

Integração de OBIA, árvores de decisão e classificação hierárquica para mapeamento de voçorocas

Andrea Tedesco¹
Alzir Felipe Buffara Antunes²

^{1,2} Programa de Pós-Graduação em Ciências Geodésicas - UFPR
Caixa Postal 19.001 - 81.531-990 – Curitiba - Pr, Brasil

¹ atedesco@uepg.br

² felipe@ufpr.br

Abstract. Gully erosion is a mix of soil, vegetation and water, which complicates its classification using only spectral information. It presents considerable altimetric variation, with high boundary declivity and distinct spatial features. Due to this problem, object-based image analysis, including altimetric data, shape and context information, are more appropriated to discriminate gully erosion. Mapping gully erosion is a need for monitoring, prevention, correction and mitigation, what motivated this research. The aim of this work is to propose an approach to classify gully erosion, integrating spectral data (Ikonos imagery) and intensity and altimetric data (Airborne LASER Scanner – ALS), using object-based image analysis. Relevant input data, segmentation parameters, attributes and fuzzy membership functions were identified to discriminate gully erosion. The objects were generated by multiresolution segmentation (Fractal Net Evolution Approach – FNEA) and afterwards they were classified by hierarchical and decision tree (Classification And Regression Trees - CART) classification. The hierarchical classification required specialist's knowledge about the target of interest. The tree decision classification selected attributes and input data automatically. The hierarchical and tree decision classification results were very similar, showing that is possible to use CART to predefine parameters and attributes and afterwards refine them to improve hierachical classification.

Palavras-chave: object-based classification, data mining, gully erosion, classificação baseada em objeto, mineração de dados, voçoroca.

1. Introdução

As voçorocas são processos erosivos caracterizados por depressões profundas, circundadas por vertentes quase verticais, que se alargam nas proximidades das cabeceiras devido à intensa atividade erosiva regressiva e se afunilam junto à foz do curso-d'água que a percorre, cortando sedimentos arenosos de fraca coesão, sem apresentar forte declive longitudinal (VIEIRA, 1978). Para que ações preventivas, mitigatórias ou corretivas possam ser aplicadas, faz-se necessário o mapeamento e monitoramento dessas feições erosivas, os quais podem ser realizados por medidas conduzidas no local ou por sensoriamento remoto.

As medidas efetuadas no local, apesar de amplamente utilizadas, requerem o caminhamento e instalação de equipamentos nas bordas e interior das voçorocas, o que pode agravar os processos erosivos já instalados e representar risco para os executores do levantamento. Dentre as técnicas mais comuns estão o estaqueamento (HESSEL e VAN ASCH, 2003; GUERRA, 2005; MORGAN, 2005) e levantamentos topográficos convencionais ou com uso de equipamento terrestre de varredura a LASER (*Light Amplification by Stimulated Emission of Radiation*) (PERROY et al., 2010). Alternativamente aos levantamentos no local, técnicas aerofotogramétricas podem ser usadas (MARTÍNEZ-CASASNOVAS et al., 2004; MARZOLFF e POESEN, 2009). A identificação de áreas com risco de erosão (KING et al., 2005; VRIELING et al., 2005), o cálculo volumétrico e estimativas de perdas de solo (JAMES et al., 2007; EUSTACE et al., 2009) podem ser realizados com base em imagens de satélite e em dados de ALS (*Airborne Laser Scanner*). A busca por métodos automáticos ou semiautomáticos, em abordagem orientada a

objeto, tem sido alvo de estudos sobre voçorocas (PERROY *et al.*, 2010; SHRUTHI *et al.*, 2011 e 2012; JOHANSEN *et al.*, 2012; WANG *et al.*, 2014; ALKHASAWNEH *et al.* (2014).

As voçorocas são constituídas pela mistura de solo, vegetação e água, o que dificulta a classificação delas utilizando apenas informações espectrais. Elas apresentam considerável variação altimétrica, com alta declividade nas bordas, e características espaciais bastante distintas. Portanto, análises orientadas a objeto, incluindo dados altimétricos e informações de forma e contexto, são mais apropriadas para a discriminação de voçorocas.

Nesse contexto, esse estudo teve por objetivo propor uma metodologia para a classificação de voçorocas, integrando dados espectrais (imagem Ikonos) e dados de intensidade e altimétricos (ALS), em análise orientada a objeto. Identificaram-se os dados de entrada mais relevantes, os parâmetros ótimos para a segmentação e quais atributos são capazes de extrair a classe voçoroca dentre as demais. Os objetos foram gerados por segmentação multirresolução (FNEA - *Fractal Net Evolution Approach*), sendo posteriormente classificados usando os métodos de classificação hierárquica e por árvores de decisão (algoritmo CART - *Classification And Regression Trees*).

2. Materiais e Métodos

2.1 Área de Estudo

O estudo foi conduzido para uma área localizada na zona rural do município de Uberlândia - MG, na Fazenda Experimental do Glória, pertencente à Universidade Federal de Uberlândia. A figura 1 ilustra a localização da área de estudo e algumas características da voçoroca, como a mistura de solo, vegetação e água (foto 1 – cabeceira da voçoroca) e a variação altimétrica (foto 2 - canal principal da voçoroca). As direções das tomadas das fotos são indicadas em azul e vermelho.

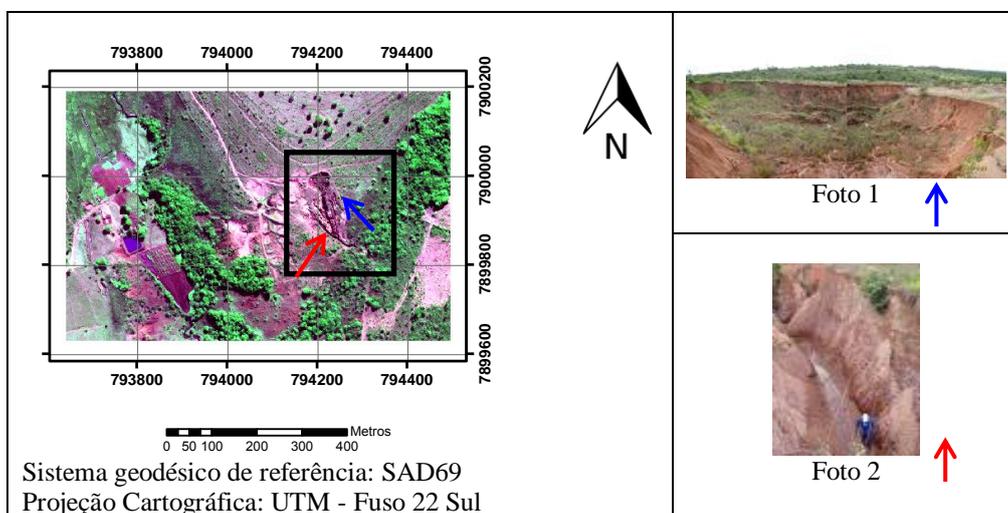


Figura 1. Localização e registros fotográficos da área de estudo.

2.2 Materiais

Para o desenvolvimento da metodologia foram usados:

- Dados ALS (ALTM 2025 - *Airborne LASER Terrain Mapper* - Optech Inc.): voos efetuados de janeiro a junho de 2004, com altura de voo de 1.000 m; densidade média de 1,89 pontos por m²; distância média entre pontos de 1,23 m; qualidade planimétrica após processamento (segundo o fabricante) de 0,5 m a 1,0 m;
- Imagem Ikonos multiespectral: bandas 1, 2, 3 e 4 (Blue, Green, Red e Near InfraRed) fusionadas com a Pan (pancromática), com resolução espacial de 1 m.

Para o processamento dos dados foram utilizados os programas: ENVI - *The Environment for Visualizing Images 4.7*; ALDPAT - *Airborne LiDAR Data Processing and Analysis Tools* (ZHANG e CUI, 2007); eCognition Developer 8.8; ERDAS Imagine 9.1 e ArcGis 9.

2.3 Métodos

2.3.1 Pré-processamento dos dados

Os dados ALS foram usados para gerar o Modelo Digital de Superfície - MDS e a imagem de intensidade. Sobre os dados altimétricos realizou-se a triangulação de Delaunay e, posteriormente, a interpolação linear. Sobre os dados de intensidade, aplicou-se a interpolação linear. Os procedimentos permitiram a transformação dos pontos irregularmente espaçados em uma grade regular com resolução espacial de 1 m.

O Modelo Digital de Terreno - MDT (figura 2a) foi gerado a partir do MDS, aplicando-se o filtro ATIN - *Adaptive TIN Filter* (AXELSSON, 2000). O Modelo Digital de Superfície Normalizado - MDSN (figura 2b) foi obtido usando-se o princípio de matemática de bandas, subtraindo-se o MDT do MDS. A partir do MDT gerou-se o mapa de contornos (curvas de nível) e, a partir dele, o mapa de declividade (figura 3a). O NDVI - *Normalized Difference Vegetation Index* (figura 3b) foi obtido usando-se o princípio de matemática de bandas, sendo $NDVI = (B4-B3)/(B4+B3)$ ¹.

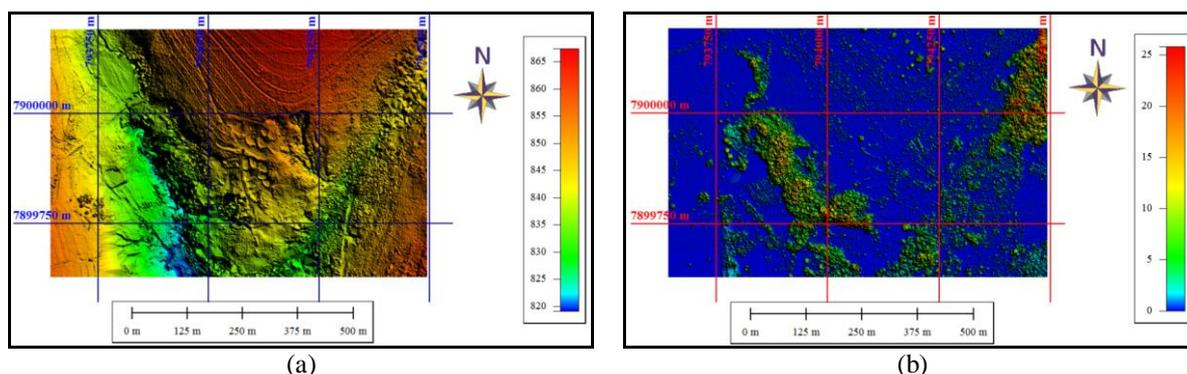


Figura 2. (a) Modelo Digital de Terreno; (b) Modelo Digital de Superfície Normalizado.

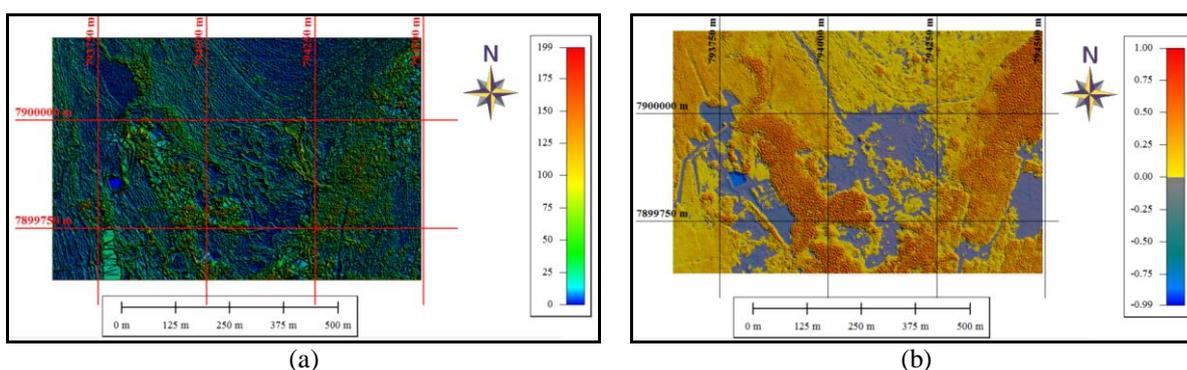


Figura 3. (a) Mapa de Declividade; (b) Imagem NDVI.

O MDSN foi usado na discriminação entre vegetação arbórea e rasteira, a declividade na discriminação da classe voçoroca (borda com declividade acentuada). O NDVI foi usado para melhorar a discriminação das áreas vegetadas e não vegetadas.

Os dois conjuntos de dados (ALS e IKONOS) foram registrados, obtendo-se erro máximo do registro em X (E_x) = 0,8653 m e em y (E_y) = -1,3139 m. O erro total foi de 0,8698 m, inferior a um *pixel*.

¹ banda 4 = infravermelho próximo; banda 3 = vermelho

2.3.2 Segmentação

A segmentação dos dados foi realizada pela técnica FNEA, usando-se o algoritmo de segmentação multirresolução. Foram utilizados como dados de entrada as quatro bandas da imagem IKONOS, a imagem NDVI, a imagem de intensidade do ALS, o MDT, o MDSN e o mapa de declividade. Variou-se o fator de escala (FE) de 5 a 100 (com intervalos múltiplos de 5, ex: 5-10-15-20...), e FE=125 e FE=150. Os pesos atribuídos aos parâmetros de entrada foram: cor = 0,9; forma = 0,1; compacidade = 0,5 e suavidade = 0,5.

2.3.3 Planejamento da Classificação

Por se tratar de ambiente rural, poucas classes foram identificadas para o interesse desse estudo, sendo essas: vegetação (arbórea e rasteira), corpos-d'água e solo (solo exposto e voçoroca). A partir das chaves de interpretação analisadas para cada classe, foram definidos os descritores de cada classe, bem como as funções de pertinência fuzzy, a serem usados na classificação hierárquica. A figura 4 ilustra as funções fuzzy utilizadas, o intervalo de valores, classe para a qual foi utilizada e banda à qual foi aplicada.

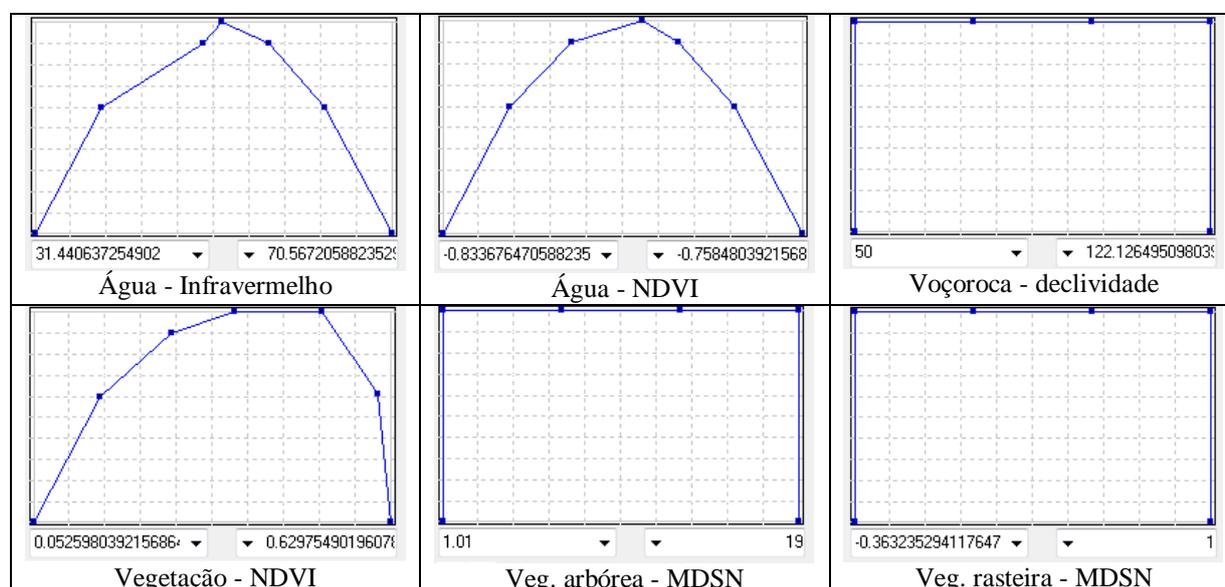


Figura 4. Funções de pertinência fuzzy utilizadas.

2.3.4 Classificação Hierárquica

A classificação hierárquica foi realizada com auxílio do programa eCognition Developer 8.8. A figura 5a apresenta a rede hierárquica estabelecida em função dos níveis de segmentação com FE=150, FE=30 e FE=20. No FE=150 foi realizada a classificação das feições em vegetação e não vegetação, usando-se os valores digitais da imagem NDVI. A partir dessa primeira classificação, no nível com FE 30, foram separados os objetos de vegetação em arbórea ou rasteira, usando-se as alturas dos objetos constantes no MDSN. Os objetos classificados como não vegetação foram reclassificados como água ou solo, em função dos valores digitais da banda do infravermelho próximo e da imagem NDVI. No nível com FE 20 foram reclassificados os objetos de solo, detalhando-os em solo exposto ou voçoroca. O atributo determinante nessa classificação foi a declividade dos objetos. A figura 5b apresenta os descritores e algoritmos classificadores (*nearest neighbour* e/ou fuzzy) utilizados para cada classe.

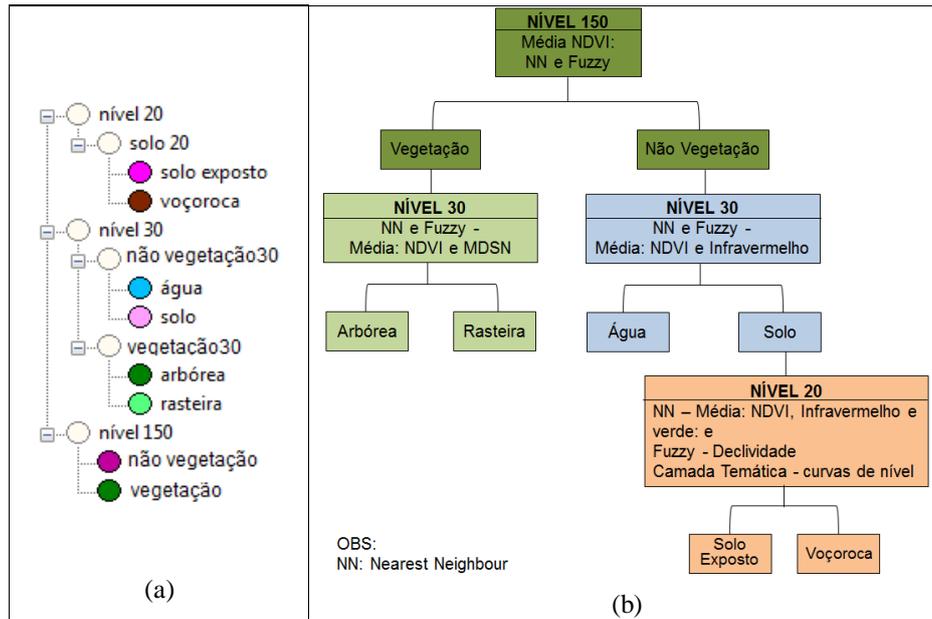


Figura 5. (a) Rede Hierárquica; (b) descritores das classes.

2.3.5 Classificação por Árvores de Decisão

A classificação por árvores de decisão foi realizada com auxílio do algoritmo CART, implementado no programa eCognition Developer 8.8. Como dados de entrada foram usados: as quatro bandas da imagem IKONOS, a imagem NDVI, a imagem de intensidade, o MDSN, o MDS e o mapa de declividade. A classificação foi realizada sob os objetos originados pela segmentação com FE=50. Foram selecionadas amostras para as classes definidas na rede semântica, visando o treinamento da árvore de decisão. Posteriormente, a classificação foi realizada pela árvore de decisão, em função das regras de decisão estabelecidas pela mineração dos dados.

3. Resultados e Discussão

3.1 Classificação Hierárquica

Analisando-se as matrizes de confusão, a classificação dos elementos de “vegetação” e “não vegetação”, no nível com FE=150, não apresentou confusão entre as duas classes. A reclassificação dos elementos de “vegetação” nas classes “arbórea” e “rasteira”, bem como os de “não vegetação” nas classes “água” e “solo”, no nível com FE=30, também não apresentou confusão entre as classes. A matriz de confusão da reclassificação dos objetos “solo” em “solo exposto” e “voçoroca”, no nível com FE=20, apresentou duas amostras selecionadas como solo exposto sendo classificadas como voçoroca. A acurácia global nesse nível hierárquico foi de 95,12% e o índice Kappa = 0,90.

A declividade foi utilizada como atributo mais significativo na separação das classes solo exposto e voçoroca. Entretanto, a declividade no interior da cabeceira não é acentuada como é nas bordas. Devido a isso, ocorreram confusões no interior da voçoroca com a classe solo exposto. Entretanto, o atributo permitiu a delimitação do contorno da feição erosiva em estudo. Uma alternativa para melhorar este resultado é o uso de informações espaciais, a inserção de informações de contexto e atributos geométricos. Pode-se verificar também o erro de inclusão de elementos de vegetação arbórea na classe voçoroca, devido à declividade existente entre a copa das árvores e o solo nas bordas das áreas vegetadas (áreas circuladas em amarelo). A figura 6 ilustra a classificação resultante.

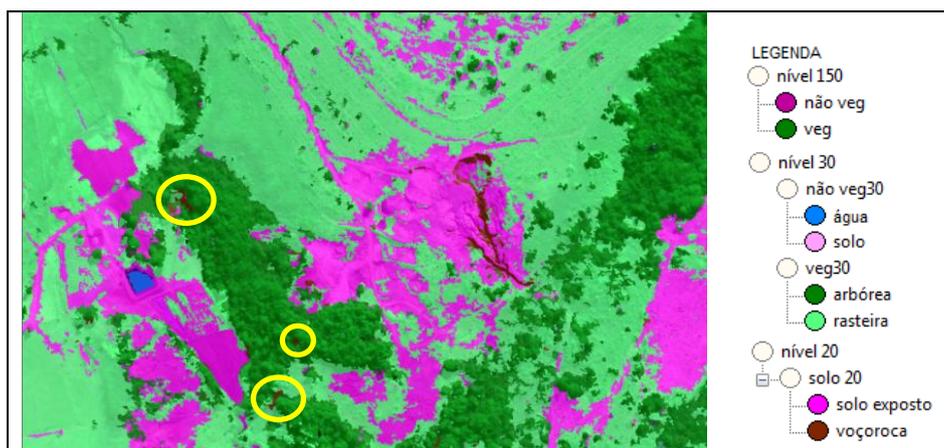


Figura 6. Classificação hierárquica.

3.2 Classificação por árvore de decisão

Comparando-se os atributos do primeiro ramo da árvore de decisão (figura 7) com os atributos definidos para a classificação hierárquica (figura 5b), nota-se que os mesmos são similares, para a separação entre as classes “vegetação” e “não vegetação” (em função do NDVI). No ramo correspondente à “vegetação”, a separação entre as classes “arbórea” e “rasteira” foi realizada em função do MDSN, enquanto que na classificação hierárquica usou-se também o NDVI. No ramo correspondente à “não vegetação”, a separação entre as classes “água” e “solo” foi realizada em função apenas do NDVI, enquanto que na classificação hierárquica usou-se também a banda do infravermelho. Neste mesmo ramo, a separação entre as classes “solo exposto” e “voçoroca” se deu em função da declividade, enquanto que na classificação hierárquica foram usadas também as informações do NDVI e das bandas do verde e infravermelho. Na classificação hierárquica também foi usada, em algumas etapas da classificação, lógica fuzzy, diferenciando-se da classificação por árvore de decisão.

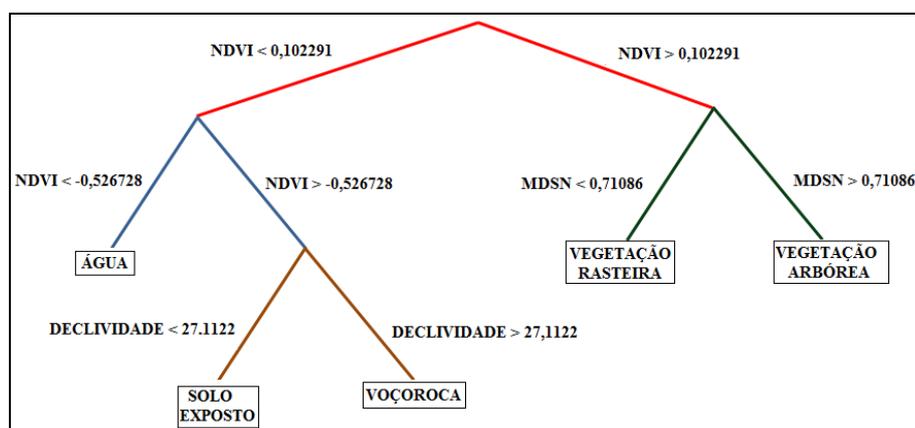


Figura 7. Árvore de decisão gerada com o algoritmo CART.

A matriz de confusão mostrou que uma das amostras selecionadas como solo exposto foi classificada como voçoroca. A qualidade global foi de 98,8%. Analisando a classificação resultante (figura 8), em comparação com o resultado da classificação hierárquica, percebe-se um aumento das áreas classificadas como solo exposto em contraposição às de vegetação rasteira. Essa confusão entre as classes ocorreu nas regiões da cena em que as culturas aparecem em estágio inicial, nas quais é possível entrever o solo abaixo da vegetação. Para a classe voçoroca, as regiões internas da voçoroca foram classificadas corretamente, entretanto, para o entorno da mesma, houve um acréscimo na classificação, alterando o contorno da feição. Além disso, regiões de solo exposto foram incorporadas à classe.

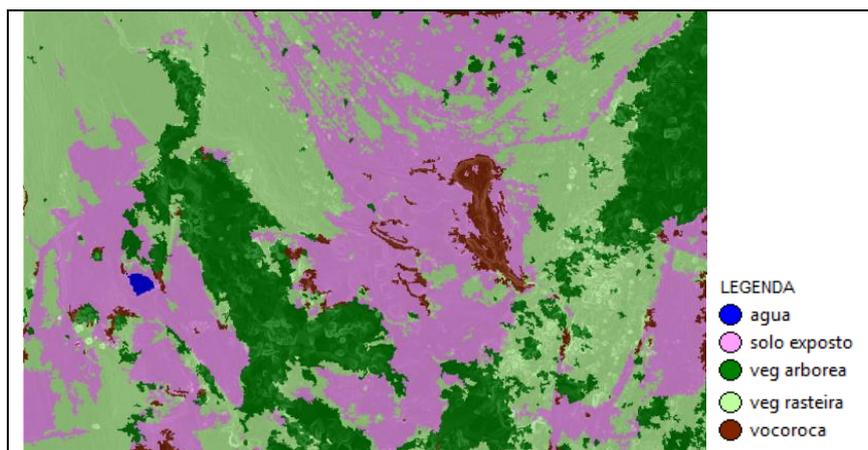


Figura 8. Classificação por árvore de decisão.

4. Conclusões

A classificação resultante do processamento baseado em objetos (regiões), e realizada de forma hierárquica, mostrou a separação entre as classes e um delineamento da feição voçoroca. Esse resultado comprova a eficácia do método proposto, bem como da rede hierárquica estabelecida e dos atributos selecionados para discriminação da feição de interesse. A classificação hierárquica foi desenvolvida com base no pré-conhecimento dos autores, requerendo maior tempo de análise para geração das segmentações, seleção de atributos e seus limiares, e para estabelecimento das regras de pertinência a cada classe. O uso do algoritmo CART agilizou consideravelmente o procedimento de classificação, e a árvore de decisão gerada automaticamente é muito semelhante às regras de decisão definidas para a classificação hierárquica. Considerando-se que a classificação resultante do algoritmo CART foi menos fidedigna à realidade do que a obtida de forma hierárquica, recomenda-se que a árvore de decisão seja gerada preliminarmente, de forma a conduzir a seleção de atributos e seus limiares, os quais devem posteriormente ser refinados de acordo com o conhecimento do intérprete.

Os resultados deste estudo comprovaram que a inserção de informações altimétricas auxilia na discriminação da voçoroca, mas não é suficiente para permitir a extração do contorno da voçoroca. Em função disso, a metodologia proposta vem sendo refinada com a inserção de atributos de forma e de contexto.

Agradecimentos

Os autores agradem à Esteio Engenharia e Aerolevantamentos S.A., que cedeu os dados de ALS e à Engesat Imagens de Satélite, que cedeu a imagem do sensor IKONOS.

Referências

Alkhasawneh, M.; Ngah, U. K.; Tay, L. T.; Isa, N. A. M.; Al-Batah, M. S. Modeling and Testing Landslide Hazard Using Decision Tree. **Journal of Applied Mathematics**, v. 2014, article ID 929768, 2014. Disponível em: <<http://dx.doi.org/10.1155/2014/929768>>. Acesso em: 11 ago. 2014.

Axelsson, P. DEM generation from laser scanner data using adaptive TIN models. **International Archives of Photogrammetry and Remote Sensing**, XXXIII, Part B3, p. 85-92, 2000.

Eustace, A.; Pringle, M.; Witte, C. Give Me the Dirt: Detection of Gully Extent and Volume Using High-Resolution Lidar. In: Jones, S.; Reinke, K. **Innovations in Remote Sensing and Photogrammetry**. New York: Springer, 2009. p. 255-269.

- Guerra, A. J. T. O início do processo erosivo. In: Guerra, A. J. T. (Org.); Silva, A. S.(Org.); Botelho, R. G. M. (Org.). **Erosão e Conservação dos Solos: Conceitos, Temas e Aplicações**. Rio de Janeiro: Bertran Brasil, 2005. cap. 1, p. 17-55.
- Hessel, R., Van Asch, T. Modelling gully erosion for a small catchment on the Chinese Loess Plateau. **Catena**, v. 54, p.131-146, 2003.
- James, L. A.; Watson, D. G.; Hansen, W. Using LiDAR data to map gullies and headwater streams under forest canopy: South Carolina, USA. **Catena**, v. 71, p. 132-144, 2007.
- Johansen, K.; Taihei, S.; Tindall, D.; Phinn, S. Object-based Monitoring of Gully Extent and Volume in North Australia using LiDAR Data. In: International Conference on Geographic Object-Based Image Analysis, 4 - GEOBIA. Maio 7-9, 2012, Rio de Janeiro-RJ. **Anais...Artigos**, p. 168-173. Pen-drive, On-line. ISBN 978-85-17-00059-1. Disponível em: <<http://mtc-m18.sid.inpe.br/col/sid.inpe.br/mtc-m18/2012/05.17.16.49/doc/052.pdf>>. Acesso em: 10 ago. 2014.
- King, C.; Baghdadi, N.; Lecomte, V.; Cerdan, O. The application of remote-sensing data to monitoring and modelling of soil erosion. **Catena**, v. 62, p. 79-93, 2005.
- Martínez-Casasnovas, J. A.; Ramos, M. C.; Poesen, J. Assessment of sidewall erosion in large gullies using multi-temporal DEMs and logistic regression analysis. **Geomorphology**, v. 58, p. 305-321, 2004.
- Marzolf, I; Poesen, J. The potential of 3D gully monitoring with GIS using high-resolution aerial photography and a digital photogrammetry system. **Geomorphology**, v. 111, p. 48-60, 2009.
- Morgan, R. P. C. **Soil Erosion and Conservation**. Oxford: Blackwell Publishing, 2005. 304 p.
- Perroy, R. L.; Bookhagen, B.; Asner, G. P.; Chadwick, O. A. Comparison of gully erosion estimates using airborne and ground-based LiDAR on Santa Cruz Island, California. **Geomorphology**, v. 118, p. 288-300, 2010.
- Shruthi, R. B. V.; Kerle, N.; Jetten, V. Object based gully feature extraction using high spatial resolution imagery. **Geomorphology**, v. 134, p. 260-268, 2011.
- Shruthi, R. B. V.; Kerle, N.; Jetten, V. Extracting gully features and its dynamics from high spatial resolution imagery using object based image analysis. In: International Conference on Geographic Object-Based Image Analysis, 4 - GEOBIA. Maio 7-9, 2012, Rio de Janeiro-RJ. **Anais...Artigos**, p. 7-12. Pen-drive, On-line. ISBN 978-85-17-00059-1. Disponível em: <<http://mtc-m18.sid.inpe.br/col/sid.inpe.br/mtc-m18/2012/05.16.19.53/doc/008.pdf>>. Acesso em: 10 ago. 2014.
- Vieira, N. M. **Estudo geomorfológico das voçorocas de Franca-SP**. 1978. Tese - Faculdade de Filosofia, Ciências e Letras da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Franca. 1978.
- Vrieling, A.; Rodrigues, S. C.; Sterk, G. Evaluating erosion from space: a case study near Uberlândia. **Revista Sociedade & Natureza**, Edição Especial, p. 683-696, 2005.
- Wang, T.; He, F.; Zhang, A.; Gu, L.; Wen, Y.; Jiang, W.; Shao, H. A Quantitative Study of Gully Erosion Based on Object-Oriented Analysis Techniques: A Case Study in Beiyanzikou Catchment of Qixia, Shandong, China. **The Scientific World Journal**, v. 2014, article ID 417325, 2014. Disponível em: <<http://dx.doi.org/10.1155/2014/417325>>. Acesso em: 25 jul. 2014.
- Zhang, K.; Cui, Z. Airborne LIDAR Data Processing and Analysis Tools. **National Center for Airborne Laser Mapping**. 2007. Disponível em: <<http://lidar.ihrc.fiu.edu/download/Doc/ALDPAT.pdf>>. Acesso em: 25 jul. 2014.