

## Estudo da configuração de diferentes arquiteturas de redes neurais artificiais MLP para classificação de imagens ópticas

Livia dos Santos Abdalla<sup>1</sup>  
Carlos Frederico de Sá Volotão<sup>1</sup>

<sup>1</sup> Instituto Militar de Engenharia – IME - Departamento de Engenharia Cartográfica  
Praça Gen. Tibúrcio, nº 80 Urca, Rio de Janeiro - RJ, 22290-270  
{abdalla.livia@gmail.com, volotao@ime.eb.br}

### Abstract

Artificial neural networks (ANN) can be used to produce several products, including remote sensing classification images. ANN may not beat the performance of traditional classification methods, but it is unique in the sense that: 1) it is not dependent on the prior knowledge of statistical model of data; and 2) it makes it possible to add unusual information by the configuration of parameters, input, hidden and output layers. Motivated by the ability to add different levels of information, including spatial and non-spatial data (e.g., Digital Terrain Models, time, date, a given classification or a segmented image), and comparing to classical methods of classification, this work test the use of ANN for image classification. Despite this capability, this work aims to compare the plain classification ability, by means of *kappa* values and training sets as a reference example when there is no ground truth available. Providing a fare alternative for image classification, the advantages of the potential enhancements are to be studied in future papers. This study explores simple architectures of MLP to identify common themes of land cover and uses, and is based on HRG/SPOT5 images. The results using *kappa* was 91% indicating that the RNA has achieved a good index of training.

**Palavras chaves:** remote sensing, artificial neural network, automatic classification, *kappa* index, sensoriamento remoto, redes neurais artificiais, classificação automática, índice *kappa*.

### 1.Introdução

Existem diversos métodos para o tratamento de imagens digitais de sensoriamento remoto que permitem realizar tarefas de manipulação, análise e compreensão de cenas. O processamento de imagens utiliza pixels como unidade de entrada e saída. A classificação de imagens utiliza pixels como entrada, faz uma análise dos pixels e gera, como saída, uma nova imagem onde existem agrupamentos que compõem classes e pode haver informações sobre as classes (ZHANG, 2009). Desse modo, tanto a entrada como a saída são imagens e a finalidade principal desse processo é obter dos dados as informações mais facilmente discerníveis e identificáveis.

As técnicas de processamento de imagens, segundo NOVO (2010), podem ser categorizadas em três grupos: técnicas de pré-processamento, técnicas de realce e técnicas de classificação. As técnicas de classificação consistem no estabelecimento de um processo de decisão, no qual um grupo de pixels é definido como pertencente a uma determinada classe. Na classificação de imagens por processamento digital, algoritmos computacionais são utilizados na identificação automática dos alvos e são baseados em classificadores *pixel a pixel* ou por regiões.

A classificação pode ser executada com ou sem revisão. No modo não-supervisionado, a classificação é feita com base em métodos estatísticos de análise de agrupamento e, durante o procedimento, não há intervenção direta do intérprete na definição das classes pertinentes a cena. O modo supervisionado, por sua vez, é baseado na coleta de amostras denominadas áreas de treinamento, referentes a classes definidas pelo usuário. Os algoritmos de classificação, tanto determinísticos como os probabilísticos, utilizam-se das amostras fornecidas pelo usuário para rotular os *pixels* de toda a imagem como pertencente a um determinado uso ou cobertura do solo.

Na classificação supervisionada, a utilização de redes neurais artificiais (RNA) pode ser uma alternativa mais vantajosa que utilizar métodos estatísticos em certos casos, como quando não há informação prévia sobre a distribuição dos dados ou há possibilidade de incorporar dados adicionais, não espectrais, à classificação (CRUZ & GALO, 2003). Como o uso de métodos baseados no modelo de distribuição da imagem já têm seu uso consagrado pela comunidade, o uso de RNA, pela flexibilidade de permitir adicionar informações e pela disponibilização de em tem sido cada vez mais

Nesse contexto, o objetivo do presente trabalho é definir um conjunto de parâmetros que viabilize uma arquitetura de RNA que adequadamente identifique diferentes feições de uso e cobertura do solo de uma imagem do sensor HRG/SPOT5.

## 2. Redes Neurais Artificiais

Segundo BRONDINO (1999), as RNAs baseiam-se em modelos abstratos do funcionamento do cérebro humano e buscam reproduzir sistemas biologicamente realísticos. O sistema de processamento de informações de uma RNA se assemelha à estrutura do cérebro humano, pois adquire e armazena conhecimentos através do processo de aprendizado, que se dá pelas conexões entre neurônios, o que também são conhecidas como sinapses. Os neurônios interconectados trabalham em paralelo para classificar e generalizar dados de entrada em classes de saída (RIBEIRO & CENTENO, 2000).

Segundo BRONDINO (1999), esta organização das camadas de uma RNA (figura 1) que definem sua estrutura topológica ou sua arquitetura (maneira como os elementos de processamento são organizados). Assim sendo, existem redes neurais de camada simples constituídas por um grupo de neurônios (*perceptrons*) arranjados em apenas uma camada, e as redes multicamadas, formadas por várias camadas escondidas ou pela combinação de várias redes de camadas simples. Nessa estrutura, a camada de entrada é aquela na qual os padrões são apresentados à rede; as camadas intermediárias (também chamadas de ocultas ou escondidas) são responsáveis por grande parte do processamento, podendo ser consideradas como extratoras de características; e a camada de saída é aquela onde o resultado final é apresentado.

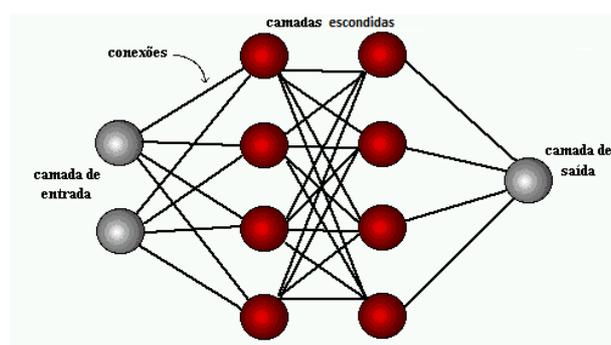


Figura 1: Arquitetura de uma RNA com duas camadas escondidas

Uma RNA é formada por um conjunto de elementos processadores simples, uma rede de interconexão e uma regra de aprendizado, sendo o processamento feito de forma paralela por cada um dos nós da rede. Assim, as RNAs conhecidas como perceptron multicamadas (MLP, do inglês *Multi-Layer Perceptron*) são organizadas em camadas conectadas: camada de entrada, uma ou mais camadas intermediárias e camada de saída. Na camada de entrada não ocorre processamento e os dados já preparados são apenas apresentados ao classificador, esta camada se comunica a uma ou mais camadas ocultas onde ocorre o processamento por meio de um sistema de conexões ponderadas. O processo de aprendizagem se concentra nas camadas ocultas da rede. A camada de saída apresenta o resultado do processamento.

É feita uma atribuição de pesos aos dados de entrada, o que irá determinar o nível de ativação da unidade na rede. A magnitude da rede é determinada por um processo de treinamento repetitivo, o qual busca a repetição com o intuito de convergir pra um esperado valor de saída conhecido do conjunto de treinamento para cada amostra definida.

O algoritmo de aprendizado mais utilizado para o seu treinamento é o algoritmo retropropagação (*backpropagation*) (RUMELHART et al., 1986 apud PICOLI, 2006). Seu treinamento é supervisionado e baseia-se numa regra de aprendizagem que compensa o erro durante o treinamento (BOCANEGRA, 2002). Na correção, o erro é retroalimentado por meio da rede e o ajuste dos pesos é feito usando um método recursivo, realinhando em cada iteração os resultados até que num futuro os resultados se aproximem do esperado numa certa tolerância pré-definida.

O algoritmo *backpropagation* permite o ajuste de pesos em cada uma das camadas da rede, sendo projetado para minimizar a soma do erro quadrático entre a saída calculada por uma estrutura multicamadas e a desejada. Durante o treinamento, a intenção é que este erro diminua, mas para que isso ocorra, é necessário que os pesos sejam iterativamente ajustados, como explicado. Os primeiros pesos a serem ajustados são os da camada de saída, posteriormente os pesos das camadas anteriores serão ajustados sequencialmente.

Um resultado significativo requer um ótimo conjunto de treinamento e que o processo de classificação necessita de um ajustamento de vários parâmetros, o que pode restringir sua aplicabilidade (BLASCHKE et al., 2000).

### **3. Metodologia**

#### **3.1. Material**

A imagem multiespectral utilizada para a realização da classificação automática com RNA foi a do satélite SPOT 5 (figura 2), sensor HRG (*High-Resolution Geometric*).

As imagens do satélite SPOT possuem aplicações em diversas áreas científicas e comerciais, no monitoramento de fenômenos e recursos naturais, acompanhamento do uso agrícola das terras, apoio ao monitoramento e definição de áreas de preservação, na atualização de mapas e cartas, entre outros.

As imagens adquiridas no modo multiespectral oferecem resolução espacial de 10 metros (com exceção da banda do infravermelho médio com resolução de 20 metros); e a banda pancromática com 5 metros.



Figura 2: Composição colorida da imagem SPOT 5

O processamento da imagem foi conduzido no aplicativo IDRISI, cujas funcionalidades são direcionadas às aplicações em SIG e Processamento Digital de Imagens, e que utiliza como entrada as bandas espectrais e canais oriundos de quaisquer origens. O programa incorpora muitas funcionalidades analíticas para imagens, consulta a banco de dados e modelagem espacial (EASTMAN, 1998).

### 3.2.Arquitetura e Treinamento da Rede Neural

As classes de saída foram definidas com base em cinco classes classificadas visualmente e foram compostas da seguinte forma: corpos d' água, culturas diversas, floresta, campo e urbano.

A obtenção das áreas de treinamento foi realizada através da extração de áreas em formato de polígonos da imagem para cada classe a ser classificada, sendo no presente trabalho selecionadas entre 5 a 7 polígonos de tamanhos variados para cada classe. RICHARDS (1993) indica um mínimo de  $(N+1)$  amostras para cada classe, sendo o  $N$  o número de bandas espectrais que integram a imagem utilizada no processo de classificação.

Após a aquisição das áreas de treinamento, foi calculada a resposta espectral média do conjunto de treinamento, no qual o programa associa cada pixel da imagem à assinatura mais similar (módulo "Makesig"). Os conjuntos de treinamento têm que ser definidos num plano de informação vetorial ou como imagem e têm que possuir um identificador inteiro. Para que o referido módulo possa funcionar, necessita-se indicar também as bandas que poderão ser usadas na classificação. Neste trabalho foram utilizadas 4 bandas (onde ocorreu uma prévia fusão IHS na banda do infravermelho médio).

Após a realização da seleção e estudo das classes a serem classificadas, que trata da distribuição de cada elemento da imagem a uma das classes previamente selecionadas, foi definida uma arquitetura para a RNA classificar as imagens. Esse procedimento foi executado utilizando o módulo de classificação por redes neurais artificiais do tipo perceptron multicamadas (MLP), do software IDRISI.

Diversos trabalhos descrevem os aspectos teóricos ou a aplicação de RNA na classificação de dados de sensoriamento remoto, e buscando uma classificação

supervisionada, foi adotado um modelo básico de rede neural com estrutura de múltiplas camadas, treinada com algoritmo *backpropagation*.

A definição da arquitetura de uma MLP implica na determinação prévia da dimensionalidade dos dados de entrada e das classes de ocorrência. O número de elementos da camada de saída foi definido com base no número de classes a serem classificadas na imagem, sendo cinco nós, com valores reconhecidos em virtude do conjunto de treinamento estabelecido nas amostras. Para dimensionar o número de camadas escondidas e o número de nós que constituem cada camada intermediária, foram realizadas simulações com diferentes arquiteturas de redes neurais. Essas simulações constaram, na verdade, do treinamento e análise do desempenho da rede treinada sobre a análise do índice estatístico *kappa*, conforme descrito por LANDS e KOCH (1977), para cada arquitetura testada. O índice é apresentado após a classificação pelo módulo em que a rede neural está implementado, resultando no *kappa* da classificação em relação ao conjunto de treinamento.

#### 4. Resultados e Discussões

Como o presente trabalho surgiu da necessidade de estudar as vantagens de utilização de um método pouco explorado que, embora não seja novo, não tem uma grande utilização no meio dos usuários de sensoriamento remoto como poderia, a despeito da disponibilização crescente nos aplicativos especializados na área. O trabalho de VOLOTÃO e TEIXEIRA (2007) exemplifica um uso diferenciado do modelo para detectar bordas e classificar imagens, indicando o grande potencial que a utilização de MLP como ferramenta de processamento de imagens de sensoriamento remoto.

Tabela 1: Alguns parâmetros utilizados na realização dos testes

n	Pixels <sup>1</sup>	Camadas Escondidas	Neurônios		Taxa de Aprendizado	Momentum <sup>2</sup>	Iterações	RMS	
			1ª	2ª Camada				Treinamento	Teste
1	320	2	10	3	0.01	0.5	10000	0.001880	0.001958
2	320	1	10	-	0.01	0.5	10000	0.001565	0.001591
3	360	2	15	5	0.01	0.5	10000	0.001051	0.001125
4	360	2	15	5	0.02	0.53	10000	0.001049	0.001129
5	360	2	15	10	0.02	0.53	10000	0.001137	0.001197

(<sup>1</sup> Número de pixels selecionados para o conjunto de treinamento e teste; <sup>2</sup> durante a busca para o valor mínimo sobre a superfície de erro evita problemas de oscilação e é usado para acelerar o processo de convergência)

Este trabalho visa, assim, servir como referência básica com um nível de detalhamento e simplicidade que estimulem a utilização desse modelo. Algumas simplificações devem ser destacadas: 1) não utilização de todos os parâmetros e recursos possíveis numa rede neural; 2) um número exaustivo de variantes das configurações possíveis do MLP e todos os seus parâmetros; 2) foi utilizada apenas produto de sensor óptico; 3) não foi realizado o teste com grande número de imagens variadas e de diferentes regiões geográficas; 4) o estudo limitou-se aos cinco temas apresentados; e 5) apenas utilizou-se como entrada as informações pixel-a-pixel das bandas multiespectrais, sem informação de contexto; e 6) não foi realizada verificação de campo nem utilizada uma verdade de campo.

Para definição da arquitetura ideal para classificação foram realizados trinta testes, no qual foram realizadas as seguintes alterações na arquitetura da rede: número de camadas; número de neurônios nas camadas; número do conjunto de pixels para o conjunto de treinamento e teste; fator de *momentum* e taxa de aprendizado. A tabela 1 apresenta a experimentação dos cinco melhores testes.

Tabela 2: Matriz de Confusão entre as Classes

	Massas d'Água	Culturas	Floresta	Campo	Urbano	Total	Erro de Comissão
Massas d'Água	722	0	1	7	235	965	25,18%
Culturas	0	1502	0	8	307	1817	17,34%
Floresta	6	0	4221	6	183	4416	4,42%
Campo	0	0	19	3468	167	3654	5,09%
Urbano	2	50	0	1	9253	9306	0,57%
Total	730	1552	4241	3490	10145	20989	
Erro de Omissão	1,10%	3,22%	0,47%	0,63%	8,79%		

Nesse sentido, duas importantes observações devem ser pontuadas: o aumento do número de camadas produziu melhores resultados, sendo o número de neurônios da primeira camada superior ao da segunda camada. Mas uma terceira camada não seria justificada pela simplicidade dos fenômenos envolvidos, além de exigir um número muito maior de amostras de treinamento, teste e validação.

Dentre as experimentações apresentadas na tabela, o 4º teste apresentou resultados mais refinados que os demais, sendo possível chegar a esta constatação a partir da análise da matriz de confusão dessa classificação (tabela 2), que representa a distribuição de percentagens de pixels classificados correta ou erroneamente (CONGALTON, 1991).

A partir da matriz de confusão, o MLP apresentou o índice *kappa* para validar o que foi constatado visualmente nas classificações. Este índice foi calculado em relação às amostras de treinamento, e obteve o valor de 0,91 o que segundo LANDIS e KOCH (1977) é um valor considerado “excelente”, portanto, satisfatório para este estudo.

Neste trabalho utilizou-se o índice *kappa*, para avaliar a acurácia temática por ser mais sensível as variações de erros de omissão e inclusão (COHEN, 1960) e pelo seu uso já consagrado pelos especialistas da área. O seu cálculo utiliza, além dos elementos da diagonal principal, todos os elementos da matriz de erro (MOREIRA, 2001). A cena classificada, resultante da aplicação da arquitetura treinada pela rede neural é apresentada na figura 3.

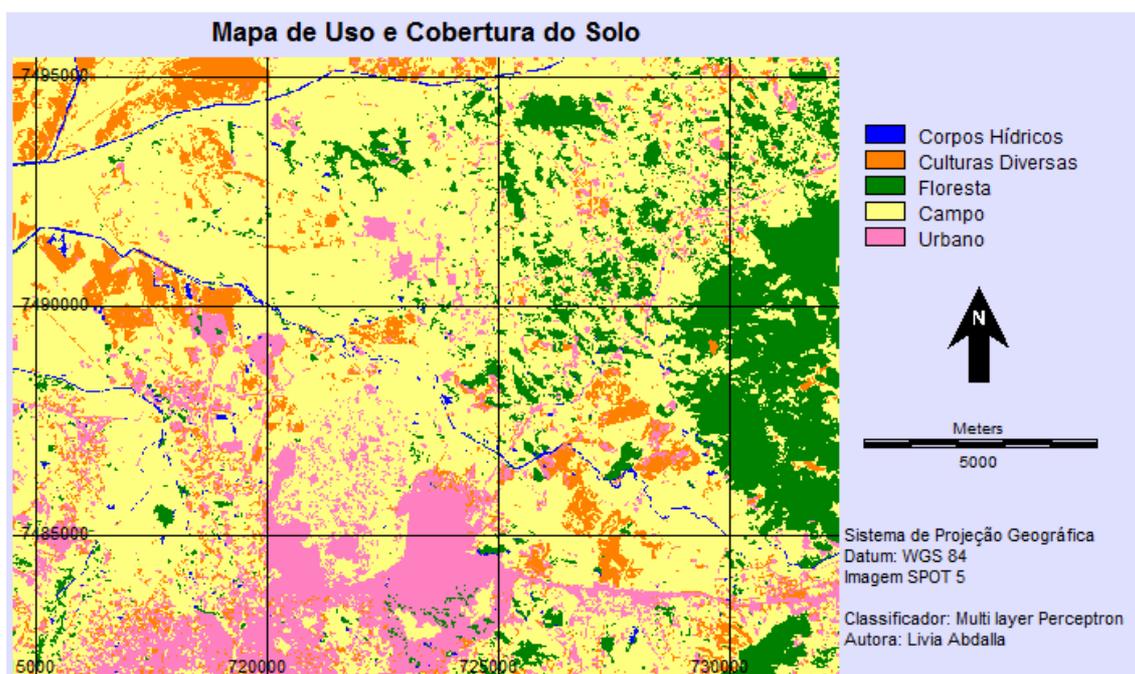


Figura 3: Resultado da classificação baseada na arquitetura de redes neurais artificiais

## 5. Considerações Finais

Neste trabalho, foi realizada a classificação de imagens de sensoriamento remoto por uma RNA do tipo MLP com o propósito de definir um conjunto de parâmetros que melhor identificassem as diferentes feições de uso e cobertura do solo para uma aplicação simples, estabelecendo um parâmetro inicial. A adição futura de outros tipos de dados vai exigir um estudo específico de parâmetros, conforme feito neste trabalho, mas exigindo uma busca mais completa e focalizada dos parâmetros.

Para os dados utilizados, o conjunto de parâmetros que apresentou melhor resultado, após diversos testes, continha duas camadas escondidas, com quinze neurônios na primeira e cinco na segunda camada, e taxa de aprendizado de 0,02 e fator de *momentum* de 0,53.

Não dispondo do mapa verdade da região, o treinamento foi feito pelo algoritmo MLP, onde os índices *kappa* foram obtidos a partir da rede neural treinada e das amostras de treinamento. Mesmo com o relevante valor do índice *kappa*, houve pequenos erros de classificação observados, os quais se justificam por regiões de mistura e fatores ambientais, como o relevo acidentado que causa sombra e culturas intercalares que interferem no padrão espectral da imagem. Para refinar o processo e corrigir esses erros pontuais, a suavização por um filtro que utiliza vizinhança do pixel no contexto da classe vizinha possibilitaria, em tese, uma classificação com menor índice de confusão.

É importante salientar a existência do caráter aleatório de inicialização do conjunto de pesos do processo de MLP, o que não garante a repetibilidade do resultado, a menos que os valores das sementes iniciais sejam registrados. O valor de *kappa* pode não ser atingido mas pode ser obtido valores ainda melhores em novas instâncias, devido ao procedimento de treinamento da rede. Assim, a apresentação do índice de desempenho da classificação é um indicador de que aquela configuração específica é satisfatória.

Apesar dessa instabilidade intrínseca ao método, sua flexibilidade na possibilidade de integração e adição de novos dados, além da possibilidade constatada neste trabalho de obtenção de resultados aceitáveis, o torna um método que precisa ser muito mais explorado pela comunidade de sensoriamento.

## Agradecimento

Os autores agradecem a UFF pelas imagens disponibilizadas pelo laboratório de Geografia Física da UFF (LAGEF/UFF).

## Referências Bibliográficas

BOCANEGRA, C.W.R. Procedimentos para tornar mais Efetivo o Uso das Redes Neurais Artificiais em Planejamento de Transportes. 2002, 146p. Dissertação (Mestrado em Engenharia Civil – Transportes) – Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos. 2003.

BLASCHKE, T.; LANG, L.; LORUP, E.; STROBL, J.; ZEIL P. **Object-Oriented Image Processing in an Integrated GIS/ Remote Sensing Environment and Perspectives for Environmental Applications**. Vol. 2, 555-570, 2000.

BRONDINO, N.C.M. Estudo da Influência da Acessibilidade no Valor de Lotes Urbanos Através do Uso de Redes Neurais. 1999, 97p. Dissertação (Doutorado em Engenharia Civil – Transportes) – Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos. 1999.

COHEN, J. A coefficient of agreement for nominal scales. **Educational and Psychological Measurement**, v. 20, n. 1, p. 37-46, 1960.

CRUZ, N. C.; GALO, M. L. B. T. Aplicação de redes neurais artificiais na classificação de imagens multitemporais e multisensor. In: Colóquio Brasileiro de Ciências Geodésicas, 3., 2003, Curitiba. **Anais**. Curitiba: UFPR, 2003. CD-ROM.

EASTMAN, R. J. Tradução Português. Hasenack, H.; Weber, E. IDRISI for Windows versão 2. **Manual do Usuário**. Versão Digital. Introdução; Exercícios Tutoriais. Porto Alegre: Centro de Recursos Idrisi, Brasil, 1998. 218p.

EMBRAPA. **Monitoramento por Satélites**. 2012. Disponível em:  
<http://www.sat.cnpm.embrapa.br/conteudo/spot.htm>. Acesso em: 15/09/2012.

LANDIS, J.R.; KOCH, G. G. **The measurement of observer agreement for categorical data**. Biometrics. New York, v. 33, n.1, p.159-174, 1977.

MOREIRA, Maurício Alves, **Fundamentos de Sensoriamento Remoto**. São José dos campos-SP: INPE, 2001

NOVO, E. M. L. M. **Sensoriamento remoto: Princípios e aplicações**. Editora Edgard Blücher LTDA. São Paulo, SP. 2010.

PICOLI, M. C. A. Estimativa da produtividade agrícola da cana-de-açúcar utilizando agregados de redes neurais artificiais: estudo de caso usina Catanduva. 2006. 90p. (INPE-14669-TDI/1222). Dissertação (Mestrado em Sensoriamento Remoto), Instituto Nacional de Pesquisas Espaciais, São José dos Campos: INPE, 2006.

RIBEIRO, S. R. A.; CENTENO J. A. S. **Clasificación del uso del suelo utilizando redes neurales y el algoritmo maxver**. Innovación, Año 14 n.1, p.55-60. Antofagasta – Chile, 2002.

RICHARDS, J. A. **Remote Sensing Digital Image Analyses: an introduction**. 2. ed. Berlin: Springer-Verlag, 1993. 340p.

VOLOTÃO, C.F.S.; GELELETE, C. Linear features detection in CCD/CBERS-2 image using neural network. In: XIII Simpósio Brasileiro de Sensoriamento Remoto, pp.1205-1210, São José dos Campos, **Anais**. INPE, 2007.