

**AVALIAÇÃO DO TAMANHO DE AMOSTRAS DE TREINAMENTO
PARA REDES NEURAIS ARTIFICIAIS NA CLASSIFICAÇÃO
SUPERVISIONADA DE IMAGENS UTILIZANDO DADOS ESPECTRAIS
E LASER SCANNER**

*Assessment of training sample size for artificial neural networks in supervised
image classification using spectral and laser scanner data*

Gabriel Henrique de Almeida Pereira

Jorge Antonio Silva Centeno

Universidade Federal do Paraná – UFPR. Pós-Graduação em Ciências Geodésicas - Departamento de
Geomática. Caixa Postal 19001, 81.531-990 Curitiba, Paraná, Brasil.

Email: pereira.gha@hotmail.com; centeno@ufpr.br

Resumo:

Técnicas de Sensoriamento Remoto tem ganhado especial interesse, uma vez que podem ser utilizadas para o monitoramento de sistemas e fenômenos em escala local ou global, de maneira contínua temporal e espacialmente. Redes Neurais Artificiais estão entre os métodos que são capazes de trabalhar com grande quantidade de dados, com diversas características e sofrer pouca influência de ruídos. Desta forma, fez-se o uso de Redes Neurais Artificiais com o propósito de classificar dados de sensoriamento remoto. Utilizou-se de dados de alta resolução espacial, como imagens espectrais de aerolevanteamento e dados altimétricos *Laser Scanner*, para a classificação do alvo “árvores”. Com isso, gerou-se RNA especialistas na detecção destes alvos. Os dados utilizados são de uma área densamente urbanizada, onde existe grande variabilidade de cotas e características espectrais. Os resultados mostraram que a classificação utilizando dados espectrais e altimétricos resultaram em melhores classificações, do que a utilização apenas de informações espectrais. Testou-se também a influência do tamanho das amostras de treinamento das Redes Neurais Artificiais, gerando assim uma “curva de aprendizado” das RNA. Percebeu-se que conforme se aumenta o tamanho das amostras de treinamento, existe uma tendência em aumentar a acurácia na classificação dos dados. Os acertos globais foram superiores a 87,5% quando utilizando apenas informação espectral e 97,5% quando utilizando dados espectrais e altimétricos.

Palavra Chave: Sensoriamento Remoto, Classificação Digital de Imagens, Redes Neurais Artificiais, Imagens de Alta Resolução Espacial, *Laser Scanner*.

Abstract:

Remote Sensing techniques has gained special interest, since it can be used for monitoring systems and phenomena in local or global scale, in a temporally and spatially continuous way. Artificial Neural Networks are able to work with large amounts of data, with different

characteristics. ANN was used in this work as the purpose of classifying remote sensing data. It was used multi sources and high-resolution spatial data, such as spectral images and Laser Scanner elevation data to classify the class “tree”. So, all the ANN created were specialist in tree class classification. In addition, the data used is from a densely urbanized area where there is a large variability of elevations and spectral characteristics. The results showed that the classification using both spectral and altimetry data resulted in better classifications than the use of only spectral information. It was also tested the influence of the size of samples for training the ANN, creating a learning curve for the ANN. It was noticed that with increasing the size of training samples there is a tendency to increase the accuracy in the classification. The global hits were above 87.5% when using only spectral data, and 97.5% when using spectral and altimetry data.

Keywords: Remote Sensing, Digital Image Classification, Artificial Neural Networks, High Spatial Resolution Images, Laser Scanner.

1. Introdução

A classificação digital de imagens de sensoriamento remoto tem ganhado, nas últimas décadas, reconhecimento como técnica para suprir a escassez dos mapeamentos temáticos. A extração automatizada de informações demanda métodos que consigam reproduzir a capacidade humana de interpretar uma imagem. Neste contexto, ganha significado a integração de novos métodos de análise como, por exemplo, o uso de técnicas de inteligência artificial, como as Redes Neurais Artificiais (RNA), para representar o conhecimento do analista humano para efetuar a interpretação e classificação de imagens (Araki, 2005).

Segundo Centeno (2009), uma maneira de tornar a interpretação de imagens mais simples consiste em separar grupos de pixels com características espectrais similares. Partindo da hipótese de que pixels cobrindo um mesmo alvo devem ter características espectrais similares, algoritmos computacionais podem ser usados para dividir os pixels de uma imagem em grupos, segundo suas leituras em cada banda espectral. Este processo é denominado classificação digital de imagens e consiste em, a partir dos valores do contador digital de uma imagem, gerar uma nova imagem, mais simples, onde cada pixel está associado a uma categoria, ou classe.

Os classificadores podem ser enquadrados em dois grupos: os não supervisionados e os supervisionados. Dentre os métodos não supervisionados estão: IsoData e K-médias; e dentre os supervisionados estão: paralelepípedo, mínima distância, Máxima Verossimilhança, mapeamento por ângulo espectral, árvores de decisão, Redes Neurais Artificiais e, mais recentemente, o *Support Vector Machines* (SVM, Máquinas de Vetor de Suporte).

A utilização de Redes Neurais Artificiais (RNA) emerge como uma alternativa vantajosa até mesmo aos classificadores de imagens baseados em conceitos estatísticos, como o da Máxima Verossimilhança. Isto porque nenhuma hipótese prévia sobre a distribuição dos dados a serem classificados é exigida e os dados de entrada podem ter escalas diferentes. Sua capacidade de distribuição dos dados de entrada é mais sutil do que os métodos estatísticos tradicionais e a capacidade de formar fronteiras de decisão não lineares no espaço de características têm aumentado o seu uso (Botelho, 2004).

Estudos descritos na literatura têm mostrado que as técnicas como Redes Neurais Artificiais, árvores de decisão e *Support Vector Machines* podem ser capazes de classificar um conjunto de dados com maior precisão do que os convencionais classificadores estatísticos (Foody e Mathur, 2004). Ainda, Paola e Schowengerdt (1995) ressaltam que as Redes Neurais mostraram-se uma ferramenta útil para sensoriamento remoto e classificação de imagens. Afirmando também que a capacidade flexível da RNA e habilidade de usar conjuntos de treinamento pequenos, fazem da rede neural uma ferramenta em potencial para classificação de imagens multiespectrais de sensoriamento remoto.

Tipicamente, para uma RNA ser aplicada, é necessário tanto um conjunto de treinamento como de teste. Ambos os conjuntos de treinamento e teste contêm pares com informações de entrada e suas respectivas saídas, tomadas a partir de dados reais. A primeira é usada para treinar a rede, e a segunda para avaliar o desempenho da rede. Na fase de teste, os dados de entrada alimentam a rede e os padrões de saída desejados são comparados com os obtidos através da rede neural. A concordância, ou não, destes dois conjuntos dá uma indicação do desempenho da rede neural (Lek e Guégan, 1999).

Lek e Guégan (1999) também relatam que o conjunto de treinamento tem ainda de ser suficientemente grande para ser representativo do problema. Desta forma, constituiu a presente pesquisa a fim de analisar as capacidades de Redes Neurais Artificiais para a classificação de imagens analisando-se diferentes tamanhos de amostras de treinamento. Ainda, a fim de se avaliar a capacidade das RNA em trabalhar com dados de diferentes fontes, de características e distribuições diferentes, utilizou-se de dados espectrais de alta resolução espacial e dados altimétricos Laser Scanner Aerotransportado de uma área densamente urbanizada, onde existe grande variabilidade de cotas altimétricas e respostas espectrais dos objetos.

Existe trabalhos na literatura que tratam da utilização de Redes Neurais Artificiais para a classificação de dados Laser Scanner, como Andrade e Centeno, 2003; Botelho, 2004; Höfle, Hollaus, e Hagenauer, 2012; Mendes e Dal Poz, 2013; Xu, Vosselman e Elberink, 2014; Görgens, Montaghi e Rodriguez, 2015. Entretanto, poucos estudos trabalharam na comparação da influência da utilização de dados espectrais ou conjuntamente com a informação altimétrica no processo de classificação utilizando RNA. São ainda mais raros os trabalhos que tratam da comparação desta influência do tipo de dados com a melhora proporcionada pelo aumento do tamanho das amostras para o processo de treinamento das Redes Neurais Artificiais, ou seja, a “curva de aprendizado” da RNA. Com isto, este trabalho enfoca na análise tanto do incremento das amostras de treinamento e geração da “curva de aprendizado” das RNA, como na melhora proporcionada pela inserção da informação altimétrica de Laser Scanner no processo de treinamento das RNA.

2. Redes Neurais Artificiais

Redes Neurais Artificiais têm se tornado o foco de muita atenção, devido a sua ampla aplicabilidade e, principalmente, por tratar de casos considerados complicados. RNA podem identificar e aprender padrões relacionando conjunto de dados de entrada e valores de saída correspondentes. Após o treinamento, RNA podem ser usadas para prever o resultado relacionado a um novo grupo de dados de entrada. Elas podem resolver problemas com dados não lineares e complexos, mesmo sendo dados imprecisos e ruidosos.

Haykin e Lippmann (1994) descreveram RNA como um processador maciçamente paralelo, constituído por unidades de processamento simples, que tem uma tendência natural para armazenar conhecimento empírico e torná-lo disponível para o uso. São poderosas ferramentas para modelagem, especialmente quando são desconhecidas as relações entre os dados (Lek e Guégan, 1999).

Embora existam várias descrições do funcionamento das funções utilizadas, a ideia principal das RNA está no cálculo inspirado no sistema biológico, no cérebro humano. Resumidamente, uma RNA compreende um grande número de unidades de processamento simples, os nós ou neurônios, organizados em camadas, conectados por ligações ponderadas de acordo com uma arquitetura especificada.

Usualmente, as camadas são classificadas em três grupos:

- Camada de Entrada: onde os padrões são apresentados à rede;
- Camadas Ocultas (Intermediárias ou Escondidas): onde é feita a maior parte do processamento, através das conexões ponderadas; podem ser consideradas como extratoras de características;
- Camada de Saída: onde o resultado final é concluído e apresentado.

A Figura 1 apresenta uma representação de RNA, com os neurônios dispostos nas camadas de entrada, camada oculta e camada de saída. O número de neurônios em cada camada pode ser variável, assim como o número de camadas ocultas.

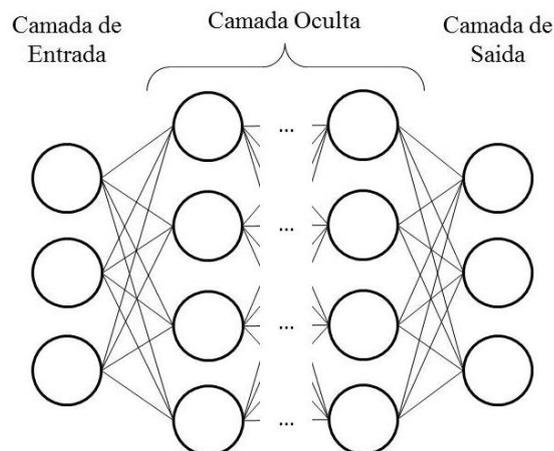


Figura 1: Representação gráfica de uma Rede Neural Artificial.

Em aplicações de sensoriamento remoto, por exemplo, como a classificação de imagens, Petropoulos, Arvanitis e Sigrimis (2012) afirmam que os nós na camada de entrada poderiam ser as bandas espectrais de uma imagem, as características de textura ou outras informações obtidas por sensoriamento remoto. Cada um dos nós de saída representa uma classe. As camadas escondidas são constituídas por vários nós. As ligações ponderadas entre os nós carregam o fluxo de informação através da rede. A aprendizagem ocorre através do ajuste dos pesos das ligações para minimizar a diferença entre o resultado proposto pela rede e o fornecido pelo usuário, ou seja, a classificação correta dos pixels da imagem.

Heermann e Khazenie (1992) afirmam que todos os nós da rede, exceto os de entrada, executam as mesmas funções: recolhem os valores transferidos da camada anterior e definem um resultado de saída. Os nós de entrada são determinados pelos dados de entrada. A função que recolhe os valores da camada anterior é dada pela Equação 1:

$$net_{pi} = \sum_j w_{ij} a_{pj} + bias_i \quad (1)$$

Onde w_{ij} representa o peso da ligação de um nó j da camada anterior com o nó i da camada atual, a_{pj} é o valor passado do nó anterior j do padrão p , e o $bias$ representa um termo independente, podendo ser considerado caso o nó atual i sempre possua ativação.

Vieira (2000) reporta que o algoritmo de aprendizado da rede neural mais utilizado na classificação de imagem oriundas do sensoriamento remoto é o *backpropagation*.

Backpropagation é um algoritmo de gradiente descendente, que minimiza o erro entre a saída obtida pela rede e os resultados fornecidos para o treinamento. Portanto, para esta etapa, um conjunto de pares de dados de entrada e saída é apresentado repetidamente à rede. Para diminuição do erro entre o resultado da rede e o esperado, os pesos vão sendo ajustados conforme a taxa de aprendizado (Pal e Mather, 2004).

Heermann e Khazenie (1992) afirmam que a técnica de rede neural de *backpropagation* parece ser viável para a classificação de imagens de satélite. O algoritmo *backpropagation* reduz a principal desvantagem das redes neurais, que é o tempo de treinamento. Os mesmos autores ainda acrescentam que uma rede *backpropagation* é facilmente modificada para acomodar mais canais ou para incluir informação espacial e temporal. A camada de entrada pode simplesmente ser aumentada para aceitar os dados adicionais.

O treinamento consiste em apresentar dados de entrada com seus respectivos resultados esperados. O objetivo do treinamento é minimizar o erro entre o resultado esperado e o obtido pela rede. Para garantir a diminuição desse erro, os pesos entre as ligações são ajustados a cada iteração.

Heermann e Khazenie (1992) afirma que a fundamentação do algoritmo de aprendizado *backpropagation* é a otimização não-linear, de gradiente descendente, pelo somatório do quadrado da diferença entre os padrões apresentados a rede O_{pi} e o resultado obtido pela rede t_{pi} . O objetivo é então minimizar o erro dado pela Equação 2:

$$E = \sum_p \sum_i (t_{pi} - O_{pi})^2 \quad (2)$$

Onde p são os padrões que se quer que a rede aprenda, e i o número de nós de saída da rede.

O treinamento da rede é concluído quando o erro é menor que um dado limiar, ou um número máximo de iterações definido. É importante evitar o excesso de treinamento, conhecido como *overfitting*, uma vez que a rede pode classificar perfeitamente os dados de treinamento, mas não funcionar com os dados de teste (Chettri, Crompt e Birmingham, 1992).

Chettri, Crompt e Birmingham (1992) comparam as RNA *backpropagation* com um classificador de Máxima Verossimilhança. Alguns dos itens avaliados foram: o tempo de treinamento, exatidão, tempo e sensibilidade da classificação. O nível de exatidão obtido, ou seja, o número de dados classificados corretamente em um conjunto de teste, foi melhor quando obtidos pela RNA do que quando obtidos pelo classificador Máxima Verossimilhança. Os requisitos de memória e tempo de classificação foram equivalentes para ambos os métodos. Já para o tempo de treinamento, o classificador Máxima Verossimilhança foi menor. Sendo concluído que as redes neurais são um candidato melhor como método para classificação supervisionada de dados de sensoriamento remoto.

Paola e Schowengerdt (1995) também compararam RNA e o classificador Máxima Verossimilhança. O algoritmo Máxima Verossimilhança foi escolhido para esta comparação por ser amplamente utilizado como classificador padrão. Uma das principais diferenças entre os algoritmos é a velocidade, sendo o treinamento da RNA mais lento. Entretanto, o tempo de classificação - depois da rede treinada - foi considerado menor para a RNA. Chegou-se a conclusão de que a rede neural, por ser não paramétrico, é mais robusta para seleção, definição e treinamento das classes, mesmo com um número menor de pixels para treinamento, enquanto que o algoritmo de Máxima Verossimilhança é mais sensível a pureza das classes.

Paola e Schowengerdt (1995) relatam que as redes neurais mostraram-se uma ferramenta útil para sensoriamento remoto e classificação de imagens. A maior desvantagem do método é o tempo de treinamento. Como os computadores se tornam cada vez mais poderosos, a velocidade de processamento aumenta e as redes neurais se tornam mais atraentes. Este aumento de velocidade, juntamente com a capacidade flexível da rede neural e habilidade de usar conjuntos de treinamento pequenos, fazem da rede neural uma ferramenta em potencial para classificação de imagens multiespectrais de sensoriamento remoto.

3. Metodologia

O problema da classificação digital de imagens pode ser resolvido de diferentes maneiras como, por exemplo, adotando o pixel como unidade ou aplicando métodos de segmentação de imagens e classificação destes segmentos - também conhecida como classificação orientada ao objeto. Neste trabalho, a classificação é feita com base na menor unidade de informação - o pixel.

Neste estudo a classificação se restringe ao pixel por considerar que esta unidade é mais eficiente no processo de integração da imagem espectral com os dados do *Laser Scanner*, uma vez que esses dados podem ser organizados em formato *raster*, ou seja, matriz ou imagem, compatível com a resolução da imagem. Desta forma, seria atribuído a cada pixel da imagem mais informação, como uma nova camada ou banda espectral, para auxiliar na classificação.

A ideia central é, neste caso, deduzir a presença de árvores, em função da variação dos valores digitais da imagem RGB e dados altimétricos. Optou-se por fazer esta classificação binária, apenas 2 classes – árvores e não árvores – a fim de se ter uma Rede Neural Artificial especialista na identificação de apenas um alvo, neste caso, árvores.

Fez-se uso de dados altimétricos de *Laser Scanner* Aerotransportado e valores no espectro do visível. Estes dados são referentes a uma área densamente urbanizada, na cidade de Curitiba, Paraná. Nesta área há significativa variação de elementos, como a diversidade de tipos de telhados, ruas, árvores, gramados, parques, estádio de futebol, etc. Além disso, esta área possui também alta variação altimétrica, seja na declividade do próprio relevo, como nas diferentes elevações de construções ou tipos de árvores.

A localização da área de estudo, a imagem espectral e os dados altimétricos utilizados neste trabalho são apresentados na Figura 2.

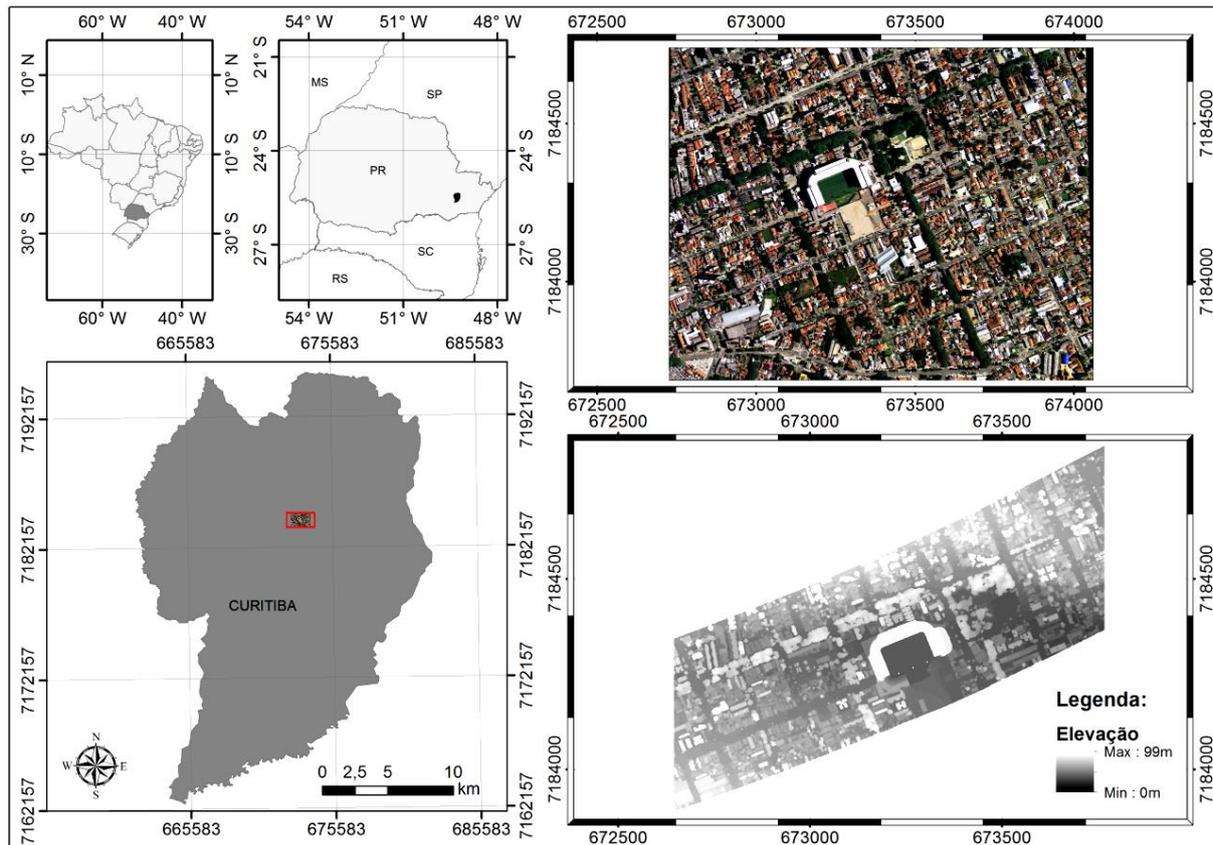


Figura 2: Localização da área de estudo, imagem espectral RGB e altimétricos *Laser Scanner*.

A imagem espectral possui as informações no espectro do visível (3 canais - RGB), obtida a partir de aerolevantamento, com resolução espacial de 30cm. Os dados *Laser Scanner* também foram organizados em formato *raster*, com a mesma resolução espacial da imagem, 30 cm.

Entretanto, para esta abordagem, onde se utilizará a informação da altura dos objetos para a utilização na classificação, precisa-se saber a altura dos objetos em relação ao solo. Ou seja, não é possível utilizar diretamente os valores altimétricos do *Laser Scanner*, é necessária a redução da influência do relevo, para utilização apenas dos valores de altura dos objetos em relação ao solo.

Conceitualmente, quando os dados se referem a pontos apenas do terreno, é possível construir um Modelo Digital do Terreno (MDT). Já quando o modelo inclui, além do terreno, os objetos a ele superpostos, tem-se o Modelo Digital da Superfície (MDS). De acordo com Weidner e Föerstner (1995), após a interpolação das grades do MDT e do MDS, uma terceira grade pode ser derivada calculando a diferença entre o MDS e o MDT. Trata-se, neste caso, de uma grade que contém apenas os objetos acima do terreno, sendo conhecida como Modelo Digital de Superfície Normalizado, ou MDSn. Neste modelo MDSn, apenas a altura dos objetos acima do terreno é representada, o que simplifica a análise. Em regiões de relevo com variação de cotas acentuada, pontos do terreno e pontos no telhado de um prédio podem ter a mesma coordenada altimétrica na grade. Após a normalização, o efeito do terreno é eliminado. A vantagem do uso do modelo normalizado no estudo de áreas urbanas é que o mesmo permite obter diretamente a altura das construções ou outros objetos. A Figura 3 traz uma representação deste problema.

A Figura 3a demonstra a representação de uma situação em que elementos que não são do terreno, como árvores, casas ou prédios, podem possuir uma cota menor que a cota do terreno na vizinhança. Nesta representação, prédios possuem cota menor que casas, ou vegetação rasteira

com cota maior que árvores. Já a Figura 3b apresenta a mesma situação após a operação de normalização, onde os efeitos do relevo são removidos, e todas as cotas das feições são maiores que o terreno, ajustando também a condição de que o prédio tem cota maior que a casa, assim como árvores com cotas maiores que vegetação rasteira, como gramados.

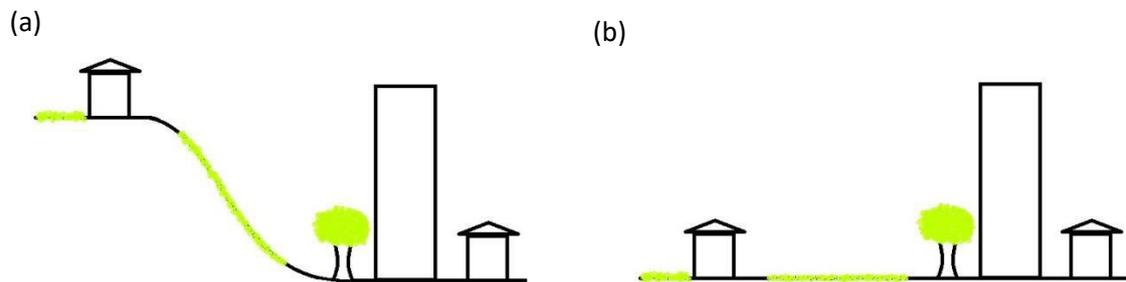


Figura 3: (a) Representação de um caso de terreno com grande variação altimétrica, (b) a mesma situação após normalização.

Segundo Zhang et al. (2003) e Centeno, Kishi e Mitishita (2009), para gerar um MDT, o terreno tem que ser identificado e os pontos que não são do terreno devem ser retirados, sendo esta filtragem considerada como desafiadora por muitos autores. Na literatura encontram-se descritas diferentes alternativas para a solução deste problema. A maior parte desses métodos se baseia na identificação de pontos do terreno em função da declividade local, continuidade da superfície ou da distância do ponto a um plano de referência que corresponderia ao terreno.

Zhang et al (2003) conclui que os filtros morfológicos podem remover elementos acima do terreno. Neste processo, aplica-se o método de morfologia matemática conhecida como operação de fechamento. No entanto, é difícil que todos os objetos, de diversos tamanhos, sejam filtrados utilizando um tamanho de janela fixa. Este problema pode ser resolvido aumentando o tamanho da janela de filtragem gradualmente. Os experimentos mostraram que a proposta do filtro morfológico progressivo, com incremento do tamanho da janela, geraram bons resultados na filtragem tanto em ambientes urbanos como em regiões montanhosas. Este filtro morfológico progressivo foi o método utilizado neste trabalho para a geração do MDT e, posteriormente, geração do MDSn.

Ao final deste processo de filtragem, pode-se utilizar então a informação da cota de cada objeto, ou pixel, juntamente com a informação espectral. A partir deste processo, fez-se então a seleção das amostras de treinamento da RNA, juntamente com a seleção das amostras de validação, para verificação da acurácia da classificação.

Parte das imagens foi classificada pelos autores como sendo pixels representativos de árvores e não árvores, para esta área de estudo. Dos pixels representativos selecionados, 50% destes pixels foram utilizados para a validação dos resultados obtidos pelos classificadores. Os outros 50% dos pixels considerados representativos foram utilizados para o treinamento dos classificadores.

Para análise da influência das amostras de treinamento no processo de classificação, o grupo de treinamento foi dividido em 10 partes (de 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% 90% e 100%). Ou seja, se da amostra considerada representativa, 50% foram utilizados para validação dos resultados obtidos pelos classificadores, e os outros 50% para treinamento, tem-se que 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50% dos pixels selecionados como representativos foram utilizadas para o treinamento dos classificadores e análise da influência do tamanho das mesmas.

Como estavam disponíveis 1.904.357 pixels ao todo - com informação altimétrica e espectral - e foram classificados pelos autores 579.518 pixels (representando 30,43% dos pixels). Metade destes pixels selecionados (289.759 pixels) foram utilizados para o processo de validação dos resultados (o que representa 15,22% do total de pixels com informação altimétrica e espectral). Foram utilizados conjuntos de treinamento de 28.975 (1,52% do total de pixels), 57.951 (3,04%), 86.927 (4,56%), 115.903 (6,09%), 144.879 (7,61%), 173.855 (9,13%), 202.831 (10,65%), 231.807 (12,17%), 260.783 (13,69%), 289.759 (15,22%) pixels para treinamento dos classificadores. Estes dados foram organizados e apresentados na Tabela 1.

Tabela 1: Quantidade de pixels amostrados, com percentuais em relação à amostragem e às imagens, para as áreas de estudo.

	Quantidade (pixels)	Percentual em relação ao Amostrado	Percentual em relação à quantidade total de pixels Imagem
Valor total Amostrado	579.518	100%	30,43%
Conjuntos de Treinamento	28.975	5%	1,52%
	57.951	10%	3,04%
	86.927	15%	4,56%
	115.903	20%	6,09%
	144.879	25%	7,61%
	173.855	30%	9,13%
	202.831	35%	10,65%
	231.807	40%	12,17%
	260.783	45%	13,69%
	289.759	50%	15,22%
Validação	289.759	50%	15,22%

Também, de forma a avaliar se a utilização de dados altimétricos do Laser Scanner podem contribuir nos resultados das classificações, as Redes Neurais Artificiais foram treinadas utilizando apenas os dados com informação espectral (RGB); e espectral juntamente com a altimétrica (RGB+*Laser Scanner*).

Para implementação de Redes Neurais Artificiais utilizou-se da *Neural Networks Toolbox*, do software *Matlab*.

Como o intuito deste trabalho é avaliar a influência do tamanho das amostras de treinamento, assim como a utilização de dados multi sensores, os demais parâmetros das RNA foram constantes para todos os testes. Foram criadas RNA multicamadas (2 camadas ocultas), com 10 neurônios em cada camada oculta. O algoritmo utilizado para o treinamento das RNA foi o *backpropagation*. No processo de treinamento, as RNA deveriam alcançar um erro menor que 0,00001, ou o máximo de 1000 interações, sem limite de tempo (infinito). A taxa de aprendizagem foi de 0,01 e com momentum de 0,9.

4. Resultados

Inicialmente, fez-se o processo de filtragem da influência do terreno. A partir dos dados altimétricos *Laser Scanner*, fez-se a organização dos valores em formato *raster*, gerando o MDS. Através do método de morfologia matemática progressiva, obteve-se apenas os valores do terreno. Após a filtragem dos dados *Laser Scanner*, os mesmos foram interpolados para preenchimentos dos espaços vazios gerados. Foi utilizado o método de interpolação bilinear. Com isso, com esse modelo totalmente preenchido, tem-se assim um Modelo Digital do Terreno, MDT. Com a subtração do MDS e MDT (MDS-MDT), tem-se o MDSn, que contém apenas os valores dos objetos acima do terreno, eliminando assim a influência do mesmo. O resultado deste processo é apresentado na Figura 4.

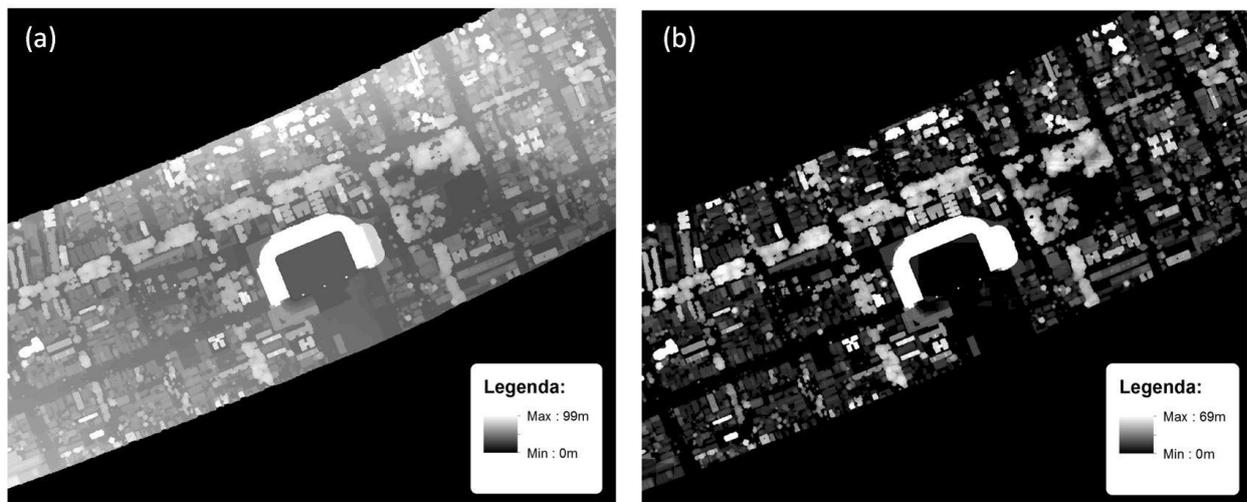


Figura 4: Modelos Digitais de Elevação da área de estudo a partir de dados *Laser Scanner* Aerotransportado. (a) Modelo Digital de Superfície, MDS; (b) Modelo Digital de Superfície Normalizado, MDSn.

A Figura 4a apresenta o MDS original, e a Figura 4b apresenta o MDSn - com a informação já normalizada, sem a influência do terreno. Na Figura 4 é possível destacar a tendência mais escura ao centro e mais claras nas extremidades na imagem da Figura 4a, e a ausência da variação de informação nas ruas na Figura 4b, por exemplo. Ainda, pela legenda, é possível destacar a redução de 30 metros de influência do terreno na cota mais alta. Desta forma, o modelo apresentado na Figura 4b apresenta apenas as cotas dos objetos acima do terreno.

Após este processo fez-se o treinamento e classificação com as Redes Neurais Artificiais, utilizando então apenas os valores espectrais (RGB) ou espectrais e cotas (RGB + *Laser Scanner*) em cada pixel. Para validação e avaliação dos resultados, foram geradas matrizes de confusão, onde as linhas representam as classes classificadas pelos autores, e as colunas são as classes classificadas pelo classificador RNA. Seria interessante que todos os valores se encontrassem na diagonal principal desta matriz, representando que todos os pixels foram classificados corretamente pelo classificador em relação aos autores. A porcentagem apresentada no canto inferior direito representa o acerto global do classificador, enquanto que os acertos parciais são apresentados nas porcentagens da linha e coluna final da matriz. As Tabelas 2 a 5 apresentam as respectivas matrizes de confusão, utilizando-se amostras de treinamento de 5% e 50%, com a utilização de dados RGB ou RGB e *Laser Scanner*.

Tabela 2: 05% de amostras de Treinamento, utilizando apenas dados RGB

	Árvores	Não Árvores	
Árvores	187811	28646	86,76%
Não Árvores	6877	66425	90,62%
	96,47%	69,87%	87,74%

Tabela 3: 50% de amostras de Treinamento, utilizando apenas dados RGB.

	Árvores	Não Árvores	
Árvores	189033	27424	87,33%
Não Árvores	6282	67020	91,43%
	96,78%	70,96%	88,37%

Tabela 4: 05% de amostras de Treinamento, utilizando dados RGB e *Laser Scanner*

	Árvores	Não Árvores	
Árvores	211674	4783	97,79%
Não Árvores	1750	71552	97,61%
	99,18%	93,73%	97,75%

Tabela 5: 50% de amostras de Treinamento, utilizando dados RGB e *Laser Scanner*.

	Árvores	Não Árvores	
Árvores	212434	4023	98,14%
Não Árvores	1690	71612	97,69%
	99,21%	94,68%	98,03%

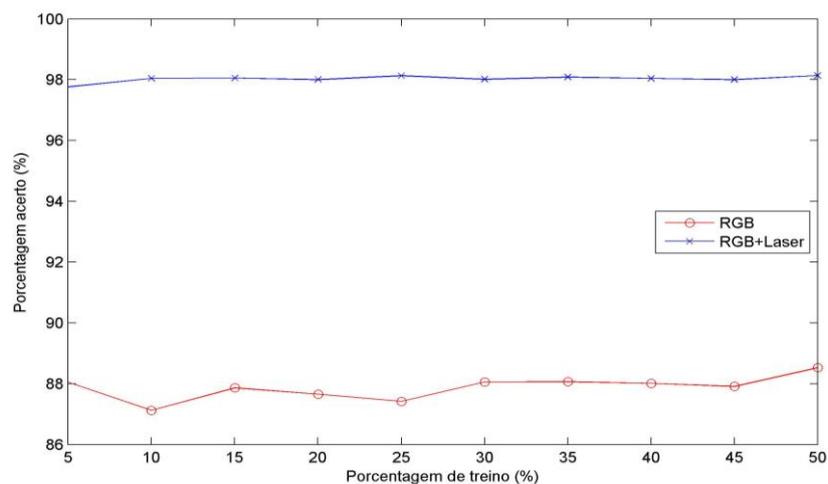
Ao analisar as tabelas, percebe-se que os acertos das RNA, treinadas com tanto com 5% a 50%, tiveram bons resultados se comparados com a classificação de referência feita pelos autores, sobretudo no que diz respeito ao acerto global. Os acertos globais de todas as RNA foram acima de 87%. Ou seja, todas as Redes Neurais Artificiais treinadas obtiveram concordância em mais de 87% dos pixels com as amostras de validação selecionadas pelos autores. Atenta-se especialmente para os valores dos acertos globais utilizando dados RGB e *Laser Scanner*, com valores de acertos globais acima de 97,5%.

Atenta-se também para os pixels classificados erroneamente pelas RNA, uma vez que estes valores podem expressar as principais dificuldades e confusões geradas pelos classificadores. Quando comparado os valores classificados pelos autores como 'árvores' e pela RNA como 'não árvores' (falso negativo), percebe-se que, ao utilizar apenas informação do RGB esses pixels representavam aproximadamente 28 mil pixels, e quando a informação do Laser Scanner foi adicionada ao processo, este valor diminuiu para 4 mil pixels. Ainda, comparando os valores dos falso positivos (onde as RNA incluíam pixels na classe 'árvore' e os autores consideraram como 'não árvore'), é possível verificar que apenas com a informação RGB esses valores eram de 6 mil pixels, caindo para 1 mil pixels quando inserida a informação do Laser Scanner. Com isso, é possível verificar que a ordem de grandeza desses valores permaneceu similar entre as classificações com tamanhos de treinamento de 05% e 50%, sofrendo maior influência quando adicionada ou não a informação do altimétrica do Laser Scanner.

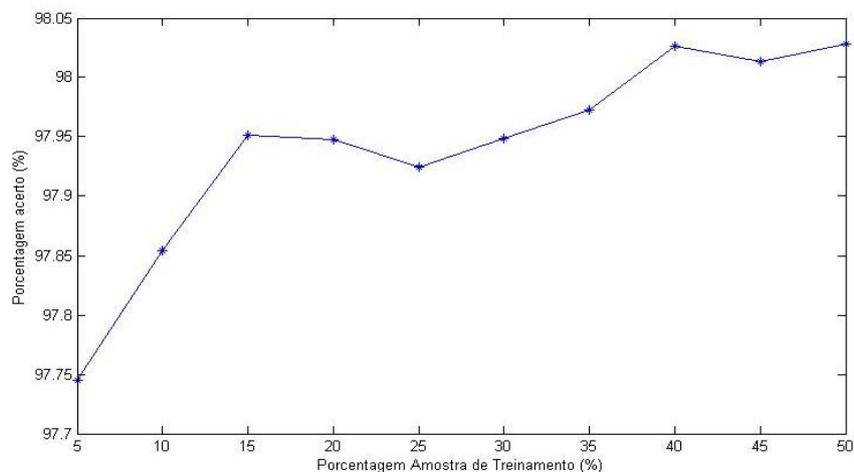
A fim de sintetizar e comparar os resultados das classificações, os acertos globais foram agrupados em função do tamanho das amostras de treinamento utilizadas e construiu-se a Tabela 6 e Figura 5. Percebe-se que os acertos globais foram elevados, com pouca variação entre os diferentes tamanhos de amostras de treinamento, mas com visível diferença entre a utilização ou não de dados altimétricos na classificação. As diferenças nos acertos globais em função do tamanho das amostras de treinamento foram na ordem de 1 ponto percentual, enquanto que a diferença entre a utilização de apenas informação espectral (RGB) contra espectral e altimétrica (RGB+*Laser Scanner*) foi na ordem de 10 pontos percentuais.

Tabela 6: Acertos Globais da RNA, conforme Porcentagem de Amostras utilizadas no Treinamento, utilizando apenas informação RGB ou RGB e dados *Laser Scanner*

Porcentagem Treinamento	Acertos Globais (%)	
	RGB	RGB + <i>Laser Scanner</i>
5%	87,74	97,75
10%	87,68	97,85
15%	87,82	97,95
20%	87,93	97,95
25%	88,51	97,92
30%	87,72	97,95
35%	88,05	97,97
40%	88,35	98,03
45%	87,70	98,01
50%	88,37	98,03

**Figura 5:** Acertos Globais da RNA, em função do tamanho das amostras de treinamento, utilizando apenas dados RGB, ou RGB e altimétricos *Laser Scanner*.

De maneira a melhor ilustrar a “curva de aprendizado” das RNA em função do tamanho das amostras de treinamento, apresenta-se na Figura 6 os acertos globais para as Redes Neurais Artificiais utilizando dados RGB e *Laser Scanner*.

**Figura 6:** Acertos Globais da RNA, em função do tamanho das amostras de treinamento, utilizando dados RGB e altimétricos *Laser Scanner*.

Ao se analisar os valores dos acertos globais em função do tamanho das amostras de treinamento apresentados na Tabela 6 e gráficos da Figura 5 e 6, é possível identificar um crescimento nos acertos globais conforme se aumenta o número de amostras de treinamento. Isto porque, aumentando a quantidade de elementos na amostra de treinamento, aumenta-se a variabilidade dos valores dos pixels fornecidos a RNA, e aumenta-se a capacidade de discriminação de pequenas variações em valores.

Entretanto, percebe-se que há sim uma tendência de crescimento dos acertos, mas não necessariamente linear e constante com o aumento das amostras de treinamento. Isto porque, conforme se aumenta a variabilidade de informações dos pixels apresentados à RNA, melhora-se a capacidade de discriminar novos elementos, porém, pode-se também estar adicionando pixels de ruído, resultado em certa confusão por parte da RNA na classificação.

Contudo, ressalta-se que, apesar da RNA ter sido treinada com menor quantidade de amostra de treinamento (apenas 1,52% do total de pixels da imagem com informação espectral e cota), esta já apresentou resultados expressivos, com 97,75% de acerto global. Ou seja, com a seleção de apenas 1,52% dos pixels da imagem com informação altimétrica e espectral para treinamento, foi possível obter uma classificação com Redes Neurais Artificiais com acertos globais na ordem de 97,75%.

Desta forma, pode-se concluir que, mesmo com amostras de treinamento pequenas, é possível ter resultados muito bons utilizando Redes Neurais Artificiais na classificação supervisionada, sobretudo com a utilização de dados espectrais e altimétricos. Ainda, amostras de treinamento pequenas diminuem o esforço na seleção de grande quantidade de pixels representativos para cada classe, facilitando e agilizando o processo de classificação de imagens.

Como forma de apresentar o resultado final da classificação através de RNA, apresenta-se a Figura 7. A Figura 7 traz um comparativo entre a imagem espectral original, com 30cm de resolução espacial (Figura 7a), e a identificação das árvores resultante do treinamento da RNA com apenas 1,52% da imagem total, com informação espectral e altimétrica do *Laser Scanner* (Figura 7b).



Figura 7: Resultado da classificação de árvores em ambiente urbano, utilizando Redes Neurais Artificiais, com imagens espectrais de alta resolução e dados *Laser Scanner* e amostra de treinamento de apenas 1,52% do total da imagem. (a) Imagem Espectral; (b) identificação de pixels referente a árvores.

5. Conclusões

Neste trabalho gerou-se RNA especialistas na identificação da classe “árvores”, apresentando bons resultados, com acertos globais de até 98,03%. Contudo, tal metodologia pode ser aplicada a diferentes alvos, como estradas, telhados, água, solo exposto, vegetação rasteira, etc., desde que com a utilização de pixels representativos de cada classe para o treinamento das RNA. A vantagem em se ter classificadores especialistas para cada classe está na simplicidade do processo, uma vez que pode-se estar interessado em apenas um alvo ou característica. Ou, ainda, com a geração de várias RNA especialistas para cada alvo ou uso do solo para determinada cena ou conjunto de dados, pode-se executar posterior comparação e fusão dos resultados obtidos por cada RNA especialista, determinando assim os pixels e classes de maior confusão entre os resultados obtidos.

Pode-se concluir com este estudo que os resultados das classificações utilizando RNA, dados espectrais de alta resolução e dados *Laser Scanner* foram muito bons. Ainda, foi possível verificar que, conforme se aumenta o número de amostras de treinamento, existe uma tendência em aumentar a acurácia das classificações. Isto porque, aumentando as amostras de treinamento, aumenta-se a variabilidade de informações oferecidas a RNA, aumentando sua capacidade de diferenciação dos valores a serem classificados. Porém, aumenta-se também a possibilidade de inclusão de dados ruidosos. Contudo, verificou-se que pequenas amostras de treinamento já foram capazes de obter ótimos resultados. Ou seja, utilizando-se apenas 1,52% da imagem com informações altimétricas e espectrais, foi possível obter acertos na ordem de 97,75%.

Ao se comparar a utilização apenas de dados espectrais para classificação, ou dados espectrais juntamente com dados altimétricos, concluiu-se que houve uma expressiva melhora nas classificações com a utilização de dados multi fontes. Esta utilização de dados multi fontes, como no caso de espectrais e altimétricos, trouxe melhores resultados do que o acréscimo no tamanho das amostras de treinamento. Enquanto o aumento do tamanho das amostras de treinamento trouxeram ganhos na ordem de 1% no acerto global da classificação, a utilização de dados altimétricos contribuiu para uma melhora na ordem de 10% no acerto global. Ou seja, para classificações supervisionadas utilizando Redes Neurais Artificiais, pode-se concluir que é mais importante oferecer mais informações sobre os objetos que serão classificados, do que outras informações de outros objetos similares também presentes nesta classe.

Com isso, percebe-se que Redes Neurais Artificiais são poderosas ferramentas para a área de sensoriamento remoto e geoprocessamento, uma vez que podem trabalhar com grande quantidade de dados, dados de diferentes fontes e características, e ter pouca influência de dados ruidosos, podendo facilmente trabalhar com dados de grande variabilidade, como no caso de dados de alta resolução espacial de áreas urbanas. Desta forma, conclui-se que RNA para classificação de imagens apresenta-se como uma forma robusta para geração de informação a partir de dados geográficos, sobretudo para o processo de mapeamento, atualizações cartográficas, e trabalhos com grande volume de dados.

AGRADECIMENTOS

Os autores agradecem ao CNPq pelo suporte financeiro e à empresa Esteio Engenharia e Aerolevantamentos pelo fornecimento da imagem e dados *Laser Scanner*.

REFERÊNCIAS BIBLIOGRÁFICAS

- Andrade, A. F., & Centeno, J. A. S. 2003. Integração de informações espectrais e de forma na classificação de imagens com redes neurais. *Boletim de Ciências Geodésicas*, 9(2).
- Araki, H. 2005. Fusão de Informações Espectrais, Altimétricas e de dados auxiliares na classificação de Imagens de Alta Resolução Espacial. 136 p. Tese de doutorado, *Curso de Pós-Graduação em Ciências Geodésicas, UFPR*.
- Botelho, M. F. 2004. Aplicação de Redes Neurais na Classificação de Imagens de Alta Resolução Espacial e Dados do *Laser Scanner*, Usando uma Abordagem Orientada a Regiões. Dissertação de Mestrado, *Curso de Pós-Graduação em Ciências Geodésicas, UFPR*.
- Centeno, J. A. S. 2009. *Sensoriamento Remoto e Processamento de Imagens Digitais*. Curitiba: Universidade Federal do Paraná.
- Centeno, J. A. S., Kishi, R. T., & Mitishita, E. A. 2009. Aplicação de laser scanner para detecção de áreas para coleta de água de chuva. *SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO*, 14, 5009-5013.
- Chettri, S. R., Crompton, R. F., & Birmingham, M. 1992. Design of neural networks for classification of remotely sensed imagery. *Telematics and Informatics*, 9(3), 145-156.
- Foody, G. M., & Mathur, A. 2004. Toward intelligent training of supervised image classifications: directing training data acquisition for SVM classification. *Remote Sensing of Environment*, 93(1), 107-117.
- Görgens, E. B., Montaghi, A., & Rodriguez, L. C. E. 2015. A performance comparison of machine learning methods to estimate the fast-growing forest plantation yield based on laser scanning metrics. *Computers and Electronics in Agriculture*, 116, 221-227.
- Haykin, S., & Lippmann, R. 1994. Neural Networks, A Comprehensive Foundation. *International Journal of Neural Systems*, 5(4), 363-364.
- Heermann, P. D., & Khazenie, N. 1992. Classification of multispectral remote sensing data using a back-propagation neural network. *Geoscience and Remote Sensing, IEEE Transactions on*, 30(1), 81-88.
- Höfle, B., Hollaus, M., & Hagenauer, J. 2012. Urban vegetation detection using radiometrically calibrated small-footprint full-waveform airborne LiDAR data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 67, 134-147.
- Lek, S., & Guégan, J. F. 1999. Artificial neural networks as a tool in ecological modelling, an introduction. *Ecological modelling*, 120(2), 65-73.
- Mendes, T. S. G., & Dal Poz, A. P. 2013. Integração de imagem aérea de alta resolução e dados de varredura a laser na classificação de cenas urbanas para detectar regiões de via. *Boletim de Ciências Geodésicas*, 19(2), 287.
- Pal, M., & Mather, P. M. 2004. Assessment of the effectiveness of support vector machines for hyperspectral data. *Future Generation Computer Systems*, 20(7), 1215-1225.
- Paola, J. D., & Schowengerdt, R. 1995. A detailed comparison of backpropagation neural network and maximum-likelihood classifiers for urban land use classification. *Geoscience and Remote Sensing, IEEE Transactions on*, 33(4), 981-996.
- Petropoulos, G. P., Arvanitis, K., & Sigrimis, N. 2012. Hyperion hyperspectral imagery analysis combined with machine learning classifiers for land use/cover mapping. *Expert systems with Applications*, 39(3), 3800-3809.

Vieira, C. A. O. 2000. *Accuracy of remotely sensing classification of agricultural crops: a comparative study* (Doctoral dissertation, University of Nottingham).

Weidner, U., & Förstner, W. 1995. Towards automatic building extraction from high-resolution digital elevation models. *ISPRS journal of Photogrammetry and Remote Sensing*, 50(4), 38-49.

Xu, S., Vosselman, G., & Elberink, S. O. 2014. Multiple-entity based classification of airborne laser scanning data in urban areas. *ISPRS Journal of photogrammetry and remote sensing*, 88, 1-15.

Zhang, K., Chen, S. C., Whitman, D., Shyu, M. L., Yan, J., & Zhang, C. 2003. A progressive morphological filter for removing nonground measurements from airborne LIDAR data. *Geoscience and Remote Sensing, IEEE Transactions on*, 41(4), 872-882.

Recebido em 4 de novembro de 2015.

Aceito em 11 de junho de 2016.