



**11ª Jornada Científica e
Tecnológica do IFSULDEMINAS**
& **8º Simpósio de
Pós-Graduação**

**CLASSIFICAÇÃO DE IMAGENS MULTIESPECTRAIS E HIPERESPECTRAIS
POR MEIO DE ÁRVORES DE DECISÃO**

Alana S. T. BARBOSA¹; Camila S. dos ANJOS²

RESUMO

O sensoriamento remoto é uma importante ferramenta para a aquisição de informações da superfície terrestre. Essas informações servem de apoio na tomada de decisões para profissionais de diversos segmentos. O presente trabalho teve como objetivo avaliar comparativamente a qualidade dos métodos de classificação de imagens Random Forest e C4.5. Para isso foi utilizada uma imagem RGB do sensor DiMAC ULTRALIGHT + com resolução espacial de 5 cm da cidade de Houston – Texas, Estados Unidos. Foi realizada também a extração de atributos PCA e MNF. Por último foi realizada a validação através do *software* WEKA para os dois métodos de classificação, obtendo-se a porcentagem de acertos para cada uma.

Palavras-chave:

Sensoriamento remoto; Áreas urbanas; Classificadores.

1. INTRODUÇÃO

O sensoriamento remoto desde o seu surgimento tem contribuído muito para o estudo das geociências, que fora potencializado com as constantes evoluções tecnológicas que afetou diretamente os sistemas sensores. Estes sistemas sensores se tornaram cada vez mais eficazes em obter informações remotamente, principalmente quando se trata de identificação de alvos em ambientes urbanos.

De acordo com Herold *et al.* (2003), ambientes urbanos representam uma das áreas mais desafiadoras para a análise de sensoriamento remoto pois sua diversidade espectral excede em muito a de outros ambientes antropizados e naturais. Os materiais presentes no ambiente urbano apresentam alta diversidade espacial e espectral, pois incluem uma ampla gama de telhados, estradas, calçadas, submetidos a intemperismo e de diversas idades, qualidades e composições químicas, entre outras minúcias.

¹ Bolsista PIBIC/CNPq, IFSULDEMINAS – *Campus* Inconfidentes. E-mail: alanasaritabarbosa@hotmail.com

² Orientadora, IFSULDEMINAS – *Campus* Inconfidentes. E-mail: camila.lacerda@ifsuldeminas.edu.br

Na classificação de dados de sensoriamento remoto de áreas urbanas, a identificação de objetos relativamente pequenos é importante. Por conseguinte, uma elevada resolução espacial da imagem é necessária para a classificação precisa (MELGANI; BRUZZONE, 2004).

Segundo Weber e O'Connell (2011), em certas aplicações, como, por exemplo, o mapeamento de área urbana, a integração de alta resolução espacial e espectral é importante para se atingir um mapeamento suficientemente preciso com bons resultados na discriminação de alvos.

Assim, de acordo com Anjos (2016), o quadro ideal para a identificação precisa de alvos urbanos seria a combinação dessas duas características (espacial e espectral), o que permitiria melhor detecção dos elementos que compõem o espaço urbano, sobretudo com o auxílio de classificadores automáticos de última geração.

Diante deste exposto, o presente trabalho teve como objetivo analisar comparativamente a qualidade das classificações geradas por diferentes classificadores de uma imagem RGB do sensor DiMAC ULTRALIGHT +, adquirida pela NCALM (*National Center for Airborne Laser Mapping*).

2. MATERIAL E MÉTODOS

A figura 1 apresenta o fluxograma metodológico do presente trabalho.

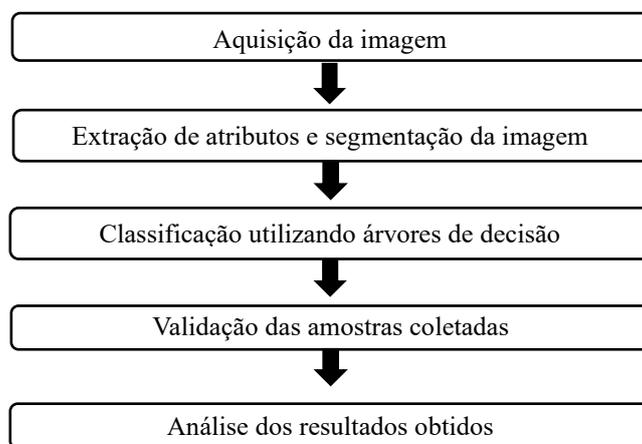


Figura 1 – Fluxograma Metodológico do Trabalho.

2.1. Aquisição da Imagem

A área de estudo deste trabalho é parte da cidade de Houston – Texas, Estados Unidos da América e a imagem de alta resolução foi cedida gratuitamente pela GRSS (*Geoscience & Remote Sensing Society*) por meio do IEEE *Data Fusion Contest* 2018. O sensor DiMAC ULTRALIGHT + possui distância focal de 70 mm e GSD de 5 cm. Os dados foram adquiridos pelo NCALM em 16 de fevereiro de 2017 entre 16:31 e 18:18 GMT, o sensor estava a bordo de um avião Navajo Chieftain da Piper PA- 31 – 350.

2.2. Extração de Atributos da Imagem

A extração de atributos é realizada para que os atributos sejam inseridos durante o processo de segmentação e classificação, resultando na melhora da classificação da imagem. Neste trabalho utilizou-se a transformação por componentes principais, PCA (*Principal Component Analysis*) e a fração mínima de ruído, MNF (*Minimum Noise Fraction*).

A transformação MNF é um procedimento estatístico que envolve a Análise por Principais Componentes (PCA) e que concilia tanto os procedimentos de segregação do componente ruído, quanto também de redução da dimensionalidade dos dados sem perda de informação (GREEN *et al.*, 1988).

A principal diferença entre os métodos MNF e PCA é que a PCA realiza o ordenamento dos dados de acordo com a variância, enquanto a MNF o realiza por meio da relação sinal/ruído. Assim, a transformação por MNF permite o ordenamento de suas componentes conforme a qualidade da imagem expressa pelo nível de ruído presente na cena (LEE *et al.*, 1990).

2.3. Classificação utilizando árvores de decisão

Ao realizar a segmentação da imagem, foram coletadas amostras para 21 classes, sendo estas, água, árvores de folhas secas, árvores de folhas verdes, assentos de estádio, calçadas, travessas, carros, estacionamentos pavimentados e não pavimentados, estradas, ferrovias, grama artificial, grama saudável, grama estressada, prédios residenciais e não residenciais, rodovias, solo exposto, trens, vias principais e amostras de sombra. Ao exportar estas amostras para o *software* WEKA, foi realizado a classificação da imagem utilizando os classificadores Random Forest e C4.5 e em seguida a visualização da classificação no *software* QGIS.

4. RESULTADOS E DISCUSSÕES

Por meio da validação no *software* WEKA (Tabela 1), nota-se que para ambos ocorreu baixa porcentagem de amostras classificadas corretamente, devido as classes apresentarem algumas confusões entre si.

Tabela 1 – Validação das Amostras

Classificador	Total de acertos (%)
C4.5	41,85
Random Forest	46,67

Para avaliar a qualidade da classificação da imagem, Landis e Koch (1977) apresentam uma tabela categorizando o índice Kappa (Tabela 2). Para este trabalho o classificador C4.5 apresentou valor do coeficiente estatístico Kappa de 0,3862 e para o Random Forest valor de 0,4369, obtendo como resultado uma classificação considerada razoável pelo classificador C4.5 e considerada boa para o Random Forest. Vale ressaltar que os valores de acertos encontrados são válidos apenas para a cena em questão.

Tabela 2 – Categorização do índice Kappa

Valores	Categorização
<0	Péssima
0 – 0,2	Ruim
0,21 – 0,4	Razoável
0,41 – 0,6	Bom
0,61 – 0,8	Muito Bom
0,81 – 1,0	Excelente

Fonte: Landis e Koch (1977).

5. CONCLUSÕES

Diante das análises realizadas pode se concluir que devido a imagem do sensor DiMAC ULTRALIGHT + possuir alta resolução espacial e baixa resolução espectral, houve algumas confusões entre as classes, sendo as mais frequentes grama estressada e grama saudável, vias principais e rodovias. Entretanto, acredita-se que o resultado da matriz de confusão esteja satisfatório, sendo assim temos o classificador por meio de árvores de decisão uma ferramenta adequada para classificação de imagens.

REFERÊNCIAS

ANJOS, C. S. Classificação de Áreas Urbanas com Imagens Multiespectrais e Hiperespectrais utilizando Métodos Não-Paramétricos - 381p. Tese (Doutorado) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2016.

GREEN, A. A.; BERMAN, M.; SWITZER, P.; CRAIG, M. D., "A transformation for ordering multispectral data in terms of image quality with implications for noise removal," IEEE Transactions on Geoscience and Remote Sensing, v. 26, n. 1, p. 65-74, 1988. DOI 10.1109/36.3001. Disponível em: <<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=3001&isnumber=170>>. Acesso em: 18 julho de 2019.

HEROLD, M.; GARDNER, E.M.; ROBERTS, D.A. Spectral resolution requirements for mapping urban areas. IEEE Transactions on Geoscience and Remote Sensing, v. 41, p. 1907-1919, 2003.

LANDIS, J. R.; KOCH, G. G. The measurement of observer agreement for categorical Data. Biometrics, v. 33, n. 1, p.159-174, 1977. DOI 10.2307/2529310. Disponível em: <<http://www.jstor.org/stable/pdf/2529310.pdf?acceptTC=true>>. Acesso em: 26 julho 2019.

LEE, J. B; WOODYATT, A. S.; BERMAN, M. Enhancement of high spectral resolution remote sensing data by a noise-adjusted principal components transform. IEEE Transactions on Geoscience and Remote Sensing, v.28, p. 295-304. 1990.

MELGANI, F.; BRUZZONE, L. Classification of hyperspectral remote sensing images with support vector machines. IEEE Transactions on Geoscience and Remote Sensing, v. 42, n. 8, p. 1778-1790, 2004.

WEBER, R.A.; O'CONNELL, K.M. Alternative futures: United States Commercial satellite imagery in 2020. Washington: Innovative analytics & training, 2011. 111 p. Disponível em: <<http://www2.gwu.edu/~nsarchiv/NSAEBB/NSAEBB404/docs/37.pdf>> Acesso em: 17 julho 2019.