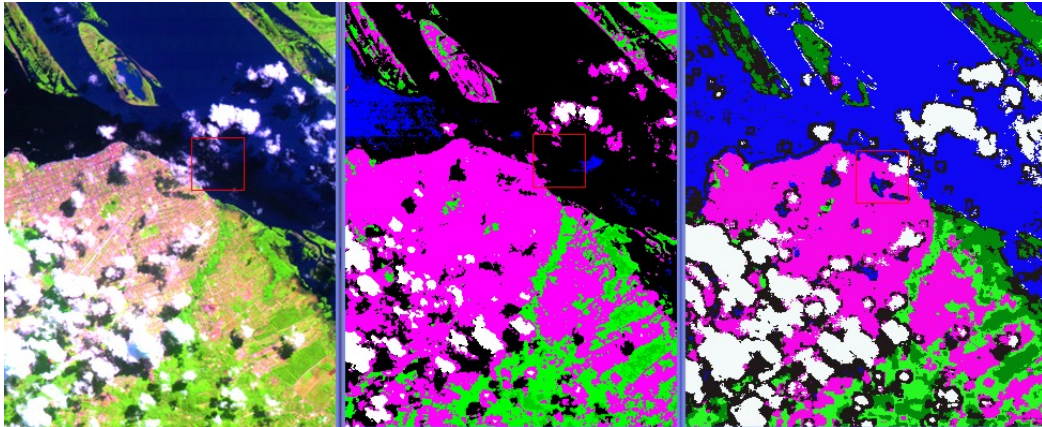


Figura 3. Resultado área 1, na seqüência: imagem, maxver e rede neural.

Outra troca de classes freqüente do classificador estatístico foi observado com a classificação de nuvens como solo, **Figura 3 e Figura 4**, semelhantemente com o ocorrido em Heermann e Khazenie (1992).

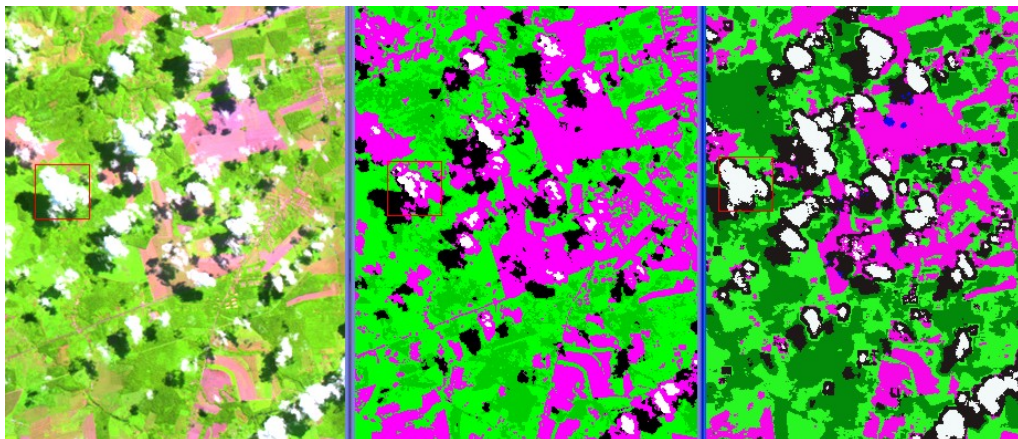


Figura 4. Resultado área 3, na seqüência: imagem, maxver e rede neural.

A rede neural reduziu significativamente a quantidade de classes de pequenas dimensões, **Figura 5**, como pequenos desmatamentos ou áreas agrícolas reduzidas e não identificou feições estreitas como estradas, **Figura 4**. Acredita-se que este erro tenha sido provocado pelas entradas relativas às janelas 3x3 e 7x7 principalmente.

Nenhum dos dois classificadores identificou os rios estreitos da região, confundindo-os como sombra, **Figura 5**.

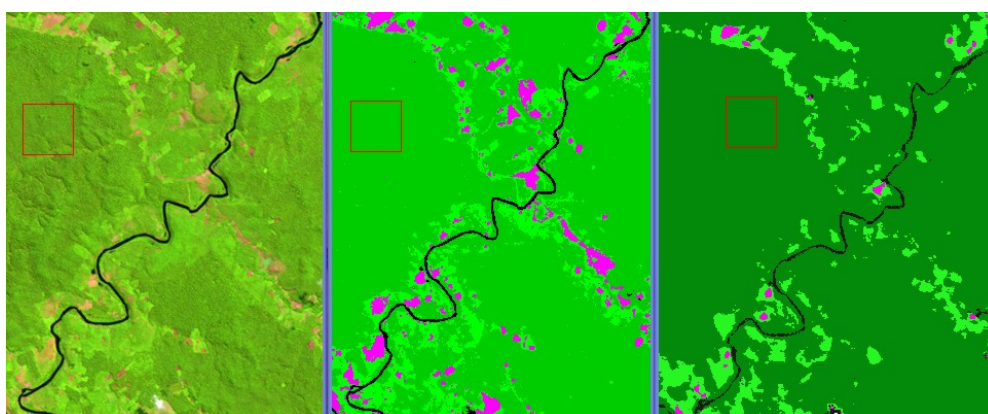


Figura 5. Resultado área 2, na seqüência: imagem, maxver e rede neural.

A **Figura 6** corresponde à imagem utilizada (esquerda) e os locais onde o sinal da classe predominante aproximou-se mais de alguma outra classe (direita), indicando as áreas onde a imprecisão da classificação da rede foi maior. As áreas com maior intensidade de vermelho apresentam os pontos de maior dificuldade para a definição da classe. Foi utilizada a soma dos quadrados das diferenças entre os valores de saída apresentados pela rede e os valores codificados “1” e “-1”. Os valores de imprecisão apontam para o rio, apesar da rede não tê-lo confundido com outra classe.

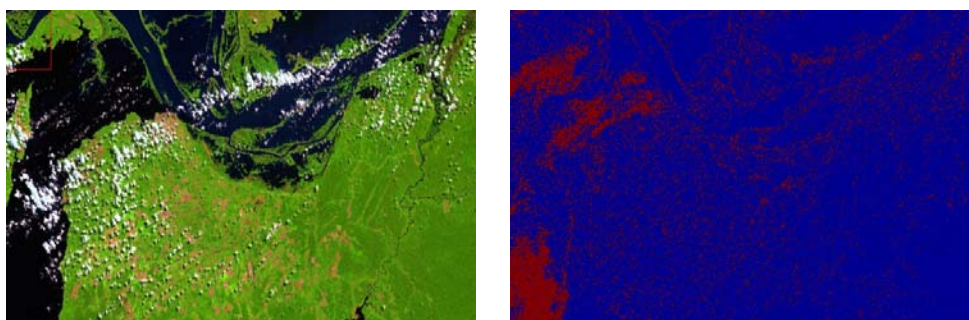


Figura 6. Imprecisão da Classificação.

## 7. Conclusões

Da forma que foi treinada, a rede apresentou dificuldades na classificação de feições estreitas, ou de pequenas dimensões, o que não significa necessariamente ser um problema. É comum ignorar este tipo de feição em alguns casos, porém pode ser necessário em outros. A sensibilidade pode ser mais trabalhada para atender às diversas finalidades, modificando-se o tamanho dos “quadrados”.

Houve menor ocorrência de confusão entre classes na classificação por Redes Neurais, apesar de algumas sombras terem sido classificadas como água. Este fato deixa claro a potencialidade das redes Neurais em identificar as classes com boa precisão, superando resultados obtidos com o classificador de Máxima Verossimilhança.

Os erros gerados pela Rede Neural podem ser trabalhados ajustando-se os valores de entrada e treinando-se a rede com mais áreas de treinamento, na busca por resultados melhores. Para certos casos, o desenvolvimento de redes associadas pode apresentar resultados ainda mais consistentes.

## Referências

- Ashish, D. **Land-use classification of aerial images using artificial neural networks**, University of Georgia, Master Thesis. Athens, 2002.
- Backer, S. D. **Unsupervised pattern recognition**. 2002. 152 p. Doctoral Thesis – University of Antwerp, Antwerp. 2002.
- Berberoglu, S.; Lloyd, C. D.; Atkinson, P. M.; Curran, P. J. The integration of spectral and textural information using neural networks for land cover mapping in the Mediterranean, Elsevier Science Ltd, **Computers and Geosciences**, v. 26, pp. 385-396, 2000.
- Bischof, H.; Schneider, W.; Pinz, A. J. Multispectral classification of Landsat-images using neural networks, **IEEE Trans. Geosc. and Rem. Sensing**, v. 30, n. 3, pp. 482-490, May 1992.
- Heermann, P. D.; Khazenie, N. Classification of multispectral remote sensing data using a back-propagation neural network, **IEEE Trans. Geosc. and Rem. Sensing**, v. 30, n. 1, pp. 81-88, January 1992.
- Queiroz, R. B.; Rodrigues, A. G.; Gómez, A. T. Estudo comparativo entre as técnicas máxima verossimilhança gaussiana e redes neurais na classificação de imagens IR-MSS CBERS 1", **WORKCOMP 2004**, 2004.
- Salu, Y.; Tilton, J. Classification of multispectral image data by the binary diamond neural network and by nonparametric, pixel-by-pixel methods, **IEEE Trans. Geosc. and Rem. Sensing**, v. 31, n. 3, May 1993.
- Todt, V. D. ; Formaggio, A. R.; Shimabukuro, Y. "Identificação de áreas desflorestadas na Amazônia através de uma rede neural artificial utilizando imagens fração derivadas dos dados do IR-MSS/CBERS. In: XI Simpósio Brasileiro de Sensoriamento Remoto, 2003, **Anais**. São José dos Campos: INPE, 2003.