



Revista Brasileira de Cartografia (2012) N<sup>o</sup> 64/5: 645-660  
Sociedade Brasileira de Cartografia, Geodésia, Fotogrametria e Sensoriamento Remoto  
ISSN: 1808-0936

## **CLASSIFICAÇÃO DE COBERTURA DA TERRA DE NOVA FRIBURGO-RJ UTILIZANDO DUAS ABORDAGENS: MINERAÇÃO DE DADOS E REDE HEURÍSTICA**

*Land Cover Classification in Nova Friburgo-RJ, Brazil, Using Two Approaches: Data Mining and Heuristic Semantic Network*

**Cristiane Nunes Francisco<sup>1</sup> & Cláudia Maria de Almeida<sup>2</sup>**

**<sup>1</sup>Universidade Federal Fluminense – UFF**  
**Instituto de Geociências – Departamento de Análise Geoambiental**  
Campus da Praia Vermelha – Boa Viagem – Niterói – RJ - 24210-310  
crisnf@vm.uff.br

**<sup>2</sup>Instituto Nacional de Pesquisas Espaciais – INPE**  
**Coordenação de Observação da Terra – Divisão de Sensoriamento Remoto – DSR**  
Av. dos Astronautas, 1758 - Jardim da Granja - São José dos Campos, SP - 12227-010  
almeida@dsr.inpe.br

*Recebido em 15 de setembro, 2011/ Aceito em 25 de novembro, 2011*  
*Received on september 15, 2011/ Accepted on november 25, 2011*

### **RESUMO**

Este artigo tem como objetivo avaliar a classificação de cobertura da terra realizada por análise de imagens baseada em objeto (*object-based image analysis* – OBIA), utilizando duas diferentes abordagens na geração da rede semântica: i) por meio de mineração de dados, e ii) de forma heurística. Para isso, utilizaram-se descritores estatísticos e texturais extraídos de imagens ALOS/AVNIR fusionadas com imagens ALOS/PRISM e dados de relevo geomorfométricos do projeto TOPODATA. A área de estudo é o município de Nova Friburgo, com área de 933 km<sup>2</sup>, localizado na região serrana do estado do Rio de Janeiro. O índice *Kappa* alcançado na classificação baseada na mineração foi de 0,82, enquanto que este valor para a rede criada heurísticamente foi de 0,85. Considerando os índices alcançados, conclui-se que ambos os resultados apresentam excelente qualidade quanto à acurácia da classificação. O teste de hipótese entre os dois índices mostra, com nível de significância de 5%, que não há diferenças entre as duas classificações quanto à sua acurácia. Esses valores de *Kappa* são compatíveis aos obtidos em outras classificações de cobertura da terra que utilizaram OBIA, com número de classes e resolução espacial semelhantes.

**Palavras chaves:** Análise de Imagens Baseada em Objeto, Classificação de Imagens, Mineração de Dados, ALOS.

### **ABSTRACT**

This paper aims to evaluate two methods for classifying land use and land cover using object-based image analysis (OBIA): i) the first one was based on a semantic network created by data mining, and ii) the second one resulting from a semantic network heuristically designed. Statistical and textural attributes extracted from ALOS/AVNIR images pan-

sharpened with ALOS/PRISM as well as relief data from TOPODATA geomorphometric database were used as input data. The study area is Nova Friburgo County, with an area of 933 km<sup>2</sup>, located in the mountainous region of Rio de Janeiro State. The Kappa index obtained by the classification based on data mining was 0.82, while the result for the heuristically elaborated network achieved 0.85. These values indicate the high accuracy of both results. The statistical hypothesis test between the two indices with a 95% confidence interval demonstrated that there are no differences between the two classification accuracies. The Kappa values obtained in this paper are analogous to the ones found by similar works in the literature, which employed a comparable number of classes and spatial resolution.

**Keywords:** Object-based Image Analysis, Images Classification, Data Mining, ALOS.

## 1. INTRODUÇÃO

As técnicas tradicionais de classificação de imagens de sensoriamento remoto baseiam-se na análise pixel a pixel ou por regiões, enfocando as diferenças espectrais para a extração de informações temáticas. Nas técnicas de análise de imagens baseada em objeto (*object-based image analysis* - OBIA), os pixels são agrupados em objetos de acordo com as suas propriedades não somente espectrais, mas também espaciais, através da técnica de segmentação, e estes passam a ser utilizados como unidades primitivas para a classificação (NAVULUR, 2006). De maneira diversa dos classificadores tradicionais, que utilizam somente a informação espectral para identificar a classe provável a ser atribuída ao elemento de análise, a OBIA baseia-se em informações extraídas de objetos que possuem identidade própria (PINHO, 2005a). Assim, além das propriedades espectrais, os objetos apresentam numerosos atributos (descritores), associados à forma, textura, relações contextuais e semânticas, que podem ser utilizados para a análise de imagens, aproximando-se dos processos cognitivos humanos de interpretação de imagens (NAVULUR, 2006; ALMEIDA *et al.*, 2009; MARPU, 2009).

A OBIA consiste basicamente na aplicação de dois procedimentos metodológicos: (1) segmentação/regionalização, técnica primária usada para conversão de uma cena ou imagem em múltiplos objetos; e (2) classificação baseada em regras de decisão, que explicitam as propriedades dos objetos expressas pelos atributos (NAVULUR, 2006; LANG, 2008). Os algoritmos de segmentação subdividem uma imagem formada por pixels em regiões, diminuindo o nível de detalhe e a complexidade da imagem, que passa a ser formada por objetos (LANG, 2008). A imagem segmentada é constituída por um conjunto de pixels contíguos,

sendo cada pixel rotulado como pertencente estritamente a um objeto (DEFINIENS, 2007). O objeto corresponde a uma região discreta de uma imagem que é internamente coerente e diferente do seu entorno (CASTILLA; HAY, 2008).

Após a segmentação, segue-se a fase de criação da base de conhecimento considerando o objeto, que é estruturada em uma rede semântica hierárquica, responsável por armazenar o conhecimento do intérprete sobre a área de estudo e com base na qual os objetos são classificados. Durante a construção da base do conhecimento, devem ser considerados os seguintes pontos: (a) definição de classes e subclasses apropriadas para o nível de segmentação; (b) determinação dos atributos dos objetos (espectrais, texturais, morfológicos, espaciais), que os classificam e são herdados pelas subclasses; e (c) determinação das funções de pertinência (tipo de função e limiares) (MAVRANTZA; ARGIALAS, 2008). Blaschke *et al.* (2008) sugerem que o foco de OBIA deva estar na incorporação e desenvolvimento de bases geográficas inteligentes, ou seja, informações apropriadas dentro de um contexto geográfico.

A construção da base de conhecimento, representada pela rede semântica, é uma das tarefas mais importantes para a classificação dos objetos e, no entanto, uma das mais trabalhosas, devido à dificuldade para selecionar, entre uma grande quantidade de atributos fornecida pelos objetos, quais descrevem melhor as classes às quais pertencem. A rede semântica pode ser gerada de forma heurística pelo usuário, testando iterativa e interativamente os descritores, as funções e os seus limiares para a discriminação adequada das classes, ou de forma automática, através da aplicação de técnicas de mineração de dados. Estas consistem na extração de conhecimento de uma base com grande volume de dados por meio de métodos inteligentes. O modelo derivado da mineração pode

ser representado de várias formas, entre elas, as árvores de decisão, representadas por um fluxograma com estrutura de árvore, e que são facilmente convertidas em regras de classificação (HAN; KAMBER, 2006).

Existem diferenças entre as redes geradas heurísticamente e por mineração de dados. Enquanto a primeira rede pode trabalhar com lógica *fuzzy* na construção das funções para discriminação das classes, a rede gerada por mineração de dados apresenta apenas funções de pertinência *crisp*. Há plataformas de OBIA que permitem operar com mais de um nível de segmentação para classificação, o que possibilita definir relações contextuais entre os níveis. As redes geradas automaticamente operam apenas com um nível de segmentação, impossibilitando esta abordagem. Por fim, a rede derivada de mineração não opera com mais de um atributo por nó e, portanto, tampouco com redundância em nível de classe ou nó, porém, há casos em que ela é desejável para discriminações finas entre classes.

Considerando essas possibilidades, este artigo tem como objetivo avaliar a classificação de cobertura da terra realizada por OBIA, utilizando duas diferentes abordagens na geração da rede semântica: i) por meio de mineração de dados, e ii) de forma heurística. Para isso, utilizaram-se descritores estatísticos e texturais extraídos de dados espectrais e geomorfométricos. Os primeiros descritores referem-se àqueles baseados no tratamento estatístico (média e desvio-padrão, no caso deste trabalho) dos valores dos pixels de cada objeto, enquanto os segundos correspondem à distribuição espacial dos valores do pixel no respectivo objeto.

A área de estudo é o município de Nova Friburgo, com área de 933 km<sup>2</sup> e população de 182 mil habitantes (IBGE, 2011a), localizado na região serrana fluminense, na unidade geomorfológica denominada como Planalto Reverso da Região Serrana (DANTAS, 2007). O relevo desta unidade é montanhoso a escarpado, com altitude variando entre 400 e 2.300 m. Os terrenos mais íngremes e elevados, devido à difícil acessibilidade, mantêm a cobertura florestal preservada, representando cerca de 70% do seu território (CIDE, 2003). O sítio urbano está instalado em alvéolos intramontanos, enquanto os núcleos agrícolas concentram-se nos fundos de

vales e baixas encostas de relevo menos acidentado. Essa unidade geomorfológica apresenta alta vulnerabilidade a eventos erosivos e movimentos de massa, devido às características geológicas e geomorfológicas (declividade e amplitude do relevo altas), além da ocupação das encostas associada à expansão das atividades econômicas na região (DANTAS, 2007).

## 2. MATERIAL E MÉTODOS

No mapeamento de cobertura da terra de Nova Friburgo, aplicou-se OBIA com o uso da plataforma Definiens Developer 7.04. Nela, estão contidos algoritmos de segmentação, entre eles o de multirresolução, e de classificação, abrangendo classificadores baseados na regra do vizinho mais próximo e em lógica *fuzzy* ou nebulosa (ECOGNITION, 2011).

O algoritmo de segmentação multirresolução gera um nível de imagem com objetos criados a partir da fusão de pixels de acordo com critérios de homogeneidade (DEFINIENS, 2007). Quando aplicado a um banco de dados com um nível de segmentação existente, irá fundir objetos em superobjetos em níveis superiores, ou então, subdividi-los em subobjetos em níveis inferiores, em ambos os casos, respeitando as fronteiras dos níveis existentes, o que possibilita a construção de uma rede hierárquica entre os níveis de segmentação.

Os critérios de homogeneidade utilizados na segmentação multirresolução são estabelecidos a partir da combinação das propriedades espectrais (cor) e de forma dos pixels ou dos objetos. Os parâmetros cor e forma somam 1,0 e determinam, respectivamente, o quanto da informação de cor e forma é usada no processo de segmentação. O parâmetro forma é subdividido em compacidade e suavidade. Um alto valor de compacidade produz segmentos menores e mais compactos, tipicamente relativos a alvos antrópicos, como construções. Por outro lado, um alto valor de suavidade gera segmentos com bordas suaves, características de alvos naturais (KRESSLER; STEINNOCHER, 2006).

O parâmetro escala é um importante valor utilizado pelo algoritmo de segmentação multirresolução e determina a média do tamanho do objeto a ser criado (DEFINIENS, 2007). Quanto mais alto o valor atribuído ao parâmetro,

maiores serão os objetos e, com isto, menor a quantidade de objetos criados.

Deve ser atribuído, ainda, peso às imagens de acordo com o seu grau de importância no processo de segmentação. O peso varia entre 0 e 1, e quanto maior o valor, maior o grau de importância conferida à respectiva imagem de entrada, sendo que o valor 0 implica a desconsideração da respectiva imagem no processo de segmentação (DEFINIENS, 2007).

Para a mineração de dados, foi utilizado o programa WEKA 3.6.4, desenvolvido pela Universidade de Waikato, Nova Zelândia. Este programa de código aberto apresenta uma coleção de algoritmos de aprendizagem, constituída por ferramentas de pré-processamento, classificação, regressão, agrupamento e regras de associação, que também podem ser aplicadas diretamente no conjunto de dados (HALL *et al.*, 2009).

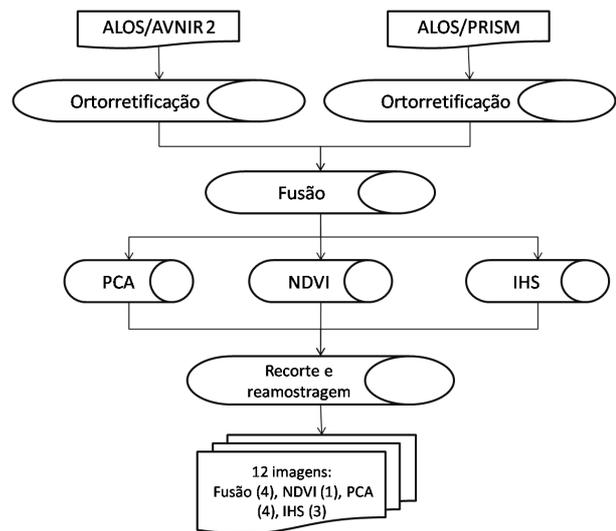
O banco de dados foi composto por imagens orbitais e dados geomorfométricos (declividade e altimetria) do projeto TOPODATA, disponível em <http://www.dsr.inpe.br/topodata> (VALERIANO, 2005). As imagens foram obtidas em agosto de 2009 pelo sensor AVNIR-2 do satélite ALOS, com resolução espacial de 10 m, correspondendo a três bandas do visível e uma do infravermelho próximo, fusionadas com a imagem pancromática do sensor PRISM (ALOS), com resolução espacial de 2,5 m.

As etapas gerais do trabalho, descritas a seguir, consistiram na preparação da base de dados, segmentação das imagens, criação da rede semântica, classificação das imagens e validação da classificação.

## 2.1. Preparação da base de dados

A preparação do banco de dados consistiu no processamento das imagens ALOS e do modelo digital de elevação (MDE) do TOPODATA. As imagens multiespectrais e pancromática foram inicialmente ortoretificadas e fusionadas. Novas imagens também foram geradas por meio de operações e transformações matemáticas (figura 1).

Como os coeficientes polinomiais racionais (*Rational Polynomial Coefficients* - RPC) das imagens ALOS são fornecidos apenas para dados obtidos no continente asiático, foram coletados 40 pontos de controle nas ortofotos 1:25.000 (IBGE, 2011b) com 1 m de resolução espacial para



**Fig. 1** - Etapa de preparação da base de dados.

ortoretificar as cenas AVNIR (70x70 km). A altitude foi extraída dos dados do MDE do TOPODATA. Para a imagem PRISM (70x35 km), foram utilizados 20 pontos. A ortoretificação foi executada no programa ENVI 4.7.

O passo seguinte foi a fusão de imagens AVNIR-2 e PRISM utilizando o método *Gram-Schmidt*, disponível no ENVI 4.7. Este método consiste na simulação de uma banda pancromática com as multiespectrais, aplicando, então, a transformação de *Gram-Schmidt* sobre as multiespectrais e a pancromática simulada. Ao final, uma transformação inversa é efetuada, com a substituição da primeira banda pela pancromática. O método *Gram-Schmidt* apresenta maior acurácia do que o de Análise das Componentes Principais (*Principal Components Analysis* – PCA), uma vez que utiliza a função da resposta espectral do sensor para simular a banda pancromática (ENVI, 2009). Em trabalho realizado por Pinho *et al.* (2005b), objetivando a avaliação de métodos de fusão aplicados às imagens QuickBird, concluiu-se que os métodos “Componentes Principais e *Gram-Schmidt*, quando aplicados em fusões com as quatro bandas multiespectrais, apresentam os melhores resultados”. Neste trabalho, os autores também constataram a fidelidade espectral da imagem sintética com relação à imagem multiespectral original.

A partir das bandas fusionadas, foram criadas novas imagens resultantes da aplicação PCA e da conversão do sistema RGB para o IHS (*Intensity* ou Intensidade – I, *Hue* ou Matiz – H, *Saturation* ou Saturação – S). Também foi calculado o Índice

de Vegetação por Diferença Normalizada (*Normalized Difference Vegetation Index - NDVI*). Depois disso, as imagens foram recortadas pelo limite da área de estudo e reamostradas para a resolução espacial de 2,5 m. Esse processamento foi executado no sistema ENVI 4.7.

O passo seguinte foi a geração da grade de declividade utilizando o interpolador TOPOTORASTER, disponível na extensão *3D Analyst* do ArcGIS 9.0, e o MDE TOPODATA. As bases topográficas foram igualmente recortadas pelo limite da área de estudo e reamostradas para 2,5 m.

No total, o banco de dados foi composto por 14 planos de informação assim distribuídos: quatro bandas multiespectrais fusionadas do sistema ALOS, quatro componentes principais (CP), três componentes IHS, NDVI, MDE e grade de declividade.

## 2.2. Segmentação das imagens

Foram realizados três níveis de segmentação consecutivos, com a diminuição do fator de escala na execução de cada novo procedimento, originando, assim, um novo nível com maior número de objetos de tamanho reduzido.

Devido à resolução espacial das imagens utilizadas na classificação de cobertura da terra neste trabalho, o fator espectral ou cor tende a ser mais importante na segmentação do que a forma (tabela 1). Por isto, nas três segmentações efetuadas, ao parâmetro forma foi atribuído o valor 0,1 e, por

**Tabela 1** - Parâmetros dos três níveis de segmentação.

Nível	Fator escala	Imagens peso 1	Peso forma	Peso compacidade
3	500	NDVI	0,1	0,5
2	250	NDVI	0,1	0,5
1	160	4 bandas multiespectrais	0,1	0,5

consequente, a cor assumiu o fator 0,9. Para o parâmetro compacidade, atribui-se o valor 0,5 e, assim, o fator suavidade assumiu o mesmo valor, visto que é complementar de 1.

Para o primeiro e segundo níveis de segmentação, foi atribuído o peso 1 apenas para a imagem de NDVI, objetivando separar os objetos com vegetação daqueles sem vegetação. No terceiro nível de segmentação, foram utilizadas as

quatro bandas multiespectrais, visando discriminar os diversos tipos de alvos presentes na cena.

## 2.3. Geração da rede semântica

As classes foram estabelecidas com base na análise visual das imagens fusionadas, sendo posteriormente ajustadas de forma iterativa com os resultados preliminares da classificação. Desse modo, foram definidas sete classes de cobertura: afloramento rochoso, floresta, vegetação herbácea, queimada, reflorestamento, sombra e área urbana (quadro 1).

Para a geração da rede semântica por mineração de dados, foram selecionadas 195 amostras de treinamento, correspondentes a sete classes de cobertura da terra, ou seja, 30 amostras por classe, com exceção da queimada, pois, como apresenta área reduzida, possui pequeno número de objetos. O conjunto inicial de dados para mineração, composto por atributos estatísticos e texturais extraídos do primeiro nível de segmentação, foi gerado no Definiens 7.04. Depois dos procedimentos de importação e limpeza dos dados (para remoção de ruídos e inconsistências), executados no Weka 3.6.4, este conjunto constituiu-se de 130 descritores estatísticos e texturais (tabela 2).

Para o cálculo dos descritores texturais, feito pelo Definiens 7.04, são utilizadas as matrizes de co-ocorrência do nível de cinza (*Gray Level Cooccurrence Matrix - GLCM*) propostas por Haralick *et al.* (1973). Esta abordagem extrai a textura pela relação espacial existente entre os níveis de cinza ao longo de diferentes direções. Devido à sua simplicidade e eficácia, é considerada por muitos autores como referência (ITO *et al.*, 2011). Entre os 14 parâmetros de textura disponíveis, foram utilizados quatro considerados de maior relevância (BARALDI; PARMIGGIANI, 1995; ITO *et al.*, 2011):

- Segundo momento angular ou Energia – mede a uniformidade da textura, ou seja, a quantidade de repetição de pares de pixels. Valores elevados significam que a distribuição do nível de cinza é constante, isto é, grande repetição na variação dos níveis de cinza. Para uma matriz normalizada, os valores são positivos e menor ou igual a 1.
- Entropia – mede a desordem da imagem, ou seja, corresponde a uma medida

inversamente proporcional ao segundo momento angular. Valores elevados indicam que a imagem não possui textura uniforme.

- **Contraste** – corresponde à diferença entre os valores mais altos e mais baixos de um conjunto de pixels contíguos. Elevados valores de contraste correspondem a imagens compostas por níveis de cinza de grande amplitude, ou seja, com textura rugosa.
- **Homogeneidade** – mede a homogeneidade da imagem. Valores elevados correspondem a pequenas diferenças tonais entre os pixels vizinhos. Este parâmetro é

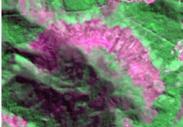
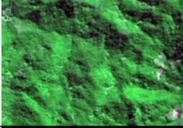
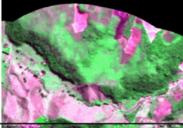
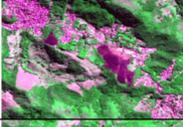
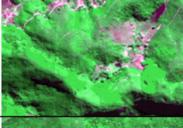
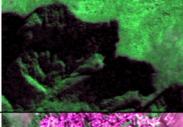
**Tabela 2** - Descritores utilizados na mineração de dados.

Descritor	Tipo	Total
Média	Estatístico	16
Desvio padrão		12
GLCM 2º Momento Angular	Textural	27
GLCM Contraste		28
GLCM Entropia		30
GLCM Homogeneidade		17

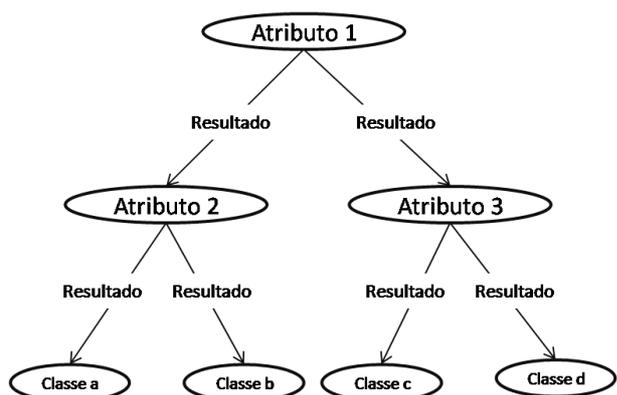
inversamente proporcional ao contraste e à energia.

As regras de classificação por mineração de dados foram estabelecidas pelo algoritmo C4.5, criado por Quinlan (1993) e implementado como classificador tree.J48 no programa Weka 3.6.4. Este algoritmo constrói árvores de decisão a partir de amostras de treinamento. As árvores são expressas por um fluxograma, onde o nó interno denota um teste com o atributo, o ramo representa o resultado do teste, e o nó externo (folha) exibe a classe esperada (figura 2). Para cada nó, o algoritmo escolhe o melhor atributo para separar os dados em classes individuais. Os atributos que não aparecem na árvore são considerados irrelevantes (HAN; KAMBER, 2006). O algoritmo permite construir árvores menores, por serem mais facilmente compreendidas e apresentarem bom desempenho. Para isto, utiliza a entropia para medir o quanto o nó é informativo. Quanto menor a entropia, menos informação será utilizada para descrever os dados (SILVA, 2006).

**Quadro 1** - Classes de cobertura da terra.

Classes	Descrição	Chave de interpretação	ALOS/AVNIR 3R_4G_2B
Afloramento rochoso	Rocha exposta com presença de pouca ou nenhuma vegetação.	<ul style="list-style-type: none"> <li>• Cor – magenta</li> <li>• Textura – intermediária</li> <li>• Forma – arredondada</li> <li>• Localização – encostas íngremes</li> <li>• Padrão – listrado</li> </ul>	
Floresta	Vegetação arbórea, em estágio médio e avançado de regeneração, e arbórea-arbustiva, em estágio inicial e médio.	<ul style="list-style-type: none"> <li>• Cor – verde escuro</li> <li>• Textura – intermediária a rugosa</li> </ul>	
Herbácea	Vegetação predominantemente composta por gramíneas e ervas.	<ul style="list-style-type: none"> <li>• Cor – verde claro (herbácea) e magenta claro (herbácea rala)</li> <li>• Textura – lisa</li> </ul>	
Queimada	Área atingida por fogo.	<ul style="list-style-type: none"> <li>• Cor – roxo escuro</li> <li>• Textura – lisa</li> </ul>	
Reflorestamento	Vegetação predominantemente arbórea e composta por eucaliptos.	<ul style="list-style-type: none"> <li>• Cor – verde intermediário</li> <li>• Textura – lisa</li> </ul>	
Sombra	Áreas sem informação devido à ausência de incidência solar.	<ul style="list-style-type: none"> <li>• Cor – preta</li> <li>• Textura - lisa</li> <li>• Localização – encostas</li> </ul>	
Urbana	Predominância de construções.	<ul style="list-style-type: none"> <li>• Cor – magenta</li> <li>• Textura – rugosa</li> <li>• Padrão – reticulado</li> </ul>	

**Fig. 2** – Esquema estrutural de uma árvore de decisão gerada pelo algoritmo C4.5.



A rede semântica gerada de forma heurística foi obtida diretamente no programa Definiens 7.04. As funções e seus limiares foram definidos iterativa e interativamente pelo usuário através da análise da distribuição espacial, por objeto, dos valores médios dos descritores em tons de cinza (comando *Feature View*), e dos histogramas da função de pertinência, por atributo, dos objetos para cada classe.

#### 2.4. Classificação de cobertura da terra

Foram geradas classificações baseadas em redes semânticas geradas heurísticamente e por mineração de dados. As árvores de decisão geradas no Weka 3.6.4 foram implementadas em redes semânticas na plataforma Definiens 7.04. Foi selecionada, para posterior validação, a árvore de decisão que apresentou simultaneamente uma estruturação lógica e um bom resultado de classificação.

#### 2.5. Avaliação da acurácia da classificação de cobertura da terra

Para a avaliação da acurácia das classificações, foram coletadas 1,4 mil amostras de segmentos estratificadas pela área esperada das sete classes. Tentou-se obedecer ao número mínimo de 50 amostras por classe, definido por Congalton e Green (2009), para mapas cobrindo menos de um milhão de acres e com menos de 12 classes. No entanto, devido à reduzida área de algumas classes, não foi possível a observância deste número para todas elas. Congalton e Green (2009) também afirmam que a unidade da amostra deve ser um polígono, caso o mapa a ser avaliado tenha esta estrutura, como aqueles oriundos de processos de segmentação.

A seguir, foram construídas as matrizes de erros compostas, nas colunas, pelas amostras de referência, identificadas com base nas ortofotos do projeto RJ-25 (IBGE, 2011b), e nas linhas, pelas classes obtidas na classificação. As ortofotos foram geradas por levantamento aerofotogramétrico realizado em 2005, com resolução espacial de 0,7 m e escala aproximada de 1:30.000. Na matriz, são expressos os erros de omissão, ou seja, amostras que não foram classificadas de acordo com as classes de referência, e os erros de comissão, correspondentes a amostras de referência classificadas erroneamente como pertencentes a outras classes. A partir das matrizes, são calculados os seguintes índices: (a) exatidão global - relação entre o número de amostras classificadas corretamente sobre o número total de amostras de referência; (b) exatidão do produtor – relativa aos erros de omissão, a qual representa a relação entre o número de amostras classificadas corretamente da classe *k* e o número total de amostras de referência da classe *k*, (c) exatidão do usuário - referente aos erros de comissão, a qual representa a relação entre o número de amostras classificadas corretamente da classe *k* e o número total de amostras classificadas da classe *k*; (d) *Kappa* (*K*); e (e) *Kappa* condicional (CONGALTON; GREEN, 2009).

O índice *Kappa* varia de -1 a 1, e quanto mais próximo de 1, maior a precisão da classificação. De acordo com Landis e Koch (1977), a qualidade da classificação pode ser definida em função dos seguintes intervalos de índice *Kappa*: [ $<0,00$ ]: muito ruim; [ $0,00 - < 0,20$ ]: ruim; [ $0,20 - < 0,40$ ]: aceitável; [ $0,40 - < 0,60$ ]: bom; [ $0,60 - < 0,80$ ]: muito bom; [ $0,80 - 1,00$ ]: excelente.

O índice *Kappa* condicional avalia a acurácia de cada classe e é calculado com base no mesmo princípio do índice *Kappa* utilizado para a avaliação global da classificação, assim como a interpretação dos resultados obedece aos mesmos princípios. O índice *Kappa* condicional do usuário avalia os erros de comissão (equação 1), e o do produtor, os erros de omissão (equação 2) (ROSSITER, 2004).

$$K_u = \frac{nn_{ii} - n_{i+} n_{+i}}{nn_{i+} - n_{+i} n_{i+}} \quad (1)$$

$$K_p = \frac{nn_{ii} - n_{ii} n_{+i}}{nn_{i+} - n_{+i} n_{i+}} \quad (2)$$

em que:

$n_{ii}$  – número total de amostras corretamente classificadas da classe  $k$ ;

$n_{i+}$  – número total de amostras classificadas da classe  $k$ ;

$n_{+i}$  – número total de amostras coletadas da classe  $k$ ; e

$n$  – número total de amostras.

Por fim, para testar a significância estatística da diferença entre os índices *Kappa*, calculados para a classificação por redes semânticas geradas por mineração de dados e de forma heurística, foi utilizado o teste *Z* (CONGALTON; GREEN, 2009), de acordo a seguinte expressão (equação 3):

$$Z = \frac{K_2 - K_1}{\sqrt{\sigma_{K_2}^2 + \sigma_{K_1}^2}} \quad (3)$$

$K1$  = índice *Kappa* da classificação 1;

$K2$  = índice *Kappa* da classificação 2; e

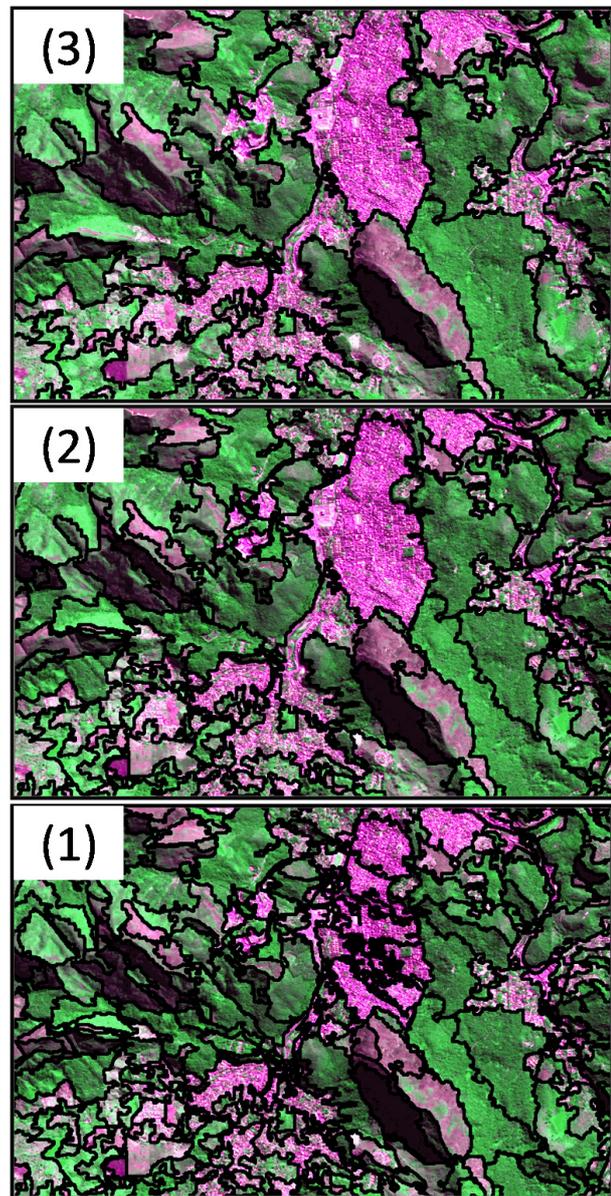
$s2$  = variância do índice *Kappa*.

### 3. RESULTADOS

#### 3.1 Cobertura da terra de Nova Friburgo

A figura 3 apresenta os três níveis de segmentação efetuados para a rede semântica heurística. No nível 3, correspondente ao fator de escala 500 e peso 1 para a camada NDVI, verifica-se a delimitação de objetos com e sem vegetação. No nível 2, com fator de escala 250 e peso 1 para a camada NDVI, os segmentos delimitam objetos de vegetação de acordo com a iluminação das encostas. No nível 1, com fator de escala 160 e peso 1 para as bandas multiespectrais, observa-se a maior discriminação de alvos da cena.

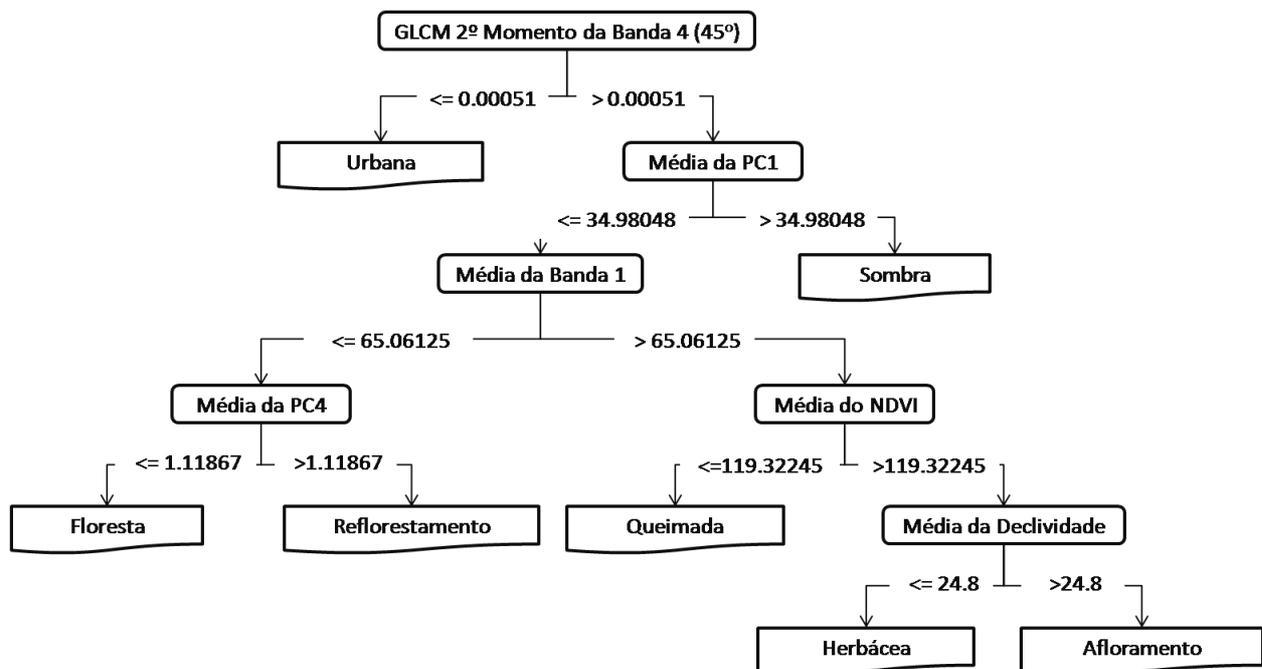
A figura 4 apresenta a árvore de decisão, treinada com base em amostras extraídas do nível 1 e referentes ao conjunto de descritores estatísticos e texturais. As regras de decisão fornecidas por esta árvore constituíram limiares *crisp* da rede semântica hierárquica criada para o processo de classificação de cobertura da terra de Nova Friburgo (figura 5) pelo método de mineração de dados. Dos 130 descritores, os seguintes foram utilizados para descrever os padrões das classes de cobertura da terra: segundo momento angular, primeira componente principal (CP1), banda 4, NDVI, quarta componente principal (CP4) e declividade.



**Fig. 3** - Níveis de segmentação por fator de escala: 500 (3), 250 (B) e 160 (C).

Inicialmente, a árvore discrimina a área urbana a partir do segundo momento angular da banda 4. Conforme anteriormente descrito, este descritor mede a uniformidade da textura, sendo que os valores próximos de 0 indicam áreas heterogêneas ou com textura rugosa (BARALDI; PARMIGGIANI, 1995). No segundo ramo da árvore, há a discriminação de sombra a partir da primeira componente principal, a qual possui a maior parte da variância total dos dados, ou seja, a informação que é comum a todas as bandas originais, contendo “a informação de brilho associada às sombras de topografia” (SPRING, 2011).

A partir do terceiro nível, a árvore divide-se em dois grandes ramos. No ramo esquerdo, estão



**Fig. 4 -** Árvore de decisão com descritores estatísticos e texturais.

presentes as classes com vegetação de porte arbóreo-arbustivo, discriminadas pela banda 1. A classe reflorestamento é discriminada de floresta pela quarta componente principal. O reflorestamento possui baixa reflectância e alta correlação entre as bandas do visível, assemelhando-se à resposta espectral da floresta. Os altos valores de reflectância do reflorestamento no infravermelho próximo, mais elevados do que os de floresta, fazem com que a quarta componente destaque esta classe. No quarto ramo do lado direito da árvore, os menores valores de NDVI são utilizados para discriminar a classe queimada de afloramento rochoso (que por vezes apresenta esparsa vegetação rupestre) e de herbácea, sendo que estas últimas são diferenciadas pela declividade, o que indica a localização dos afloramentos rochosos em encostas de maior declive.

A figura 6 apresenta a classificação de cobertura da terra de Nova Friburgo gerada através de rede semântica heurística constituída por três níveis de segmentação. No terceiro nível, correspondente à segmentação com fator de escala 500 (figura 7), o brilho foi utilizado para discriminar as áreas de sombra e não-sombra. O segundo nível utilizou a segmentação com fator de escala 250 (figura 8), e as classes arbórea-arbustiva e não arbórea-arbustiva foram discriminadas a partir dos descritores referentes à soma das médias das bandas 2 e 3 e à média da segunda componente principal.

No primeiro nível (figura 9), com fator de escala 160, as classes floresta e reflorestamento, pertencentes à classe arbórea-arbustiva, foram discriminadas a partir da soma das médias da segunda e quarta componentes. Para discriminar o afloramento rochoso, foram utilizados os descritores NDVI, soma das médias das bandas 2, 3 e 4, e a média da declividade. A queimada foi discriminada a partir da média da segunda componente principal e média da banda 4. Para a classe urbana, foram utilizados os descritores contraste da banda 4, desvio-padrão da terceira componente principal e soma das médias da primeira e terceira componentes principais, componente intensidade I (IHS) e banda 1. A classe herbácea correspondeu à não-existência de todas as outras classes.

### 3.2. Acurácia das classificações de cobertura da terra

A tabela 3 refere-se à matriz de erros da classificação de cobertura da terra gerada por árvore de decisão com os descritores texturais e estatísticos. O índice de exatidão global alcançado foi de 90%, e o índice *Kappa* de 0,82, este último considerado de excelente qualidade de acordo com Landis e Koch (1977). Apenas a classe urbana apresentou erros de omissão superiores a 50%, ou seja, exatidão do produtor de 40%, confundindo-se com herbácea e queimada. Por outro lado, esta mesma classe não apresentou erros de comissão,

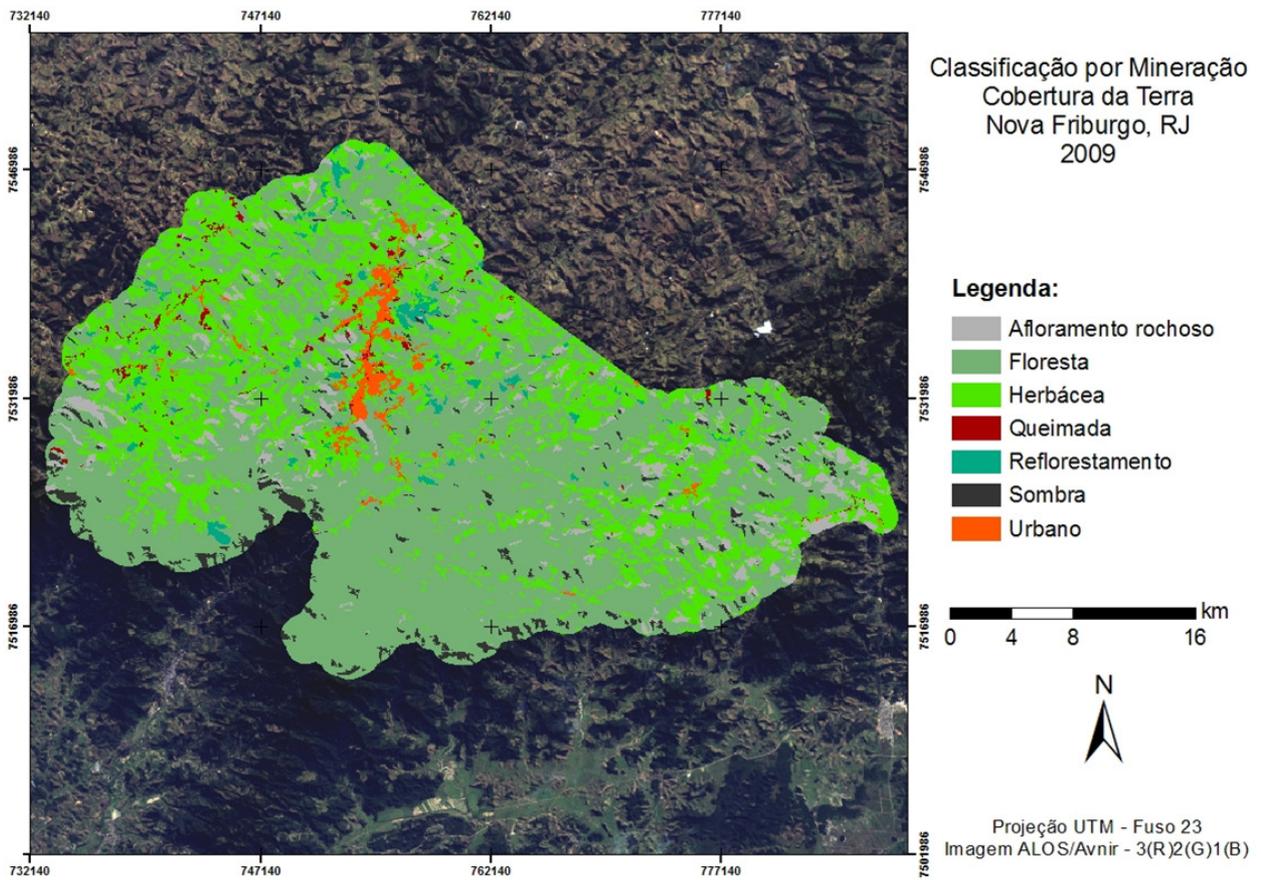


Fig. 5 - Classificação de cobertura da terra de Nova Friburgo por mineração de dados.

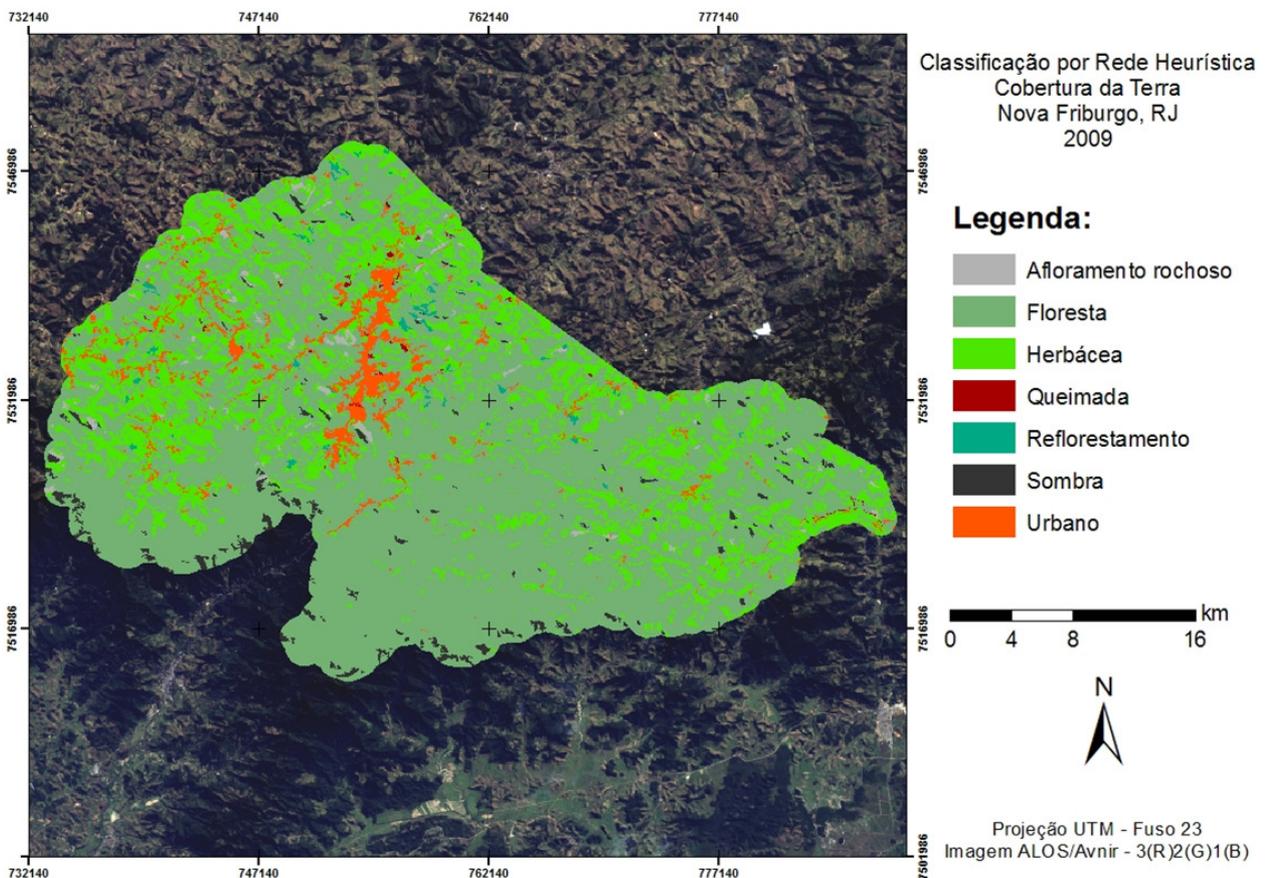


Fig. 6 - Classificação de cobertura da terra de Nova Friburgo por rede heurística.

ou seja, a exatidão do usuário foi de 100%. A classe queimada, por sua vez, apresentou baixo índice de exatidão do usuário, devido aos erros de classificação das amostras da classe urbana previamente relatados.

A tabela 4 refere-se à matriz de erros da classificação de cobertura da terra gerada por rede heurística. O índice de exatidão global alcançado foi de 91%, e o índice *Kappa* de 0,85, igualmente considerado de excelente qualidade de acordo com Landis e Koch (1977). Todas as classes ficaram com índice de exatidão do produtor acima de 50%, com exceção do afloramento rochoso, que apresentou uma alta confusão com a classe herbácea. Por outro lado, esta mesma classe não apresentou erros de comissão, ou seja, a exatidão do usuário foi de 100%. Em relação a esta taxa, todas as classes ficaram com valor acima de 70%. O teste de hipótese feito entre os índices *Kappa* das duas classificações mostrou, com nível de significância de 5%, que não há diferença entre suas acurácias. No entanto, na análise do *Kappa* condicional, verifica-se que há diferença de acurácia entre as duas classificações (tabela 5). Para a classificação gerada por mineração de dados, os valores do *Kappa* condicional (usuário) ficaram entre 0,34 e 1,0, enquanto que para a rede

heurística, os valores situam-se entre 0,65 e 1,0. Considerando-se este índice, a classificação heurística apresentou melhor desempenho nas classes afloramento rochoso e queimada. O desempenho das classes restantes não foi significativamente diferente, com nível de 5%, nas duas classificações, de acordo com os testes de hipótese aplicados.

Em relação ao *Kappa* condicional (produtor), tanto para a classificação por mineração quanto por rede heurística, os valores situaram-se na faixa entre 0,38 e 0,98. No entanto, na classificação por rede heurística, a classe afloramento rochoso apresentou valor inferior a 50%, ao passo que por mineração, a classe urbana apresentou o menor valor. Por outro lado, as classes afloramento rochoso e sombra apresentaram melhor desempenho na classificação por mineração, enquanto que na classificação heurística, as classes herbácea e urbana apresentaram os melhores resultados. O teste de hipótese demonstrou, com nível de significância de 5%, que o *Kappa* condicional (produtor) das classes floresta, queimada e reflorestamento não são significativamente diferentes.

Por fim, as diferenças entre a área das classes de cobertura da terra nas classificações testadas variaram aproximadamente, em módulo, de 15% a 560% (tabela 6). As classes floresta e herbácea obtiveram os menores valores na taxa de variação. Com relação à classe floresta, este fato é respaldado pelos respectivos índices *Kappa* condicional (usuário e produtor), que indicam os menores erros de omissão e comissão para esta classe nas duas classificações. A vegetação herbácea apresentou índices *Kappa* condicional semelhantes nas duas classificações, sendo que para o *Kappa* (usuário),

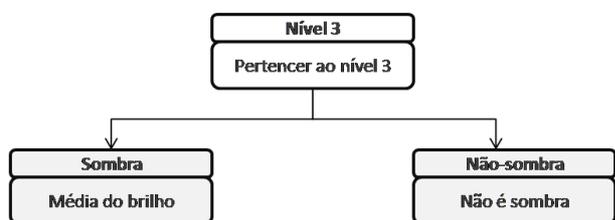


Fig. 7 - Rede semântica heurística do terceiro nível de classificação de cobertura da terra.

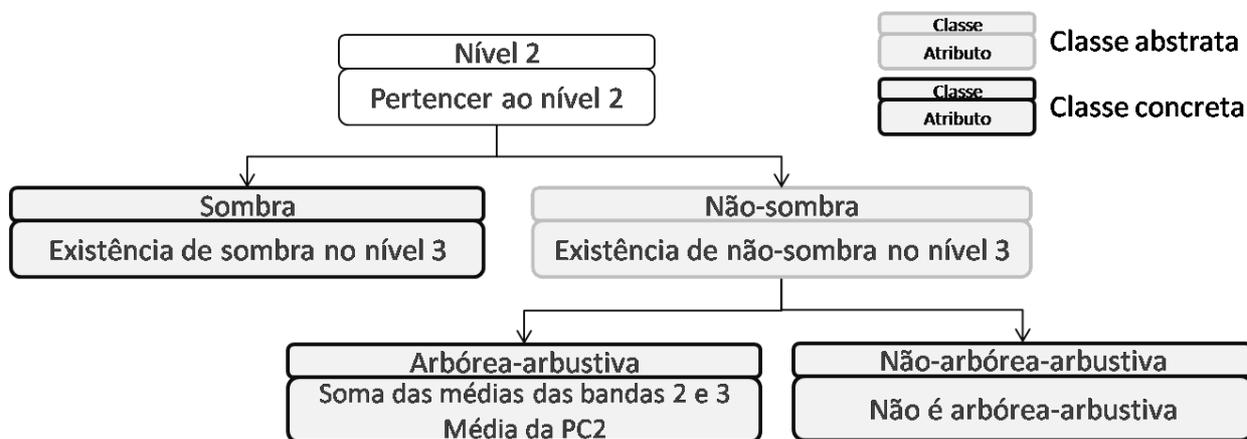


Fig. 8 - Rede semântica heurística do segundo nível de classificação de cobertura da terra.

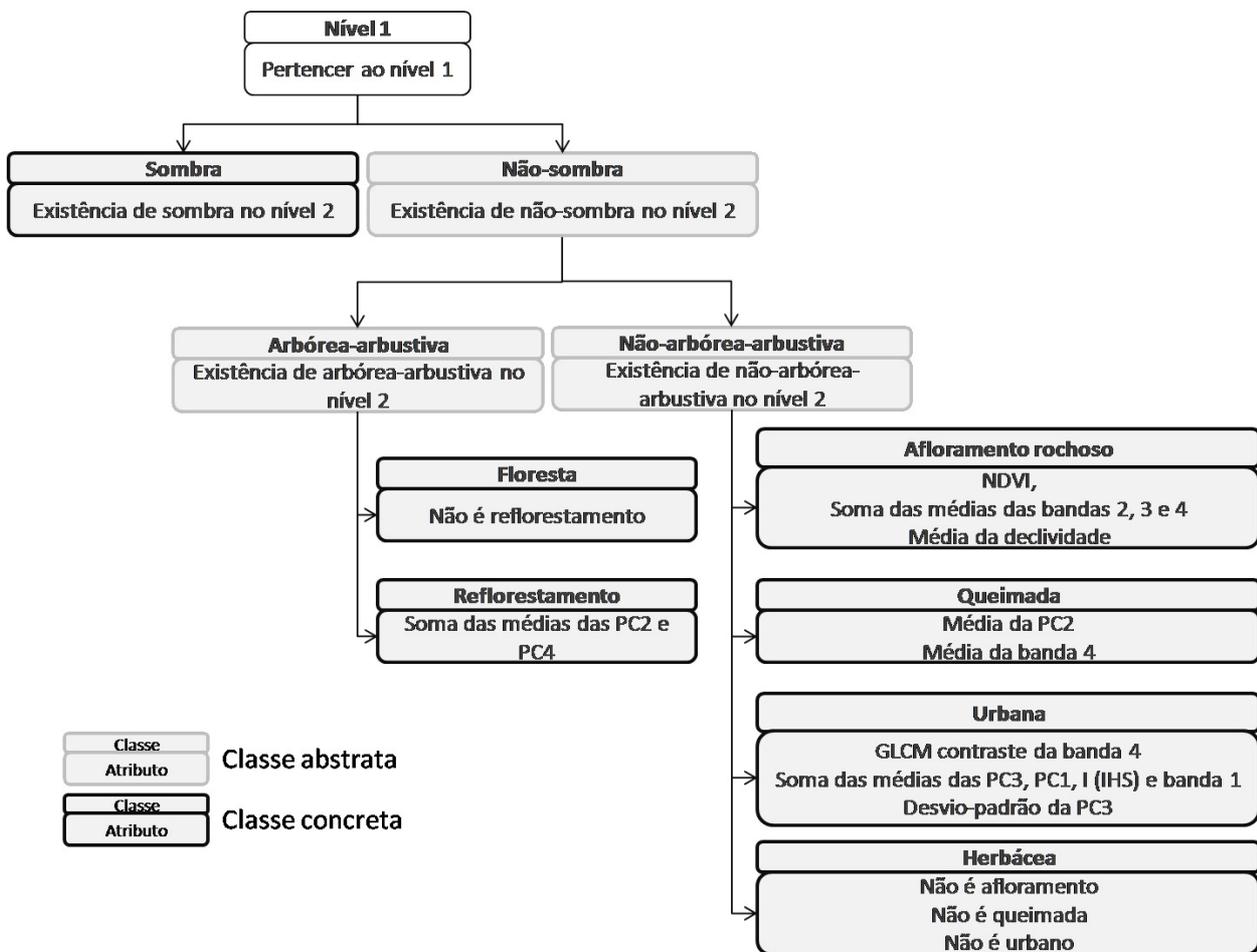


Fig. 9 - Rede semântica heurística do primeiro nível de classificação de cobertura da terra.

os índices não são significativamente diferentes. As maiores taxas de variação ocorreram nas classes queimada e afloramento rochoso, com diferenças de cerca de 348% e 560%, respectivamente. A primeira apresentou valor elevado de erro de comissão, confundindo-se com a classe urbana na classificação por mineração, explicando assim o aumento de área nesta classificação. Como a mineração separa a área urbana apenas pela textura, as áreas urbanas não rugosas apresentam mais semelhança com a queimada do que com as demais classes em vista dos descritores selecionados para discriminação da classe queimada (altos valores de CP1, baixos de NDVI e de declividade). Por sua vez, o afloramento rochoso apresentou elevado erro de omissão, confundindo-se principalmente com a vegetação herbácea na classificação por rede heurística, explicando assim a diminuição de área para esta classificação. No entanto, deve ser salientado que as maiores diferenças ocorreram nas classes com área menos expressiva, e as menores, em classes com áreas mais expressivas.

#### 4. CONCLUSÕES

Considerando o índice *Kappa* global, pode-se afirmar que as duas classificações avaliadas, por mineração e por rede heurística, apresentaram excelente qualidade quanto à acurácia da classificação. O teste de hipótese entre estes dois índices mostra que não há diferenças significantes entre eles. Os valores de *Kappa* alcançados neste trabalho foram semelhantes aos obtidos em outras classificações de cobertura da terra (REGO, 2003; FÖRSTER *et. al.*, 2010; KANJIR *et. al.*, 2010; KASS, 2010) utilizando OBIA com número de classes e resolução espacial semelhantes aos adotados neste trabalho.

Em relação às classes de cobertura da terra avaliadas pelo índice *Kappa* condicional, pode-se afirmar que a qualidade da classificação é de muito boa a excelente, pois todos os índices foram superiores a 0,6, com exceção das classes afloramento rochoso, queimada e urbana, que apresentaram desempenho aceitável em pelo menos um dos índices *Kappa* condicional avaliados. O

Tabela 3 - Matriz de erros da classificação gerada por mineração de dados.

Classificadas	Amostras de Referência								
	Classes	Afloramento	Floresta	Herbácea	Queimada	Refloresta- mento	Sombra	Urbana	Total classificadas
	Afloramento	25		9					34
	Floresta	1	420	4	1	2	3		431
	Herbácea	2	2	84		2		24	114
	Queimada	1			7			12	20
	Reflorestamento		1			17			18
	Sombra	1	2				42		45
Urbana							24	24	
Total coletadas	30	425	97	8	21	45	60	686	
Exatidão do Produtor	83%	99%	87%	88%	81%	93%	40%	90%	
Exatidão do Usuário	74%	97%	74%	35%	94%	93%	100%		

Tabela 4 - Matriz de erros da classificação gerada pela rede heurística.

Classificadas	Amostras de Referência								
	Classes	Afloramento	Floresta	Herbácea	Queimada	Refloresta- mento	Sombra	Urbana	Total classificadas
	Afloramento	12							12
	Floresta	2	422	2		6	7		439
	Herbácea	16	3	90	1	1	10	8	129
	Queimada				7				7
	Reflorestamento					14			14
	Sombra	1					28		29
Urbana		1	1				53	55	
Total coletadas	31	426	93	8	21	45	61	685	
Exatidão do Produtor	39%	99%	97%	88%	67%	62%	87%	91%	
Exatidão do Usuário	100%	96%	70%	100%	100%	97%	96%		

Tabela 5 - Índice Kappa condicional das classificações por rede heurística e por mineração de dados.

Classe	Kappa - Usuário		Kappa - Produtor	
	Rede heurística	Rede mineração	Rede heurística	Rede mineração
Afloramento rochoso	1,00	0,72	0,38	0,83
Floresta	0,90	0,90	0,98	0,97
Herbácea	0,65	0,69	0,96	0,80
Queimada	1,00	0,34	0,87	0,87
Reflorestamento	1,00	0,94	0,66	0,80
Sombra	0,96	0,93	0,61	0,93
Urbana	0,96	1,00	0,86	0,38

Tabela 6 - Área das classes de cobertura da terra.

Classe	Classificação heurística		Classificação mineração		Variação (%)
	Área (km <sup>2</sup> )	Área (%)	Área (km <sup>2</sup> )	Área (%)	
Floresta	761,1	69,1	648,0	61,7	14,9
Herbácea	272,6	24,8	330,7	31,5	21,3
Urbana	42,6	4,0	19,1	1,9	55,3
Sombra	16,9	1,5	28,7	2,8	69,7
Reflorestamento	5,6	0,5	15,8	1,4	182,2
Queimada	1,9	2,5	8,7	6,6	347,9
Afloramento rochoso	8,9	100,0	58,7	100,0	560,2

melhor desempenho coube à classe floresta, que obteve *Kappa* igual ou superior a 0,9 em todos os casos avaliados (usuário e produtor), indicando a excelente qualidade de classificação em ambos os métodos.

O desempenho das duas classificações foi semelhante para o reflorestamento e a floresta, considerando-se que os índices *Kappa* condicional para estas duas classes, tanto do produtor quanto do usuário, não são significativamente diferentes. Em relação às classes vegetação herbácea, urbana e sombra, o desempenho das duas classificações apresentou-se semelhante apenas quando avaliadas pelo *Kappa* condicional (usuário), ou seja, os erros de comissão não são significativamente diferentes. Já o desempenho em relação à classe queimada, de forma diversa, apresentou-se semelhante nas duas classificações apenas quando avaliadas pelo *Kappa* condicional (produtor), ou seja, os erros de omissão não são significativamente diferentes. Por sua vez, a qualidade da classificação do afloramento rochoso foi significativamente diferente tanto em erros de omissão quanto de comissão.

Em termos absolutos, a classificação por rede heurística apresentou o melhor desempenho geral, justificado pelos maiores valores de exatidão global e índice *Kappa*. Em termos relativos, o desempenho de ambas as classificações quanto ao *Kappa* condicional (produtor) mostrou-se equivalente, o que não se verificou para o *Kappa* condicional (usuário), em vista da comprovada superioridade da classificação por rede heurística.

O desempenho da rede semântica derivada de mineração de dados é função da qualidade das amostras de treinamento, que devem ser suficientemente representativas da variabilidade de cada classe e, ao mesmo tempo, capazes de detectarem um comportamento característico para cada classe. Da mesma forma, a habilidade do modelador na definição dos parâmetros ótimos para a geração da árvore de decisão é crucial para a qualidade dos resultados da classificação, a exemplo do número mínimo de amostras a serem atribuídas a cada classe, o que definirá o grau de generalização da árvore a ser produzida.

Ambos os métodos possuem etapas dispendiosas de tempo. Na abordagem heurística, grande esforço é demandado do intérprete na seleção de atributos e na construção e ajuste de curvas *fuzzy*. Na abordagem por mineração de

dados, observa-se um elevado consumo de tempo para a coleta e seleção de amostras, bem como para o refinamento das árvores por meio do ajuste de seus parâmetros internos.

Em termos operacionais, contudo, a abordagem por mineração ainda se mostra comparativamente mais vantajosa, por permitir a automatização dos procedimentos de seleção de atributos e definição das regras de decisão, e consequentemente, mostrar-se menos sujeita à subjetividade do intérprete.

## AGRADECIMENTOS

As autoras agradecem ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPQ), pela bolsa de pós-doutorado, e à Fundação de Auxílio à Pesquisa do Rio de Janeiro (FAPERJ), pelo financiamento deste trabalho de pesquisa.

## REFERÊNCIAS BIBLIOGRÁFICAS

ALMEIDA, C. M.; SOUZA, I. M. E.; ALVES, C. D.; PINHO, C. M. D.; FEITOSA, R. Q. Métodos Cognitivos de Classificação Aplicados a Imagens QuickBird para a Detecção de Áreas Residenciais Homogêneas. **Revista Brasileira de Cartografia**, v. 61, p. 1-12, 2009.

BARALDI, A.; PARMIGGIANI, F. An Investigation of the Textural Characteristics Associated with Gray Level Co-occurrence Matrix Statistical Parameters. **IEEE Transactions on Geoscience and Remote Sensing**, 33(2) p.293–304, 1995.

BLASCHKE, T.; LANG, S.; HAY, G.J. **Object-Based Image Analysis Spatial Concepts for Knowledge-Driven Remote Sensing Applications**. Berlin: Springer. 2008. 828p.

CASTILLA, G., HAY, G. J. Image objects and geographic objects. In: BLASCHKE, T.; LANG, S.; HAY, G.J. (Eds.) **Object-Based Image Analysis Spatial Concepts for Knowledge-Driven Remote Sensing Applications**. Berlin: Springer, p.91-110, 2008.

CIDE. Fundação Centro de Informações e Dados do Rio de Janeiro. **Índice de Qualidade de Municípios Verde II**. Rio de Janeiro: Secretaria de Estado de Planejamento, Desenvolvimento Economico e Turismo. 2003. 154p.

- CONGALTON, R.; GREEN, K. **Assessing the Accuracy of Remotely Sensed Data: Principles and Practices**. Boca Raton: CRC/Taylor & Francis. 2009. 183p.
- DANTAS, M.E. Geomorfologia do estado do Rio de Janeiro. In: Silva, L.C.; Cunha, H.V.S. **Geologia do Estado do Rio de Janeiro: texto explicativo do mapa geológico do Estado do Rio de Janeiro**. Brasília: CPRM. 2001. 63p.
- DEFINIENS. **Definiens Developer 7.04 User Guide**. 2007. 51p.
- ECOGNITION. Disponível em: <http://www.ecognition.com/>. Acesso em 20 de julho 2011.
- ENVI. **ENVI User's Guide - ENVI Version 4.7**. ITT Visual Information Solutions. 2009. 212p.
- FÖRSTER, M.; FRICK, A.; SCHUSTER, C.; KLEINSCHMIT B. Object-based change detection analysis for the monitoring of habitats in the framework of the natura 2000 directive with multi-temporal satellite data. **The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences**, Vol. XXXVIII-4/C7. 2010.
- HALL, M.; FRANK, E.; HOLMES, G; PFAHRINGER, B; REUTEMANN, P; WITTEN, I. The WEKA Data Mining Software: An Update. **SIGKDD Explorations**, Volume 11, Issue 1. 2009.
- HAN, J.; KAMBER, M. **Data Mining: Concepts and Techniques**. San Francisco: Morgan Kaufmann Publishers. 2006. 743p.
- HARALICK, R. M; SHANMUGAM, K; DINSTEIN, I. Textural Features for Image Classification. **IEEE Transactions on Systems, Man and Cybernetics**, 3, p. 610-620. 1973.
- KOHAVI, R.; QUINLAN, R. Decision Tree Discovery. In: Klossgen, W.; Zytkow, J. M. (Ed.). **Handbook of data mining and knowledge discovery**. Oxford: Oxford University, 2002, p. 282-288.
- IBGE. *Cidades @*. Disponível em: <http://www.ibge.gov.br/cidadesat>. Acesso em 09 de agosto de 2011a.
- IBGE. *Ortofotos 1:25.000 - Rio de Janeiro*. Disponível em: [ftp://geoftp.ibge.gov.br/mapas/ortofoto/Projeto\\_RJ25/TIF/](ftp://geoftp.ibge.gov.br/mapas/ortofoto/Projeto_RJ25/TIF/). Acesso em fevereiro de 2011b.
- ITO, R. H.; KIM, H. Y.; SALCEDO, W. J. **Classificação de Texturas Invariante a Rotação Usando Matriz de Co-ocorrência**. Disponível em: <http://www.lps.usp.br/~hae/Ito-Kim-Salcedo-Textura.pdf>. Acesso em 20 de julho de 2011.
- KASS, S.; NOTARNICOLA, C.; ZEBISCH, M. Object-based approach to map semi-natural areas in mountain region with high spatial resolution images. In: **Geographic Object-Based Image Analysis, 2010**, Ghent, Belgium. **Proceedings...**, 2010.
- KRESSLER, F.P.; STEINNOCHER, K. Image data and LIDAR – an ideal combination matched by object oriented analysis. In: **Geographic Object-Based Image Analysis, 1, 2006**, Salzburg University, Austria. **Proceedings...**, 2006.
- LANDIS, J.; KOCH, G. The measurement of observer agreement for categorical data. **Biometrics**, v. 33, p. 159 - 174, 1977.
- LANG, S. Object-based image analysis for remote sensing applications: modeling reality – dealing with complexity. In: BLASCHKE, T.; LANG, S.; HAY, G.J. (Eds.) **Object-Based Image Analysis Spatial Concepts for Knowledge-Driven Remote Sensing Applications**. Berlin: Springer. p. 3-28. 2008.
- MAVRANTZA, O.; ARGIALAS, D. An object-oriented image analysis approach for the identification of geologic lineaments in a sedimentary geotectonic environment. In: Blaschke, T.; Lang, S.; Hay, G.J. **Object-Based Image Analysis Spatial Concepts for Knowledge-Driven Remote Sensing Applications**. Berlin: Springer. 2008. p.383-398.
- MARPU, P. R. **Geographic Object-based Image Analysis**. Freiberg, 2009. 121f. Tese (Doutorado em Engenharia) - Faculty of Geosciences, Geo-Engineering and Mining Technische Universität Bergakademie Freiberg. 2009
- NAVULUR, K. **Multispectral image analysis using the object-oriented paradigm**. Boca Raton: Taylor & Francis Group, 2006. 165p.
- PINHO, C. M. D. **Análise orientada a objetos de imagens de satélite de alta resolução espacial aplicada à classificação de cobertura do solo no espaço intraurbano: o caso de São José dos Campos – SP**. 2005. 179 p. Dissertação.

(Mestrado em Sensoriamento Remoto) – INPE. São José dos Campos, 2005a.

PINHO, C.M.D.; Rennó, C.D.; Kux, H.J.H. Avaliação de técnicas de fusão aplicadas à imagem QuickBird. In: Simpósio Brasileiro de Sensoriamento Remoto, 12, 2005a, Goiânia. **Anais...** São José dos Campos: INPE, p. 4225-4232. 2005b.

QUINLAN, R. **C4.5: Programs for Machine Learning**. San Francisco: Morgan Kaufmann. 1993. 302 p.

REGO, L.F.G. **Automatic land-cover classification derived from high-resolution ikonos satellite image in the urban atlantic forest in Rio de Janeiro, Brasil by means of an objects-oriented approach**. 2003. 179p. Tese (Doutorado em Ciências Florestais). Albert-Ludwigs-Universität. Freiburg, 2003.

ROSSITER, D.G. **Technical Note: Statistical methods for accuracy assesment of classified**

**thematic maps**. International Institute for Geo-information Science & Earth Observation (ITC). 2004.

SILVA, M.P.S. **Mineração de padrões de mudança em imagens de Sensoriamento Remoto**. 2006. 128 p. Tese (Doutorado em Sensoriamento Remoto) – INPE. São José dos Campos, 2006.

SPRING. **Sistema de Processamento de Informações Georreferenciadas. SPRING: Tutorial de Geoprocessamento**. Disponível em: <http://www.dpi.inpe.br/spring/teoria/princo.htm>. Acesso em: 20 de julho 2011.

VALERIANO, M. M. Modelo digital de variáveis morfométricas com dados SRTM para o território nacional: o projeto TOPODATA. In: Simpósio Brasileiro de Sensoriamento Remoto, 12, 2005, Goiânia, GO. **Anais ...**, 2005. p. 1-8.