

LUÍS EDUARDO RIBEIRO DE MENDONÇA

**UTILIZAÇÃO DE ALGORITMOS DE CLASSIFICAÇÃO EM UMA IMAGEM
IKONOS PARA O MAPEAMENTO DE FEIÇÕES TERRESTRES:
UM ESTUDO DE CASO**

Dissertação apresentada à
Universidade Federal de Viçosa,
como parte das exigências do
Programa de Pós-Graduação em
Ciência Florestal, para obtenção do
título de *Magister Scientiae*.

**VIÇOSA
MINAS GERAIS – BRASIL
2006**

LUÍS EDUARDO RIBEIRO DE MENDONÇA

**UTILIZAÇÃO DE ALGORITMOS DE CLASSIFICAÇÃO EM UMA IMAGEM
IKONOS PARA O MAPEAMENTO DE FEIÇÕES TERRESTRES:
UM ESTUDO DE CASO**

Dissertação apresentada à
Universidade Federal de Viçosa,
como parte das exigências do
Programa de Pós-Graduação em
Ciência Florestal, para obtenção do
título de *Magister Scientiae*.

APROVADA: 31 de agosto de 2006.

**Prof. José Marinaldo Gleriani
(Co-Orientador)**

**Prof. Guido Assunção Ribeiro
(Co-Orientador)**

Prof. Herly Carlos Teixeira Dias

Prof. Carlos Antônio Oliveira Vieira

**Prof. Vicente Paulo Soares
(Orientador)**

Esta tese é dedicada aos meus pais, Ary e Nivalda, pela coragem com que enfrentam a vida, sempre muito difícil, mas igualmente feliz.

Aos meus irmãos, Marco Aurélio, Ana Flávia e Lucas, que mostraram ao longo da vida que a raça é hereditária, pois nunca deixaram a "peteca cair".

As minhas irmãzinhas, Kethlen e Kelly, as menininhas que vieram para mim...

Ao Sr. Jannotti e Sra. Juraci, que me acolheram em sua casa com muito carinho e compreensão.

A minha Querida e Amada Juliana, sempre companheira em todos os momentos de nossas vidas.

AGRADECIMENTOS

Ao Departamento de Engenharia Florestal da Universidade Federal de Viçosa, pela oportunidade ímpar de realizar este trabalho e por disponibilizar o programa computacional Idrisi.

Ao Professor Vicente Paulo Soares, não só pelas idéias no decorrer dos estudos, mas, principalmente, pela confiança e serenidade com que conduziu todo o processo.

Ao Professor José Marinaldo Gleriani, por estar sempre disposto a ajudar e contribuir com sua experiência e conhecimento.

Ao Professor Guido Assunção Ribeiro, por colaborar e enriquecer este trabalho com suas observações peculiares.

Aos Professores Herly Carlos Teixeira Dias e Carlos Antônio Oliveira Vieira, por participarem da banca de defesa, fazendo do questionamento uma oportunidade para melhorar sempre, aprimorando assim as informações discorridas neste trabalho.

Ao Professor Elias Silva, pelas sábias palavras em um momento de decisão importante na minha vida.

Aos meus amigos do D.C. pelos momentos de descontração, importantes durante minha morada em Viçosa.

A todos os amigos que fiz em Viçosa e que vou guardar pra sempre.

BIOGRAFIA

Luís Eduardo Ribeiro de Mendonça nasceu em 22 de março de 1974, em Belo Horizonte, Minas Gerais. Filho de Ary Ribeiro de Mendonça e Nivalda Maria Ribeiro de Mendonça, graduou-se Engenheiro Agrimensor pela FEAMIG em dezembro de 2001. No ano seguinte iniciou e concluiu uma pós-graduação em Engenharia de Segurança do Trabalho, na mesma instituição da graduação.

Em agosto de 2003 mudou-se para Viçosa com o propósito de ingressar no programa de pós-graduação da Universidade Federal de Viçosa. Tal fato foi ocorrer somente em agosto de 2004, quando iniciou o curso de Mestrado no Departamento de Engenharia Florestal, sob a orientação do Professor Vicente Paulo Soares.

Em um final movimentado e emocionante, já trabalhando na Companhia Vale do Rio Doce e de casamento marcado para o dia 02 de setembro de 2006, defendeu a tese em 31 de agosto de 2006.

LISTA DE QUADROS

	Pág.
Quadro 1 - Características do sensor do satélite IKONOS II.....	05
Quadro 2 - Representação genérica uma matriz de erros.....	17
Quadro 3 - Intervalo de aceitação dos resultados do índice Kappa.....	18
Quadro 4 - Relação de classes temáticas utilizadas nas classificações.....	24
Quadro 5 - Dados utilizados para a segmentação da imagem.....	26
Quadro 6 - Quantidades amostradas e as perdas por sobreposição.....	29
Quadro 7 - Valores dos índices de Kappa e suas respectivas variâncias, considerando os diferentes tamanhos de amostras.....	42

LISTA DE FIGURAS

Figura 1 - Configuração das Redes Neurais Artificiais (RNAs).....	11
Figura 2 - Localização da área de estudo.....	23
Figura 3 - Imagem da área de estudo, composição R (3), G (2) e B (1), com as áreas de coleta das amostras de treinamento.....	25
Figura 4 - Configurações do algoritmo de Redes Neurais no programa computacional Idrisi – <i>The Kilimanjaro Edition</i>	26
Figura 5 - Adensamento dos pontos amostrais para os testes de acurácia: (a) 250 (b) 500 (c) 750 e (d) 1.000 amostras.....	28
Figura 6 - Mapa Temático gerado pelo algoritmo de Máxima Verossimilhança.....	31
Figura 7 - Mapa Temático gerado pelo algoritmo de Distância Euclidiana Mínima.....	31
Figura 8 - Mapa Temático gerado pelo algoritmo de Redes Neurais.....	32
Figura 9 - Mapa Temático gerado pelo algoritmo Bhattacharya.....	32
Figura 10 - Mapa Temático gerado pelo algoritmo Isoseg.....	33
Figura 11 - Mapa Temático gerado pelo algoritmo Histograma.....	33
Figura 12 - Erros de comissão e omissão para as feições classificadas pelo algoritmo da Máxima Verossimilhança.....	36
Figura 13 - Erros de comissão e omissão para as feições classificadas pelo algoritmo da Distância Euclidiana Mínima.....	37
Figura 14 - Erros de comissão e omissão para as feições classificadas pelo algoritmo de Redes Neurais.....	38
Figura 15 - Erros de comissão e omissão para as feições classificadas pelo algoritmo Bhattacharya.....	39
Figura 16 - Erros de comissão e omissão para as feições classificadas pelo algoritmo Isoseg.....	39
Figura 17 - Erros de comissão e omissão para as feições classificadas pelo algoritmo do Histograma.....	40
Figura 18 - Valores calculados de Kappa para cada quantidade amostral....	41

SUMÁRIO

RESUMO	ix
ABSTRACT	x
1 – INTRODUÇÃO	1
1.1 – OBJETIVOS	3
2 – REVISÃO DE LITERATURA.....	4
2.1 – Satélite IKONOS	4
2.2 – Classificação de Imagens Digitais	5
2.2.1 – Classificação Supervisionada	5
2.2.2 – Classificação Não Supervisionada.....	7
2.2.3 – Algoritmos de Classificação por Pixel	7
2.2.3.1 – Algoritmo da Máxima Verossimilhança	8
2.2.3.2 – Algoritmo da Distância Euclidiana Mínima	8
2.2.3.3 – Algoritmo de Redes Neurais	9
2.2.4 – Algoritmos de Classificação Por Regiões	12
2.2.4.1 – Algoritmo Bhattacharya.....	13
2.2.4.2 – Algoritmo Isepeg	14
2.2.4.3 – Algoritmo do Histograma.....	15
2.3 – Avaliação das Imagens Classificadas	15
2.3.1 – Índice Kappa	17
2.4 – Aplicação de Imagens de Alta Resolução.....	19
3 – MATERIAL E MÉTODOS.....	22
3.1 – Localização e Caracterização da Área de Estudo	22
3.2 – Equipamentos e Sistemas Computacionais.....	22
3.3 – Dados.....	22
3.4. Procedimentos Metodológicos	23
3.4.1. Classificação da Imagem IKONOS	23
Item.....	24
Descrição	24
3.4.2 – Quantidade de Amostras Testes Para Avaliação das Imagens Classificadas.....	27
3.4.3. As comparações entre as quantidades de amostras e índices Kappa	29

4 – RESULTADOS E DISCUSSÃO	30
4.1 – Avaliação Qualitativa das Imagens Classificadas	30
4.2 – Avaliação Quantitativa da classificação	35
4.2 – A significância entre as diferentes quantidades amostrais	41
5 – CONCLUSÃO	44
6 – RECOMENDAÇÕES	45
REFERÊNCIAS BIBLIOGRÁFICAS.....	46
APÊNDICE I.....	49
APÊNDICE II.....	62

RESUMO

MENDONÇA, Luís Eduardo Ribeiro de, M.Sc., Universidade Federal de Viçosa, agosto de 2006. **Utilização de algoritmos de classificação em uma imagem IKONOS para o mapeamento de feições terrestres: um estudo de caso.** Orientador: Vicente Paulo Soares. Co-Orientadores: José Marinaldo Gleriani e Guido Assunção Ribeiro.

O presente trabalho consistiu em classificar parte de uma imagem IKONOS II, localizada no município de Serra do Salitre, estado de Minas Gerais, utilizando seis algoritmos de classificação, sendo três por *pixels* e três por regiões, visando o mapeamento da cobertura terrestre. Os algoritmos foram: Máxima Verossimilhança, Distância Euclidiana Mínima, Redes Neurais, Bhattacharya, Isoseg e Histograma. O cálculo das acurácias de classificação, pelo índice Kappa, foi feito com 250, 500, 750 e 1000 *pixels*, para verificar se haveria perda de exatidão ao se utilizar um menor número de amostras. Para essa comprovação, verificou-se se existiam diferenças estatísticas significativas entre os índices Kappa encontrados após as classificações. Os resultados indicaram que as classificações geradas pelos algoritmos que trabalham com uma imagem segmentada geraram imagens temáticas com qualidades visuais muito superiores aos que não utilizam este procedimento. Além do que, não houve diferença estatística significativa, a 95% de probabilidade, entre os índices Kappa, ao se variar o tamanho das amostras de testes. Houve diferença estatística significativa, a 95% de probabilidade, ao se comparar os índices Kappa gerados pelas classificações por *pixel* e por região, o que não ocorreu ao se comparar os três algoritmos por *pixel*, entre si, e os três algoritmos por regiões, também entre si. Os algoritmos de classificação por *pixels* apresentaram valores de índice Kappa considerados razoáveis enquanto que os algoritmos por regiões apresentaram resultados excelentes, com o Isoseg obtendo os melhores índices. O aumento do tamanho das amostras de testes para avaliação das imagens classificadas não demonstrou, estatisticamente, melhorias no cálculo do índice de exatidão kappa.

ABSTRACT

MENDONÇA, Luís Eduardo Ribeiro de, M.Sc., Universidade Federal de Viçosa, August 2006. **The use of classification algorithms in an IKONOS image for the mapping of ground features: a study case.** Adviser: Vicente Paulo Soares. Co-Advisers: José Marinaldo Gleriani and Guido Assunção Ribeiro.

The present work consisted in classifying part of an IKONOS II image, from the municipality of Serra do Salitre, Minas Gerais State, Brazil, using six classification algorithms, three of them by pixels and three by regions, to map the ground cover. The algorithms were: Maximum Likelyhood, Minimum Euclidian Distance, Neural Nets, Bhattacharya, Iseog and Hystogram. The calculations of the accuracy classification, made by the Kappa index, were performed with 250, 500, 750 and 1,000 pixels, to check any accuracy loss if smaller number of samples were used. For this check up it was verified if there were significant statistical differences among the kappa indexes found after the classifications or not. The results indicated that the classifications generated by the algorithms that work with a segmented image generated thematic images with a much higher visual quality than those that did not use this method. Besides, there was not a significant statistical difference, at 95% probability, among the kappa indexes, when the size of the test samples were varied. There was a significant statistical difference, at 95% probability, when comparing the kappa indexes generated by the classifications by pixel and by region. This did not occur when comparing the three algorithms by pixels with one another, and the three algorithms by regions also between each other. The classification algorithms by pixels showed values of kappa indexes considered fairly good while the algorithms by regions showed excellent results, with the Iseog obtaining the best indexes. The increase in the size of the samples of the tests for the evaluation of the classified images did not statistically show improvements in the kappa accuracy index calculation.

1 – INTRODUÇÃO

Imagens geradas por sensores remotos a bordo de satélites evoluíram muito desde o início da década de 70, quando foi lançado o primeiro, para observação da Terra. A coleta de dados separados em faixas do espectro eletromagnético e a possibilidade de se obter imagens periódicas, características dos sistemas orbitais, tornam o sensoriamento remoto um conjunto de técnicas que auxiliam no gerenciamento de recursos naturais.

Nas últimas décadas, têm sido lançadas no mercado uma enorme variedade de meios e aplicativos para a execução de mapeamentos da superfície terrestre, baseados nas técnicas do sensoriamento remoto e do geoprocessamento. A maneira como são integrados os produtos das diversas plataformas orbitais disponíveis e dos aplicativos existentes são fatores diretamente relacionados à qualidade dos resultados obtidos (COUTINHO *et al.*, 1998).

O controle territorial sempre foi motivo de muitas descobertas ao longo da história, objetivando desde o aproveitamento de áreas para cultivo até estratégias de defesa. Técnicas foram desenvolvidas na intenção de se conhecer o espaço onde as civilizações se desenvolveram ou iriam se desenvolver.

O sensoriamento remoto é uma importante ferramenta ligada ao levantamento de recursos naturais e ao monitoramento do meio ambiente, o que contribui para beneficiar o desenvolvimento econômico e social de uma população. No Brasil, um país de proporções continentais, essas técnicas podem e têm sido utilizadas neste sentido (NOVO, 1992).

Imagens geradas pelos sensores remotos, que caracterizam a superfície terrestre, são cada vez mais utilizadas para a elaboração de diferentes tipos de produtos. Os dados captados pelos sensores remotos são interpretados e transformados em informações, apresentados geralmente em forma de mapas temáticos (FLORENZANO, 2002).

As radiações refletidas e emitidas da superfície terrestres são armazenadas nos sensores e transformadas em informações sobre essa

área imageada. Essas informações representam a interação da energia eletromagnética emitida, na maioria das vezes pelo sol, e os objetos sensorizados. Com base em técnicas de processamento digital de imagens é possível inferir algo sobre este objeto e, devido às suas características, eventualmente discriminá-los.

De acordo com LILLESAND *et al.* (2004), o processamento digital de imagens é uma equação ou conjunto de equações aplicadas aos dados captados pelos sensores, armazenando os resultados da computação para cada *pixel*. O resultado forma uma nova imagem digital que pode ser exibida ou armazenada em formato pictórico, para ser manipulado posteriormente em programas computacionais.

As técnicas de classificação digital são baseadas em um processo de decisão, cujo resultado visa estabelecer para cada ponto da imagem (*pixel*) uma determinada classe temática. Tais procedimentos tornam o processo de mapeamento, com base no reconhecimento de características da superfície terrestre, menos subjetivo e com maior potencial de repetição em situações subseqüentes (NOVO, 1992).

No formato digital, a imagem pode ser interpretada utilizando os programas computacionais, no intuito de aproximar o máximo possível a representação digital da cena à realidade. Neste sentido, o resultado do processo de classificação da imagem deve ser testado, a fim de verificar sua acurácia, indicando seu nível de confiabilidade.

As falhas mais comuns de erros durante o mapeamento automático são resultantes da falha na identificação das classes temáticas existentes na cena, além da delimitação dos diferentes alvos. Tais dificuldades estão ligadas diretamente à escala de trabalho, à resolução espacial do sensor e ao desenho final do mapa, considerando-se inclusive a qualidade do papel utilizado na impressão dos mapas temáticos (MOREIRA, 2005).

O mapa gerado com a classificação é submetido a um processo de validação mediante a aplicação de índices de exatidão, que comprovam a qualidade do mesmo. LILLESAND *et al.* (2004) afirmam que, infelizmente, a capacidade de produzir mapas, com base nas técnicas de classificação digital da cobertura da Terra, não dispensa as técnicas para quantificar a significância de sua precisão.

Para se avaliar uma classificação, o mapa temático é comparado com algum material que comprove a veracidade de suas informações. Esse material pode ser uma foto aérea, mapa de inventário florestal, visita a campo, ou qualquer outro produto que auxilie no processo de validação.

Os sensores de alta resolução captam variações significativas nas feições terrestres, que causam confusões durante o processo de classificação. Os avanços tecnológicos devem acontecer em todas as fases do sensoriamento remoto, desde a captação dos dados até a apresentação final do produto, cujas informações servirão para uma tomada de decisão importante e que requer informações confiáveis para obtenção de resultados positivos.

1.1 – OBJETIVOS

Geral:

Testar seis algoritmos de classificação, sendo três por *pixels* e três por regiões, visando o mapeamento da cobertura terrestre em uma imagem digital do satélite IKONOS II, de alta resolução espacial.

Específicos:

- Calcular as acurácias dos mapas temáticos gerados pelas classificações, com a utilização do índice Kappa, a partir do uso de diferentes tamanhos de amostras de teste.
- Verificar se existe diferença estatística significativa entre os índices Kappa, resultantes das classificações automáticas.

2 – REVISÃO DE LITERATURA

2.1 – Satélite IKONOS

O programa IKONOS (do grego *icon*, imagem) prevê o lançamento de uma série de satélites. O primeiro deles, o IKONOS I, por problemas técnicos, não chegou a ser colocado em órbita (MOREIRA, 2005). O satélite IKONOS II foi lançado no dia 24 de setembro de 1999, com a previsão de uma vida operacional de aproximadamente 8 anos e meio. Possui a bordo de sua plataforma sensores capazes de gerar imagens pancromáticas e multispectrais, de 1 e 4 metros de resolução espacial, respectivamente, havendo a possibilidade de combiná-las, produzindo imagens coloridas com resolução de 1 metro (SPACE IMAGING, 2006).

A alta resolução espacial muda até o modo de se utilizarem as imagens geradas pelos satélites. Se anteriormente um *pixel* continha vários objetos, agora um objeto é composto de vários *pixels*, aumentando o nível de detalhamento da cena. Uma das características do satélite IKONOS é que, além de efetuar visadas no sentido de sua órbita, captam imagens no sentido perpendicular a esta, o que aumenta a frequência de revisita e a possibilidade de aquisição de pares estereoscópicos, utilizados na restituição de mapas (MOREIRA, 2005).

As principais características técnicas do satélite IKONOS II e seus produtos estão resumidos no Quadro 1.

Quadro 1 - Características do satélite IKONOS II.

Altitude	680 km			
Inclinação	98,1°			
Velocidade	7 km / s			
Sentido da órbita	Descendente			
Duração da órbita	98 minutos			
Tipo de órbita	Sol-síncrona			
Resolução espacial	Pancromática: 1m / Multiespectral: 4m			
Bandas espectrais	Pancromática 0,45 – 0,90 μm Azul 0,45 – 0,52 μm Verde 0,52 – 0,60 μm Vermelho 0,63 – 0,69 μm Infravermelho próximo 0,76 – 0,90 μm			
Imageamento	13 km na vertical (cenas de 13 km x 13 km)			
Capacidade de aquisição de imagens	Faixas de 11 km x 100 km até 11 km x 1000 km Mosaicos de até 12.000 km ² 20.000 km ² de área imageada numa passagem			
Frequência de revisita	2,9 dias no modo Pancromático 1,5 dia no modo Multiespectral <i>Obs.: Esses valores valem para latitude de +/- 40°. A frequência de revisita para latitudes maiores será menor, e maior para as latitudes perto do Equador.</i>			
Banda	Faixa (μm)	Reg. espectral	Resolução (m x m)	Larg de faixa (km)
MS1	0,45 a 0,52	Azul	4	11
MS2	0,52 a 0,60	Verde	4	11
MS3	0,63 a 0,69	Vermelho	4	11
MS4	0,76 a 0,90	Infravermelho	4	11
Pan	0,45 a 0,90	Visível e IV	1	11

MS = Multiespectral

FONTE: Moreira (2005)

2.2 – Classificação de Imagens Digitais

2.2.1 – Classificação Supervisionada

Nas classificações supervisionadas, o analista está em constante interação com o sistema de análise de imagens digitais. Além disto, ele dispõe de informações sobre a cena, que servem de treinamento para o sistema. Essas informações são conhecidas pelo nome de amostras de treinamento e devem ser representativas das classes que deverão ser mapeadas automaticamente (NOVO, 1992).

LILLESAND *et al.* (2004) resumem a classificação supervisionada em três passos básicos, descritos a seguir:

1. Fase de treinamento: O analista identifica áreas de treinamento representativas e desenvolve uma descrição numérica (estatísticas) dos atributos espectrais de cada tipo de cobertura terrestre, na cena de interesse.
2. Fase de classificação: Cada pixel na imagem selecionada é categorizado dentro de uma classe de cobertura terrestre, de acordo com parâmetros estatísticos.
3. Fase de apresentação dos dados: Esta fase se resume na forma como serão apresentados os resultados da classificação.

Outra etapa, não menos importante, é a de qualificar as informações contidas nos mapas derivados do sensoriamento remoto. De acordo com CONGALTON e GREEN (1998), uma das importâncias desta etapa é o aumento da qualidade das informações, por identificar e corrigir as fontes de erros. A comparação entre várias técnicas, algoritmos, analistas e interpretes podem testar e identificar os melhores procedimentos a serem utilizados, além de quantificar a qualidade dessas informações, para se definir até que ponto podem ser confiáveis nas tomadas de decisão.

A fase de treinamento consiste em fornecer ao sistema um conjunto de *pixels*, representativos de cada alvo de ocupação do solo, na imagem a ser classificada. Nesta fase, as classes de uso do solo que apresentarem diferentes respostas espectrais devem ser divididas em tantas subclasses quantas forem as variações espectrais observadas na área de estudo (MOREIRA, 2005).

De acordo com LILLESAND *et al.* (2004), a qualidade do processo de treinamento é fator importante para o sucesso do estágio de classificação. O objetivo do processo de treinamento é reunir um conjunto de estatísticas que descrevem os padrões de respostas espectrais para cada tipo de cobertura a ser classificada na imagem.

Um analista pode selecionar locais de treinamento dentro da imagem, que são classes de interesse, representativas da cobertura terrestre e cujos esquemas serão adotados na classificação (JENSEN, 1996). O grande

problema nesta fase da classificação, para classificadores paramétricos, é obter amostras de *pixels* puros, representativos de cada classe espectral. A resposta espectral do alvo contido numa área do terreno nem sempre é homogênea e exclusivamente daquele alvo. Além disso, um mesmo tipo de alvo, principalmente a vegetação, pode apresentar mais de um comportamento espectral em função de ataque de pragas e doenças, épocas de plantio ou manchas de solo com deficiência de nutrientes, provocando mudanças na quantidade de energia refletida. Outro problema bastante comum, no que diz respeito a áreas vegetadas, é a existência de tipos diferentes, porém com comportamentos espectrais semelhantes (CAMPBELL, 1987).

2.2.2 – Classificação Não Supervisionada

No caso das classificações não supervisionadas, o analista tem pouco controle sobre a separação entre as classes, pois não tem conhecimento prévio da cena imageada e, além disto, ele não precisa preocupar-se com a homogeneidade das amostras (NOVO, 1992).

Classificações não supervisionadas não utilizam dados de treinamento como base para a classificação, sendo que os algoritmos examinam os *pixels* desconhecidos na imagem, associando-os à determinadas classes, baseando-se nos agrupamentos naturais presentes nos valores numéricos da imagem. A premissa básica é que valores dentro de um determinado tipo de cobertura devem estar próximos num mesmo espaço de medida, considerando que dados em diferentes classes devem estar comparativamente bem separados (LILLESAND *et al.*, 2004).

2.2.3 – Algoritmos de Classificação por Pixel

Na classificação *pixel a pixel*, após criar o conjunto de treinamento (características espectrais de cada uma das classes de ocupação do solo), o algoritmo classificador procura associar determinado *pixel* da imagem a uma das classes contidas no pacote de treinamento. Nesta situação, considera o *pixel* como uma entidade isolada e independente dos *pixels* que compõem o restante da cena (MOREIRA, 2005).

2.2.3.1 – Algoritmo da Máxima Verossimilhança

O algoritmo da máxima verossimilhança avalia a variância e a covariância das categorias de padrões de resposta espectral quando classifica um *pixel* desconhecido. Para isto, assume-se que a distribuição da nuvem de pontos que forma a categoria dos dados de treinamento é Gaussiana (distribuição normal). Assumida essa suposição, a distribuição de um padrão de resposta espectral da categoria pode ser completamente descrita por um vetor de média e a matriz de covariância. Com estes parâmetros é possível calcular a probabilidade estatística de um dado valor de *pixel* pertencer a uma classe particular de cobertura da terra (LILLESAND *et al.*, 2004).

O classificador por máxima verossimilhança define o diagrama de dispersão das classes em torno de uma média e suas distribuições de probabilidade, considerando a distribuição normal (MOREIRA, 2005).

A função de densidade de probabilidade é exibida na Equação 1:

$$P_c = [-0,5 \log_e(Det(Vc))] - [0,5(X - Mc)^T (Vc)^{-1} (X - Mc)] \quad (1)$$

Em que:

X = vetor de medidas dos *pixels* desconhecidos;

P_c = probabilidade de o vetor X ser assinalado na classe c ;

V_c = matriz de covariância da classe c contemplando todas as bandas (K, \dots, L);

$Det(V_c)$ = determinante da matriz de covariância V_c ;

Mc = vetor das médias para cada classe c ; e

T = matriz transposta.

2.2.3.2 – Algoritmo da Distância Euclidiana Mínima

Este é o algoritmo mais simples utilizado nas classificações de imagens digitais, necessitando de menos recursos computacionais, pois inclui somente três tarefas simples. Primeiro, a média dos valores nominais de uma classe, nos dados de treinamento, é calculada para todas as bandas espectrais, denominado vetor de médias; segundo, os *pixels* são

classificados em um conjunto de dados inteiros que são divididos em classes, cujos vetores de médias são mais próximos; terceiro, um limite para as classes de dados é localizado ao redor dos vetores tal que, se um *pixel* cair fora deste limite, ele será classificado como desconhecido (CURRAN, 1992).

A estratégia do algoritmo da Distância Euclidiana Mínima é matematicamente simples e computacionalmente eficiente, mas que possui certas limitações. Este algoritmo é insensível para diferenças nos graus de variância dos dados de resposta espectral. Por causa destes problemas, este classificador não é muito utilizado em aplicações onde as classes espectrais estão próximas umas das outras no espaço e possuem alta variância (LILLESAND *et al.*, 2004).

Para calcular a Distância Euclidiana Mínima de um ponto qualquer para uma determinada classe, de acordo com JENSEN (1996), utiliza-se a Equação 2.

$$D_{PC} = \sqrt{(VN_{ijk} - \mu_{ck})^2 + (VN_{ijl} - \mu_{cl})^2} \quad (2)$$

Em que:

D_{PC} = distância Euclidiana do *pixel* P até o centróide da classe C ;

VN_{ijk} e VN_{ijl} = valores numéricos dos *pixels* desconhecidos na linha i , coluna j , bandas k e l .

μ_{ck} e μ_{cl} = vetor de média da classe C na banda k e da classe C na banda l .

2.2.3.3 – Algoritmo de Redes Neurais

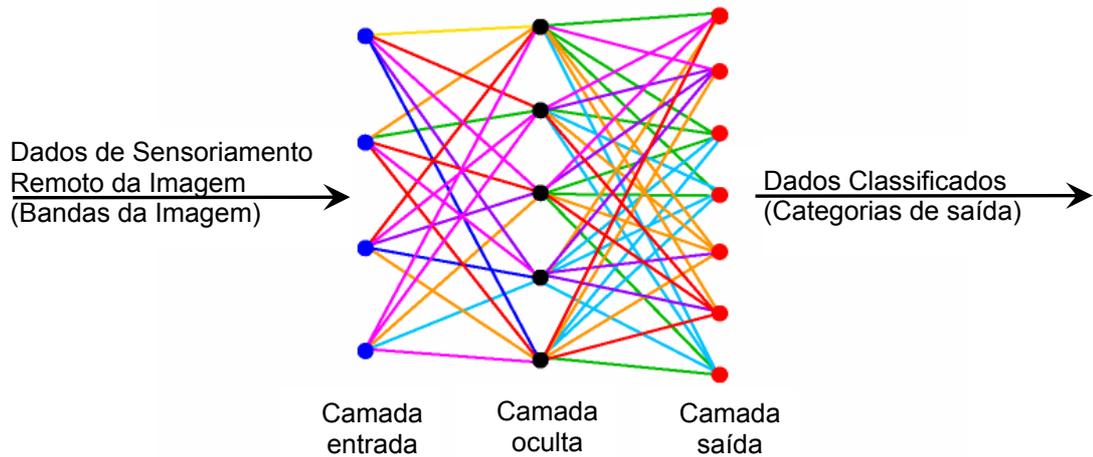
As Redes Neurais Artificiais (RNA's) são sistemas paralelos distribuídos, compostos por unidades de processamento simples, também chamadas de neurônios, que calculam determinadas funções matemáticas, normalmente não-lineares. São unidades dispostas em uma ou mais camadas, interligadas por um grande número de conexões, geralmente unidirecionais. Na maioria dos modelos estas conexões estão associadas a pesos, que configuram este modelo ponderando a entrada recebida por cada neurônio da rede. As redes neurais funcionam inspiradas em uma

concepção da natureza, que é a estrutura física do cérebro humano (BRAGA *et al.*, 2000).

As características não paramétricas das Redes Neurais fazem com que não seja necessária nenhuma suposição sobre a distribuição estatística dos dados. O desempenho da RNA se correlaciona mais com a significância das amostras de treinamento, e não com a suficiência de suposições relativas à distribuição estatística dos dados coletados. Durante a fase de treinamento as redes neurais aprendem sobre as regularidades presentes nas amostras e, com base nestas regularidades, constrói regras que podem ser estendidas aos dados desconhecidos. Além disso, o usuário deve determinar a arquitetura desta rede e também definir parâmetros para essas regras de aprendizado, que afetarão o tempo de treinamento, desempenho e a regra de convergência da RNA (MATHER e TSO, 2001).

Na fase de treinamento não se procura obter regras como na abordagem simbólica da Inteligência Artificial, mas sim determinar a intensidade de conexões entre neurônios. Um conjunto de procedimentos bem definidos para adaptar os parâmetros de uma RNA para que a mesma possa aprender uma determinada função é chamado de algoritmo de aprendizado (BRAGA *et al.*, 2000).

A Figura 1 exibe a configuração de uma Rede Neural com multicamadas, estando os algoritmos interpretados na camada oculta. Segundo RUMELHART *et al.* (1986) in MATHER e TSO (2001), o algoritmo de treinamento *Error Backpropagation* é o mais utilizado em modelos de Redes Neurais.



Fonte: IDRISI - *The Kilimanjaro Edition*

Figura 1 - Configuração das Redes Neurais Artificiais (RNAs).

O algoritmo *error backpropagation* é aplicado de forma supervisionada e, segundo VIEIRA (2000), minimiza uma função de erro interagindo os dados de saída da rede com conjuntos de saídas desejadas, treinadas com base em uma seleção de dados amostrais. O processo continua até o erro convergir para um nível baixo e aceitável. Convencionalmente, a função de erro é dada pela Equação 3.

$$E = \frac{1}{2} \sum_i (T_i - O_i)^2 \quad (3)$$

Em que:

T_i = vetor de saídas desejadas para o conjunto de treinamento i ;

O_i = vetor de saída da rede para o conjunto de treinamento dado.

VIEIRA (2000) indica três passos básicos ao treinar uma rede neural *feed-forward*, usando treinamento supervisionado, conforme abaixo:

- A rede inicia com a alocação aleatória de pesos entre os neurônios, que devem começar com valores aleatórios pequenos, a fim de evitar o problema de saturá-la com valores grandes ou a dificuldade no treinamento pela presença de valores de pesos homogêneos.
- Um padrão de entrada é apresentado à rede, sendo transmitido através da mesma até que a ativação alcance a camada de saída.

- A saída gerada pela Rede Neural é comparada com o valor desejado, correspondente à amostra de treinamento particular. O erro (Equação 3), que corresponde à diferença entre a saída gerada e a saída desejada, é então usado para computar a mudança na conexão. Desse modo, os erros são propagados no sentido inverso.

Este processo acontece para o restante das amostras de treinamento, submetendo iterativamente o conjunto até que o erro alcance um nível baixo e aceitável.

2.2.4 – Algoritmos de Classificação Por Regiões

Os classificadores por regiões utilizam como critério de decisão a informação espectral de cada *pixel* da imagem e a informação espacial que envolve a relação entre o *pixel* e seus vizinhos. Estes classificadores procuram simular o comportamento de um fotointérprete, ao reconhecer áreas homogêneas nos dados orbitais, baseados nas propriedades espectrais e espaciais dessas áreas na imagem. A classificação por regiões é dividida em fases, ou seja, segmentação, extração de regiões, classificação e mapeamento (MOREIRA, 2005).

Para os classificadores por regiões, o processo de segmentação de imagens é um dos primeiros passos em análise de imagens de sensoriamento remoto. A imagem é dividida em regiões que devem corresponder às áreas de interesse da aplicação. Atributos de regiões, tais como área, forma, parâmetros estatísticos e textura podem ser extraídos e usados posteriormente nas análises dos dados (BINS *et al.*, 1996).

Segundo BINS *et al.* (1996), a divisão em porções consiste basicamente em um processo de crescimento de regiões, uma técnica de agrupamento de dados, na qual somente as regiões adjacentes, espacialmente, podem ser agrupadas. No início da segmentação cada *pixel* da imagem é considerado uma região, chamado “semente”. Para a união de duas regiões A e B vizinhas, deve-se adotar o seguinte critério:

- A e B são similares (teste das médias);
- a similaridade, definida pelo usuário, satisfaz o limiar estabelecido;

- A e B são mutuamente próximas (dentro os vizinhos de A, B é a mais próxima, e dentro os vizinhos de B, A é a mais próxima).

Se A e B satisfazem os critérios acima, estas regiões são agregadas, caso contrário, o sistema reinicia o processo de teste de agregação.

A etapa de Extração de Regiões consiste em extrair alguns parâmetros do polígono, como: posicionamento, média de níveis de cinza e matriz de covariância, para serem utilizados na classificação (BINS et al., 1996).

2.2.4.1 – Algoritmo Bhattacharya

A medida da distância de Bhattacharya é usada neste classificador por regiões, para medir a separabilidade estatística entre um par de classes espectrais, ou seja, mede a distância média entre as distribuições de probabilidades de classes espectrais (INPE, 2004).

O classificador Bhattacharya é um algoritmo de classificação supervisionada, que requer a seleção de áreas de treinamento, podendo utilizar as regiões separadas durante o processo de segmentação ou polígonos representativos das regiões a serem classificadas. Ele utiliza as amostras de treinamento para estimar a função densidade de probabilidade das classes apontadas no treinamento. Em seguida, avalia, em cada região, a distância de Bhattacharya entre as classes, conforme Equação 4 (MOREIRA, 2005).

$$B(p_i, p_j) = \frac{1}{2}(m_1 - m_2)^T \sum (m_i - m_j) + \frac{1}{2} \ln \frac{|\sum (m_i - m_j)|}{|\sum i|^{1/2} |\sum j|^{1/2}} \quad (4)$$

Em que:

B = distância de Bhattacharya;

p_i e p_j = *pixels* nas classes i e j ;

m_i e m_j = médias das classes i e j ;

T = matriz transposta;

\ln = logaritmo neperiano;

i e j = classes dentro do contexto.

2.2.4.2 – Algoritmo Iseseg

O Iseseg é um algoritmo de agrupamento de dados não-supervisionado, aplicado sobre o conjunto de regiões que foram caracterizadas por seus atributos estatísticos (média e matriz de covariância) na fase de extração de regiões. Em outras palavras, é uma técnica para classificação que procura agrupar regiões, a partir de uma medida de similaridade entre elas. A medida de similaridade utilizada consiste na distância de Mahalanobis (Equação 5) entre a classe e as regiões candidatas à relação de pertinência com esta classe (MOREIRA, 2005).

$$D = \frac{1}{2} (X - m_i)^T C_i^{-1} (X - m_j) \quad (5)$$

Em que:

D = distância de Mahalanobis;

T = matriz transposta;

C_i = matriz de covariância;

m_i e m_j = vetor de média das classes i e j ;

X = região de análise.

O Iseseg utiliza os atributos estatísticos das regiões, matriz de covariância e vetor de média, para estimar o valor central de cada classe. INPE (2004) resume os procedimentos em três etapas, descritas a seguir:

1. O usuário define um limiar de aceitação, dado em percentagem, que por sua vez define uma distância (Mahalanobis), de forma que todas as regiões pertencentes à determinada classe estarão a uma distancia inferior à definida. Aumentando-se o limiar, maior será a distância, diminuindo também o número de classes detectadas pelo algoritmo.
2. As regiões são ordenadas de forma decrescente de área, para que se iniciem os procedimentos para agrupá-las em classes. As classes com maiores áreas, que não tenham sido associadas a nenhuma outra classe, servem como parâmetro para o agrupamento, considerando a média e a matriz de covariância como parâmetros estatísticos da região. Em seguida, associa-se a esta classe todas as regiões cuja distância de Mahalanobis for inferior a distância definida pelo limiar de

aceitação. Esta fase repete-se até que todas as regiões tenham sido associadas a alguma classe.

3. As regiões são reclassificadas, considerando-se os novos parâmetros estatísticos das classes, definidos na etapa anterior.

2.2.4.3 – Algoritmo do Histograma

Este algoritmo é baseado na técnica de crescimento de regiões tradicional, com algumas modificações as quais parcialmente resolvem o problema de dependência na ordem das uniões. O algoritmo do Histograma se encontra implementado no programa computacional Spring 4.2.

O classificador por histograma é um algoritmo de *clustering* de regiões que, ao invés de usar a distância euclidiana entre as médias das regiões, computa a diferença entre os histogramas das mesmas (BINS, 2006). O cálculo da distância entre duas regiões é feito conforme Equação 6.

$$D_{(A,B)} = \sqrt{\sum (DH_{(A,B)})^2} \quad (6)$$

Em que:

$D_{(A,B)}$ = distância entre duas regiões e / ou "clusters";

$DH_{(A,B)}$ = diferença entre dois histogramas;

2.3 – Avaliação das Imagens Classificadas

Um dos procedimentos cruciais no processamento de imagens digitais refere-se à validação das imagens temáticas geradas após o processo de classificação. Os temas que compõem um mapa, que é o produto final da classificação, traduzem de forma simplificada, o que os sensores orbitais captaram das feições terrestres.

Em geral, aplicam-se testes estatísticos para avaliar a exatidão das imagens temáticas derivadas de dados de sensores orbitais. O procedimento consiste na determinação de um número ótimo de unidades de amostras representativas dos dados de referência e compará-los às imagens temáticas, assumindo uma margem de erros permissível. Se o erro for

menor ou igual à margem assumida, a imagem é aceita como exata dentro de intervalos de confiança especificados (CONGALTON, 1998).

FITZPATRICK-LINS (1980) recomenda uma relação proveniente da teoria de probabilidade binomial (Equação 7) para determinar o número ideal de pontos (*pixels*) para serem testados em classificações de uso de terra.

$$N = [(4p)(q \approx)] / E^2 \quad (7)$$

Em que:

N = número de pontos a serem amostrados;

p = percentual esperado de exatidão;

$q \approx$ = diferença entre 100% e p ;

E = erro permissível.

MATHER e TSO (2001) propõe a Equação 8 para encontrar o número total de pixels, a serem utilizados como amostra de treinamento.

$$N = 30 \times vd \times C \quad (8)$$

Em que:

N = número de pontos a serem amostrados;

vd = variáveis discriminantes;

C = classes informacionais.

Os resultados da comparação entre a imagem classificada e os dados de referência são organizados em uma tabela, chamada Matriz de Contingência ou Matriz de Erros. O formato utilizado para organizar os resultados é mostrado no Quadro 2, de acordo com CONGALTON e GREEN (1998).

Quadro 2 - Representação genérica de uma matriz de erros.

	<i>j</i> = colunas (imagem de referência)			total linha <i>n_{i+}</i>	
	1	2	<i>K</i>		
<i>i</i> = linhas imagem classificada	1	<i>n₁₁</i>	<i>n₁₂</i>	<i>n_{1k}</i>	<i>n₁₊</i>
	2	<i>n₂₁</i>	<i>n₂₂</i>	<i>n_{2k}</i>	<i>n₂₊</i>
	<i>K</i>	<i>n_{k1}</i>	<i>n_{k2}</i>	<i>n_{kk}</i>	<i>n_{k+}</i>
	total coluna <i>n_{+j}</i>	<i>n₊₁</i>	<i>n₊₂</i>	<i>n_{+k}</i>	<i>N</i>

A partir das informações presentes na matriz de erros é possível extrair vários parâmetros estatísticos para comprovar a eficiência de uma imagem classificada, incluindo os erros de omissão e comissão, índice de exatidão global e índice Kappa (CAMPBELL, 1987).

2.3.1 – Índice Kappa

Nos últimos anos, o índice Kappa tem sido um dos índices estatísticos mais utilizados para validar imagens temáticas derivadas de dados de sensores remotos. Segundo GONG e HOWART (1990), ele tem a vantagem de representar inteiramente a matriz de erros, ao contrário do índice de exatidão global que utiliza apenas as informações que se situam na diagonal da matriz.

A partir das informações presentes na matriz de erros (Quadro 2), calcula-se o índice kappa, que é dado pela Equação 9.

$$\hat{k} = \frac{n \sum_{i=1}^k n_{ii} - \sum_{i=1}^k (n_{i+} \cdot n_{+i})}{n^2 - \sum_{i=1}^k (n_{i+} \cdot n_{+i})} \quad (9)$$

em que

k = Número de linhas da matriz de confusão.

n = número total de observações da matriz.

n_{ii} = número de observações na linha i e coluna i , na diagonal principal

$$(a) n_{i+} = \sum_{j=1}^k n_{ij} \quad (b) n_{+j} = \sum_{i=1}^k n_{ij} \quad (10)$$

De acordo com CONGALTON e GREEN (1998), as classificações das imagens geradas em sensoriamento remoto são classificadas em determinados intervalos de valores, conforme os níveis de aceitação exibidos no Quadro 3.

Quadro 3 - Intervalo de aceitação dos resultados do índice Kappa.

Índice Kappa (K)	Características
$K \leq 0,4$	Pobre
$0,4 < K \leq 0,8$	Razoável
$K \geq 0,8$	Excelente

Para se comparar dois índices Kappa quaisquer CONGALTON e GREEN (1998) propõem a Equação 11. O índice Z , resultado da equação, se maior que o Z tabelado, demonstra se existe diferença estatística significativa entre os resultados de Kappa. Para 95% de probabilidade de ocorrências, o valor tabelado é 1,96.

$$Z = \frac{|\hat{k}_1 - \hat{k}_2|}{\sqrt{\text{vâr}(\hat{k}_1) + \text{vâr}(\hat{k}_2)}} \quad (11)$$

Em que:

Z = índice de comparação;

\hat{k}_1 = coeficiente kappa do método 1;

\hat{k}_2 = coeficiente kappa do método 2;

$\text{vâr}(\hat{k}_1)$ = variância do coeficiente kappa do método 1; e

$\text{vâr}(\hat{k}_2)$ = variância do coeficiente kappa do método 2.

A variância do índice kappa é dada pela Equação 12, de acordo com CONGALTON e GREEN (1998).

$$\widehat{\text{var}}(\hat{K}) = \frac{1}{n} \left\{ \frac{\theta_1(1 - \theta_1)}{(1 - \theta_2)^2} + \frac{2(1 - \theta_1)(2\theta_1\theta_2 - \theta_3)}{(1 - \theta_2)^3} + \frac{(1 - \theta_1)^2(\theta_4 - 4\theta_2^2)}{(1 - \theta_2)^4} \right\} \quad (12)$$

Sendo que:

$$\theta_1 = \frac{1}{x} \sum_{i=1}^r x_{ii} \quad (13)$$

$$\theta_2 = \frac{1}{x^2} \sum_{i=1}^r x_{i+}x_{+i} \quad (14)$$

$$\theta_3 = \frac{1}{x^2} \sum_{i=1}^r x_{ii}(x_{i+} + x_{+i}) \quad (15)$$

$$\theta_4 = \frac{1}{x^3} \sum_{i=1}^r \sum_{j=1}^r x_{ij}(x_{j+} + x_{+i})^2 \quad (16)$$

Em que:

x = elemento da matriz de erros;

r = número de categorias presentes na matriz de erros;

x_{ii} = elementos da diagonal principal;

x_{i+} = total da linha para uma dada categoria informacional;

x_{+i} = total da coluna para uma dada categoria informacional;

x_{ij} = número de amostras classificadas em uma categoria i na classificação e categoria j nos dados de referência; e

x_{j+} = número de amostras classificadas na categoria j nos dados de referência.

2.4 – Aplicação de Imagens de Alta Resolução

Em recente trabalho desenvolvido, RIBEIRO (2003) avaliou diferentes métodos de classificação de uma imagem IKONOS II para mapeamento da cobertura terrestre. Os algoritmos utilizados, Distância Euclidiana Mínima, Máxima Verossimilhança e Redes Neurais (*backpropagation*), foram

testados com base em amostragem por pontos (pixels) e por regiões, segundo o índice Kappa. Os resultados encontrados indicaram que os métodos de Redes Neurais e Máxima Verossimilhança, ambos por regiões, foram os de melhores desempenhos para a classificação de imagens.

MEDEIROS e PETTA (2005) utilizaram imagens do satélite IKONOS II para mapear o uso e ocupação do solo em Parnamirim, estado do Rio Grande do Norte. A metodologia utilizada para a elaboração do mapa de uso e ocupação do solo consistiu na aplicação de técnicas de processamento digital de imagens, usando-se o algoritmo da máxima verossimilhança. Os autores ressaltaram que as classificações geradas não foram suficientes para elaborar o mapa de uso e ocupação do solo, devido ao fato de diferentes classes de ocupação do solo terem assinaturas espectrais semelhantes, porém, as mesmas ajudaram consideravelmente na diferenciação das classes de vegetação (vegetação de restinga e mangue).

BORGES et al. (2005) desenvolveram uma metodologia para avaliar e monitorar o processo de assoreamento em reservatórios, lagos, lagoas e rios. Para cumprir esse objetivo, os autores estudaram o reservatório Morro Grande, com 723.206,31 m² de área superficial, localizado no exutório da bacia hidrográfica do Rio Preto (1.100 km²), no município de Areal/RJ, utilizando informações de sensoriamento remoto, associadas ao método geoestatístico de krigagem Bayesiana, tendo como base uma imagem do sensor IKONOS. Os autores concluíram que a imagem demonstrou bom desempenho para auxiliar na extração de informação batimétrica, mostrando potencialidades para o estudo do assoreamento em reservatórios.

KURASZ *et al.* (2005) mapeou a paisagem no entorno da Reserva Florestal Embrapa/Epagri, município de Caçador, estado de Santa Catarina, utilizando imagem de alta resolução IKONOS II, de modo a diagnosticar o seu estado atual como zona tampão para a área da Reserva, ou seja, sua capacidade de atuar como amortecedor para as atividades fora da área protegida. O mapa de cobertura do solo foi obtido por meio da interpretação visual da imagem e poligonização das unidades de paisagem na tela do computador, considerando-se os diferentes padrões tonais, o tamanho, a forma e a textura, no software ArcView 3.2a.

NASCIMENTO *et al.* (2006) elaboraram um diagnóstico, em nível de paisagem, dos fragmentos de vegetação florestal nativa, através de uma classificação visual da imagem do satélite IKONOS II, para uma bacia hidrográfica do rio Alegre, situada no extremo sul do Estado do Espírito Santo. A imagem mostrou ser de excelente qualidade, permitindo o mapeamento de 475 fragmentos florestais.

RAMIREZ *et al.* (2006) avaliou o impacto do aumento da resolução espacial e radiométrica de uma imagem pancromática do IKONOS II na identificação de plantios de café (*Coffea arabica*), em comparação com as imagens do Landsat/ETM+. Em uma região localizada no Município de Pedregulho, Estado de SP. Os autores concluíram que a imagem pancromática do satélite IKONOS-II mostrou melhor desempenho na identificação de talhões de café com características diferentes, incluindo idade, espaçamento, estado fitossanitário, etc., em comparação ao sensor Landsat/ETM+.

RIBEIRO (2003), ao classificar uma imagem do satélite IKONOS II utilizando diferentes algoritmos de classificação encontrou os seguintes índices Kappa: Máxima Verossimilhança por pixel, 0,652; Máxima Verossimilhança por região, 0,824; Distância Euclidiana mínima por pixel, 0,537; Distância Euclidiana Mínima por região, 0,655; Redes Neurais por pixel, 0,680 e Redes Neurais por região, 0,824. Analisando os resultados encontrados, nota-se uma grande semelhança com o presente estudo ao se comparar os índices Kappa obtidos pelos algoritmos por *pixel* e por regiões.

3 – MATERIAL E MÉTODOS

3.1 – Localização e Caracterização da Área de Estudo

A área de estudo localiza-se no município de Serra do Salitre, estado de Minas Gerais, entre as coordenadas geográficas 19° 14' S e 19° 20' de latitude sul e 46° 43' W e 46° 37' de longitude oeste.

O município está inserido na Mesorregião do Triângulo Mineiro/Alto Paranaíba, cuja área da unidade territorial é de 1.298 km², possuindo uma população estimada, em 2005, de 10.271 habitantes (IBGE, 2006).

A área possui, segundo a classificação de Köppen, clima predominante tipo Cwa, temperado chuvoso, com inverno seco e verão chuvoso, com precipitação média anual de 1.600 mm e a média das temperaturas, mínima e máxima, de 14,5° e 24,8° C, respectivamente. A altitude da área de estudo varia de 850 a 970 m, onde predomina o cerrado, que se encontra bastante modificado por ações antrópicas (VILELA, 2002).

3.2 – Equipamentos e Sistemas Computacionais

Os dados foram processados utilizando um microcomputador AMD Athlon 950 Mhz, com HD de 40 GB e memória RAM de 256 MB, e os seguintes sistemas computacionais:

- Spring 4.2: Foram processadas as classificações utilizando os algoritmos Máxima Verossimilhança, Distância Euclidiana Mínima, Bhattacharya, Isoseg e Histograma.
- Idrisi – *The Kilimanjaro Edition*: As imagens temáticas classificadas foram testadas, com base no índice Kappa;

3.3 – Dados

Foi utilizada parte de uma imagem do sensor IKONOS II obtida no dia 2 de março de 2002, com área de aproximadamente 465,00 ha. A imagem contém as bandas azul (0,45 a 0,52 µm), verde (0,52 a 0,60 µm), vermelho

(0,63 a 0,69 μm) e infravermelho próximo (0,76 a 0,90 μm), com 4 metros de resolução espacial (Figura 2).

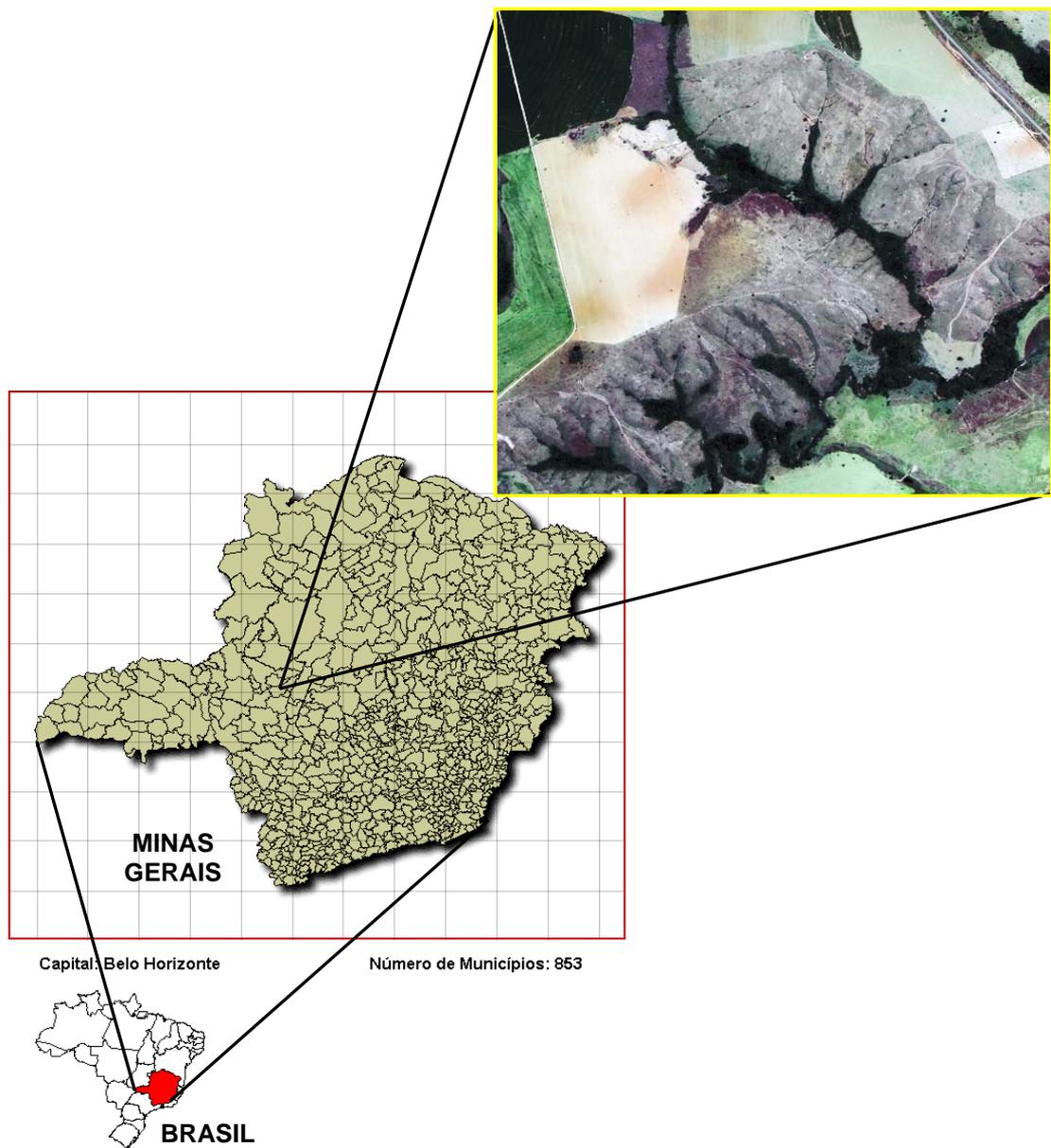


Figura 2 - Localização da área de estudo.

3.4. Procedimentos Metodológicos

3.4.1. Classificação da Imagem IKONOS

A imagem utilizada neste trabalho faz parte de um projeto que já vem sendo desenvolvido no Departamento de Engenharia Florestal da Universidade Federal de Viçosa. O motivo de sua escolha é a pouca

disponibilidade de recursos, pois não foi necessária a sua aquisição para o presente estudo.

Utilizou-se os seguintes algoritmos de classificação para geração das imagens temáticas, com os propósitos de mapeamento da cobertura terrestre: Máxima Verossimilhança, Distância Euclidiana Mínima, Redes Neurais, Bhattacharya, Isoseg e Histograma.

Os dados foram processados no programa computacional Spring 4.2, para a geração das imagens temáticas, com a exceção do algoritmo de Redes Neurais, que foi processado no programa computacional Idrisi – *The Kilimanjaro Edition*. Os algoritmos de classificação utilizados se encontram relacionados nos itens 2.2.3 e 2.2.4, que se distinguem pelos métodos de classificação por *pixels* e por regiões.

Foram identificadas, após uma visita ao local de estudo, sete classes predominantes dentro da imagem, cujos temas encontram-se descritos no Quadro 4. Foram atribuídas nomenclaturas às classes para serem utilizadas nas Matrizes de Erros.

Quadro 4 - Relação de classes temáticas utilizadas nas classificações.

Item	Nome	Cor	Descrição
1	Café		Áreas compostas por culturas perenes, já com alguns anos de implantação.
2	Solo Exposto		Áreas preparadas para o cultivo agrícola, desprovidas de qualquer tipo de vegetação.
3	Mata Ciliar		Áreas de floresta densa que acompanham o leito dos córregos e rios.
4	Campo Sujo		Áreas com predomínio de gramíneas, além de muitas espécies arbustivas de médio porte.
5	Braquiária		Áreas transformadas para pastagem, em estágio de maturidade avançada.
6	Pastagem		Áreas transformadas para pastorio, em estágio inicial de maturidade.
7	Campo Limpo		Áreas compostas predominantemente por gramíneas, com algumas espécies arbustivas.

Para os algoritmos de classificação por Máxima Verossimilhança, Distância Euclidiana Mínima, Redes Neurais, todos por pixels e o Bhattacharya por regiões, foram aplicados o método supervisionado, que envolve a coleta de amostras para treinar os algoritmos classificadores. As amostras foram retiradas de regiões representativas de cada classe, adotando-se apenas uma área para cada tipologia, considerando o tamanho da área e a quantidade de classes estudadas. Como pode ser verificado na Figura 3, os polígonos onde foram retiradas as amostras para treinar os algoritmos ficaram localizados em regiões onde havia homogeneidade dentro das classes.

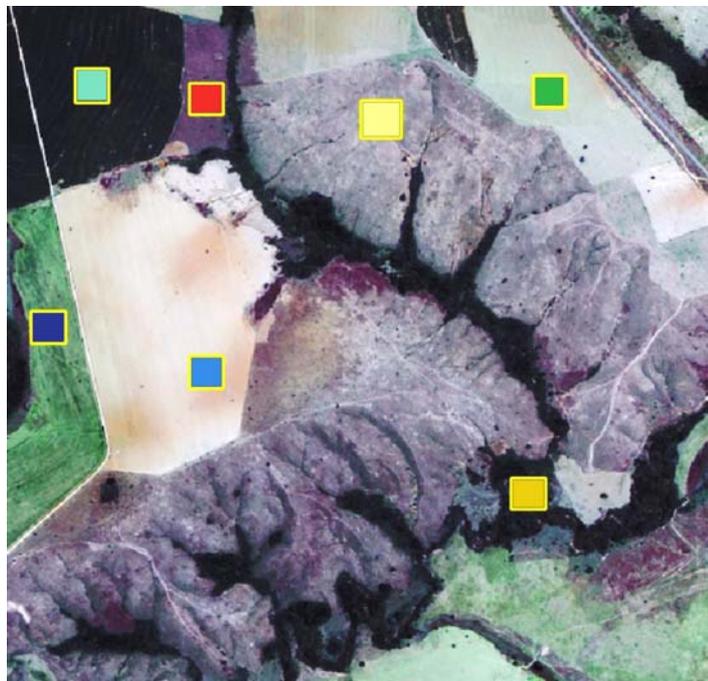


Figura 3 - Imagem da área de estudo, composição R(3), G(2) e B(1), com as áreas de coleta das amostras de treinamento.

A Figura 4 exibe as configurações utilizadas no programa computacional Idrisi para gerar a imagem classificada pelo algoritmo de Redes Neurais. Foram utilizadas na classificação as bandas 2, 3 e 4 (verde, vermelho e infravermelho próximo). A utilização das três bandas demonstrou, com base nos testes pelo índice Kappa, resultados melhores do que utilizando as quatro bandas, com a inserção da banda 1.

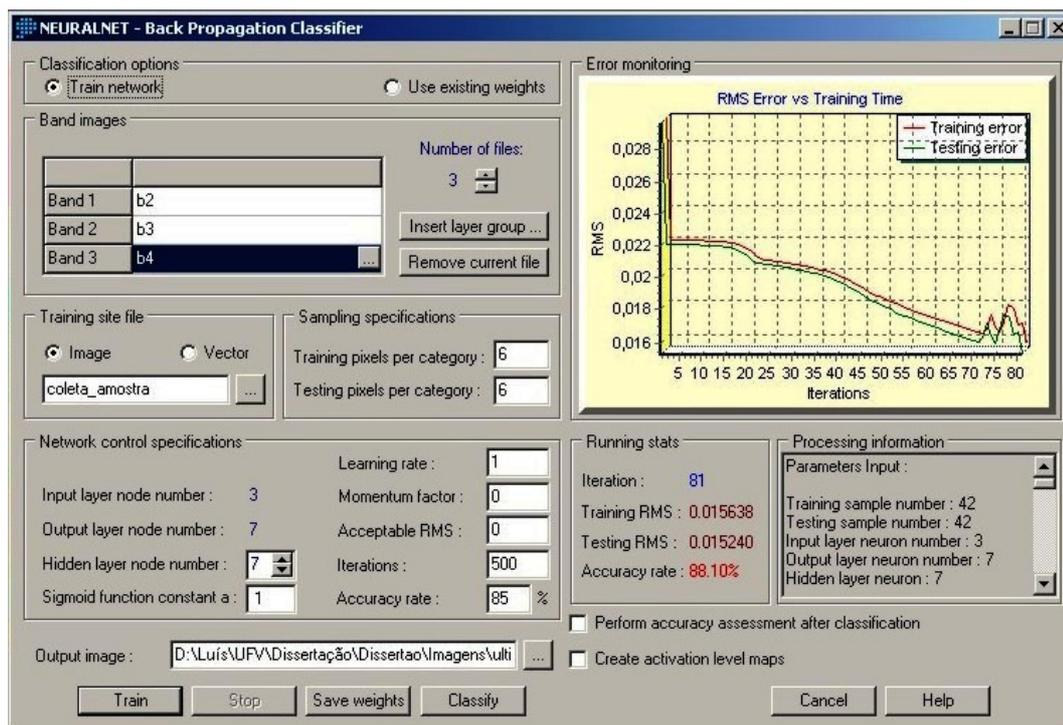


Figura 4 - Configurações do algoritmo de Redes Neurais no programa computacional Idrisi – *The Kilimanjaro Edition*.

O algoritmo Bhattacharya utiliza, além das amostras de treinamento, uma imagem segmentada para gerar a imagem classificada, cujos índices de Similaridade e Área (*pixels*) foram 70 e 200, respectivamente (Quadro 5). O Limiar de Aceitação utilizado foi de 100%. Segundo INPE (2004), o limiar dado em porcentagem define uma distância de Mahalanobis, de forma que todas as regiões pertencentes a uma dada classe estão distantes da classe por uma distância inferior a esta. Quanto maior o limiar, maior esta distância e conseqüentemente menor será o número de classes detectadas pelo algoritmo. Os dados de Similaridade, Área e o Limiar de Aceitação foram escolhidos com base nas imagens geradas, pela técnica visual. Na segmentação da imagem foram utilizadas as quatro bandas do satélite IKONOS II (azul, verde, vermelho e infravermelho próximo).

Quadro 5 - Dados utilizados para a segmentação da imagem.

Classificador	Similaridade	Área (<i>pixels</i>)
Battacharya	70	200
Isoseg	45	200
Histograma	38	200

De acordo com BINS *et al.* (1996), o segmentador implementado no programa computacional SPRING é baseado nas técnicas de segmentação tradicionais, com algumas modificações que parcialmente resolvem o problema da dependência na ordem de unir as regiões. Esta especificação impõe que as uniões iniciais são mais difíceis de realizar do que as uniões que ocorrem no final do processo de fusão.

O classificador por regiões Isepeg foi implementado em conjunto com a imagem segmentada pelos índices de Similaridade e Área (pixels) com os valores de 45 e 200, respectivamente (Quadro 5), além de ser utilizado um limiar de 75%. Assim como para o algoritmo Bhattacharya, os valores de Similaridade, Área e Limiar de Aceitação para o algoritmo Isepeg foram selecionados de acordo com a qualidade visual da imagem gerada. Para esse classificador também foram utilizadas as 4 (quatro) bandas do satélite IKONOS II. Após o processo de classificação foram feitos agrupamentos de modo que o mapa gerado permanecesse apenas com as 7 (sete) classes utilizadas neste estudo.

O algoritmo classificador por regiões Histograma também foi aplicado juntamente com a imagem segmentada, cujos índices de Similaridade e Área (pixels) aplicados foram 38 e 200, respectivamente (Quadro 5), selecionados da mesma maneira que os algoritmos por regiões, empregados anteriormente. Nesta classificação também foram aplicadas as 4 (quatro) bandas do satélite IKONOS II. Para o algoritmo Histograma é necessário somente informar a quantidade de classes a serem apresentadas ao final da classificação.

3.4.2 – Quantidade de Amostras Testes Para Avaliação das Imagens Classificadas.

O tamanho da amostra teste a ser utilizada na avaliação das imagens classificadas é altamente variável, não tendo um padrão recomendável pela comunidade científica. Como ela requer, na maioria das vezes, checagem de campo, o processo torna-se bastante oneroso se o número mínimo de *pixels* utilizados for grande.

De acordo a Equação 7, o número total de *pixels* testes deveriam estar em torno de 750. Adotando a Equação 8, o número de *pixels* a ser utilizado é 840. Nos dois cálculos analisados, considerando-se 95% de probabilidade. Entretanto, observa-se na literatura consultada que este número tem sido bem mais alto. Assim, neste trabalho, optou-se por utilizar diferentes tamanhos de amostras, a saber: 250, 500, 750 e 1000 *pixels*, para verificar se haveria perda de exatidão ao se utilizar um menor número de *pixels*. A Figura 5 mostra o adensamento das Amostras Testes, excluindo as áreas de coleta das Amostras de Treinamento.

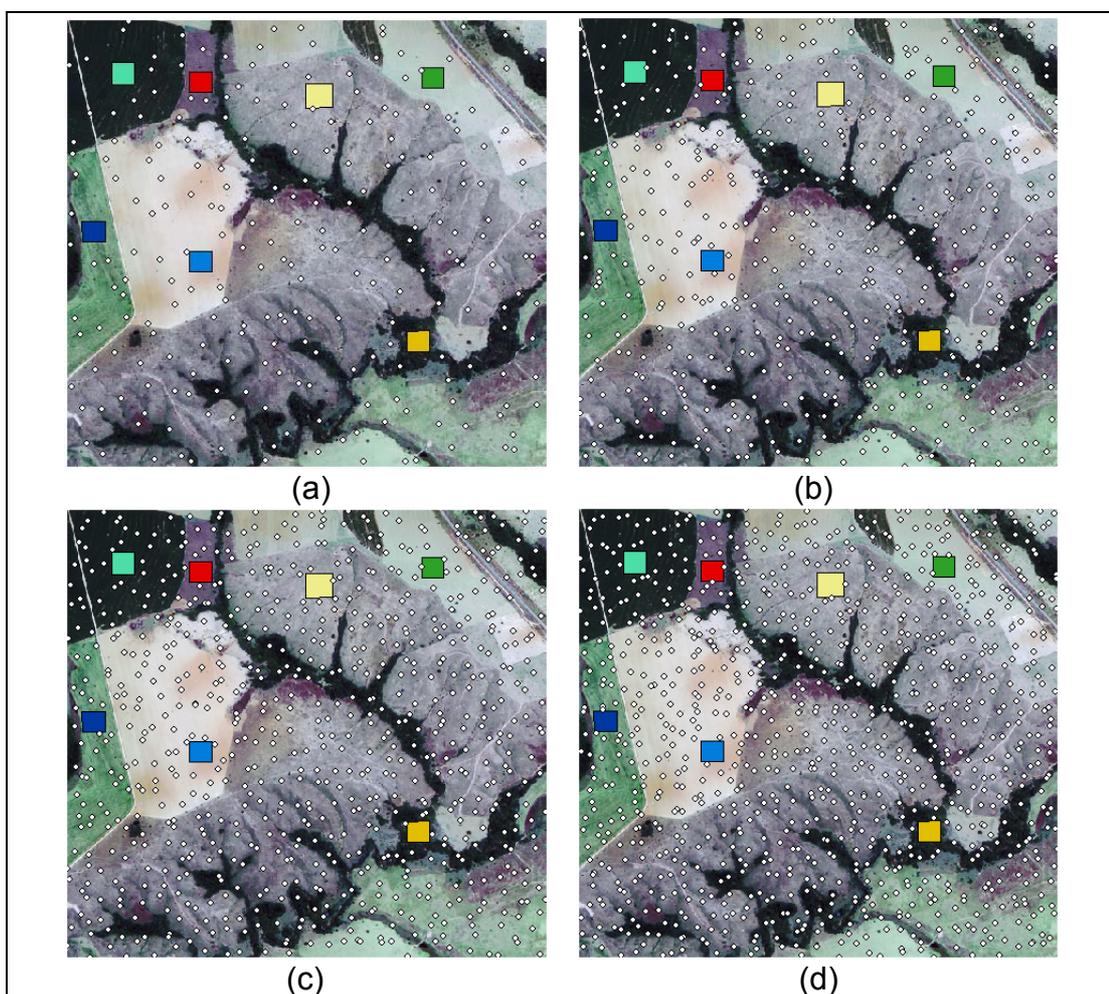


Figura 5 - Adensamento dos pontos amostrais para os testes de acurácia: (a) 250 (b) 500 (c) 750 e (d) 1.000 amostras

As Amostras testes foram coletadas de maneira aleatória, para a avaliação das várias imagens temáticas geradas durante o processo de classificação. O programa computacional Idrisi foi utilizado para a geração das amostras testes e geração das Matrizes de Erros, a partir das quais se

computaram os parâmetros estatísticos para a validação das imagens. Os valores encontrados durante o processo de validação estão sumarizados nas Matrizes de Erros localizadas no Apêndice I.

As Amostras Testes para o cálculo da exatidão que coincidiram dentro da área de coleta das amostras de treinamento foram excluídas do processo para evitar alta correlação dos dados. As perdas referentes aos *pixels* que coincidiram estão resumidas no Quadro 6.

Quadro 6 - Quantidades amostradas e as perdas por sobreposição.

Quantidade Inicial	<i>Pixels</i> que coincidiram	Quantidade Final	Perda de amostra (%)
250	12	238	4,80
500	22	478	4,40
750	19	731	2,53
1.000	30	970	3,00

3.4.3. As comparações entre as quantidades de amostras e índices

Kappa

A resultante de Z foi calculada pela razão entre a diferença de dois índices quaisquer (\hat{k}_1 e \hat{k}_2) pela raiz quadrada da soma das variâncias (σ^2_1 e σ^2_2) destes mesmos índices (Equação 10). O teste estatístico Z foi aplicado para comparar os resultados obtidos na variação do índice Kappa, entre as diferentes quantidades amostrais e técnicas de classificação.

Para verificar se havia diferença significativa entre os valores resultantes do índice Kappa, ao variar a quantidade de amostras, foi considerado cada algoritmo separadamente. Todos os algoritmos e quantidades amostradas foram comparados, a fim de verificar se havia diferença estatística significativa entre os resultados encontrados.

4 – RESULTADOS E DISCUSSÃO

4.1 – Avaliação Qualitativa das Imagens Classificadas

As Figuras 6, 7, 8, 9, 10 e 11 exibem, respectivamente, as imagens classificadas pelos algoritmos da Máxima Verossimilhança, Distância Euclidiana Mínima, Redes Neurais, Bhattacharya, Ioseg e Histograma.

Nas três classificações por *pixels* (Figuras 6, 7 e 8), as imagens apresentaram ruídos dentro das classes, com uma qualidade visual inferior às das classificações por regiões (Figuras 9, 10 e 11).

A imagem temática gerada pelo algoritmo da Máxima Verossimilhança (Figura 6) mostrou uma mistura entre as classes pastagem e braquiária para a porção superior da imagem (1) e entre solo exposto, braquiária e campo limpo na porção oeste (2) e porção leste (3). As classes café e mata ciliar, em geral, foram caracterizadas corretamente. A classe braquiária também teve uma boa classificação na parte sudeste da imagem (4).

O algoritmo da Distância Euclidiana Mínima registrou algumas confusões na classificação dos objetos da superfície terrestre. Conforme observado na Figura 7, houve confusão entre pastagem e campo limpo na porção superior da imagem (5), porém, uma boa separação entre solo exposto e braquiária na porção oeste (6). Houve uma superestimação da classe campo sujo na porção sudoeste da imagem e entorno das matas ciliares, mais localizadas ao sul (7), além de uma confusão entre braquiária, campo limpo e pastagem na porção sudeste (8). Como resultado positivo, observa-se um bom desempenho ao classificar café e mata ciliar, além de delimitar corretamente a classe solo exposto na porção leste da imagem, delineando inclusive uma estreita faixa desta classe (9).

Para o mapa temático gerado pelo algoritmo de Redes Neurais (Figura 8), as classes Café e Mata Ciliar mantiveram o bom desempenho na classificação. Outro ponto positivo observado foi na classificação de Pastagem, na parte superior da imagem (10), sendo o melhor desempenho entre todos os algoritmos. O algoritmo de Redes Neurais delimitou corretamente as classes Pastagem e Solo Exposto, na porção leste da imagem (11), assim como o algoritmo da Distância Euclidiana Mínima.

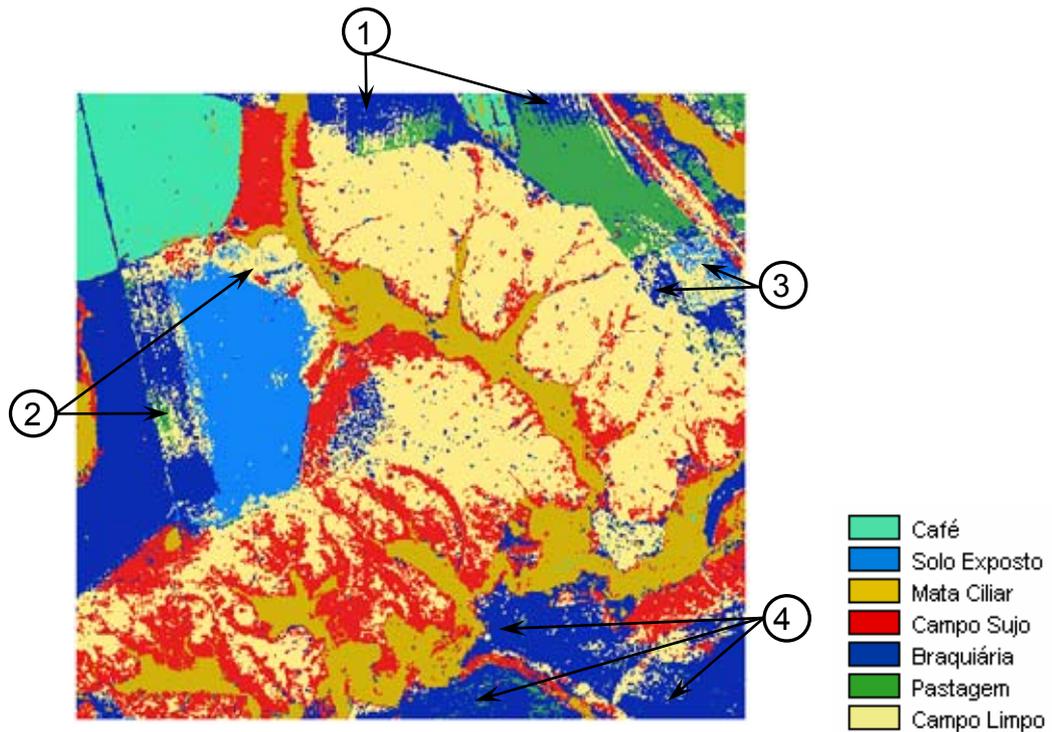


Figura 6 - Mapa Temático gerado pelo algoritmo de Máxima Verossimilhança.

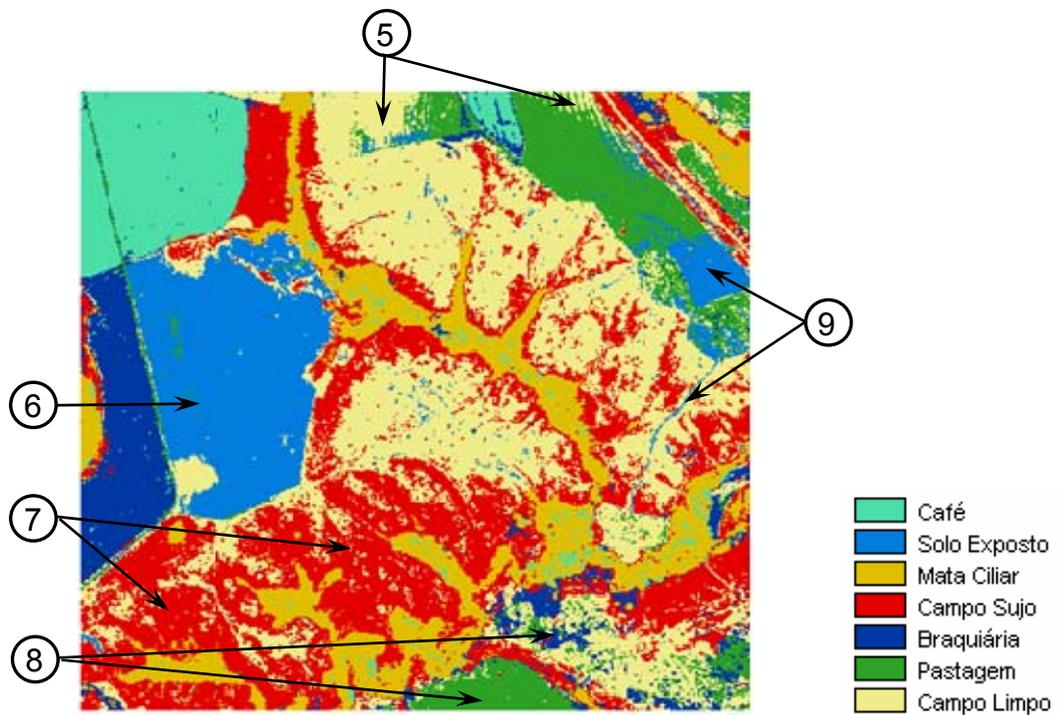


Figura 7 - Mapa Temático gerado pelo algoritmo de Distância Euclidiana Mínima.

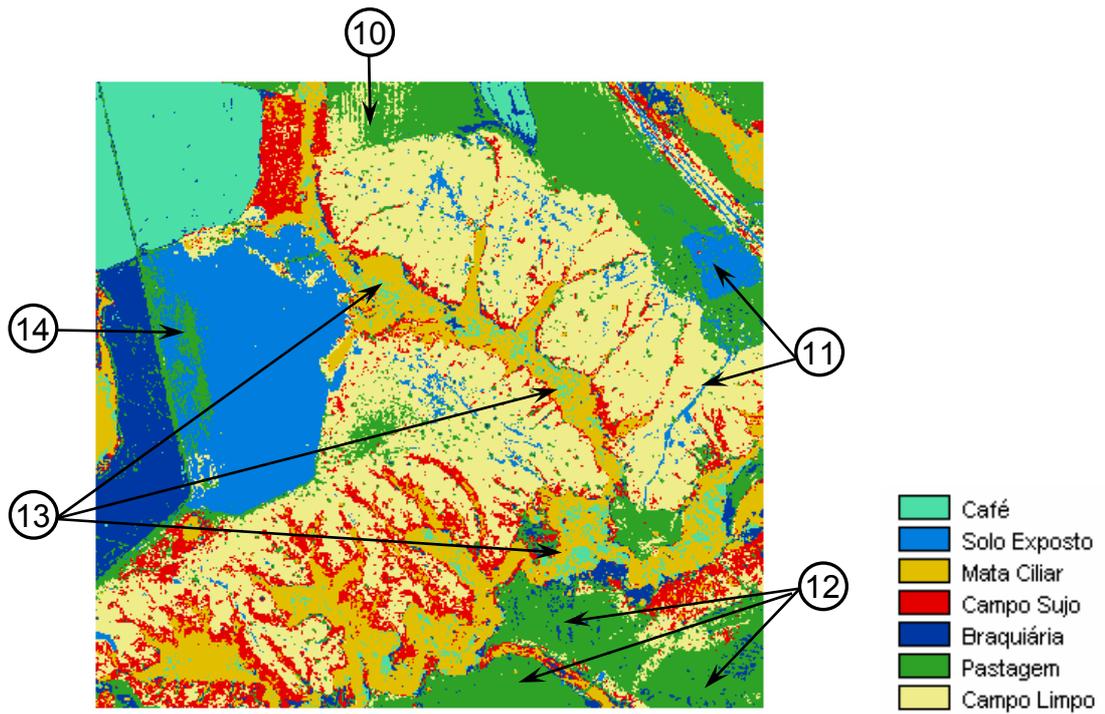


Figura 8 - Mapa Temático gerado pelo algoritmo de Redes Neurais

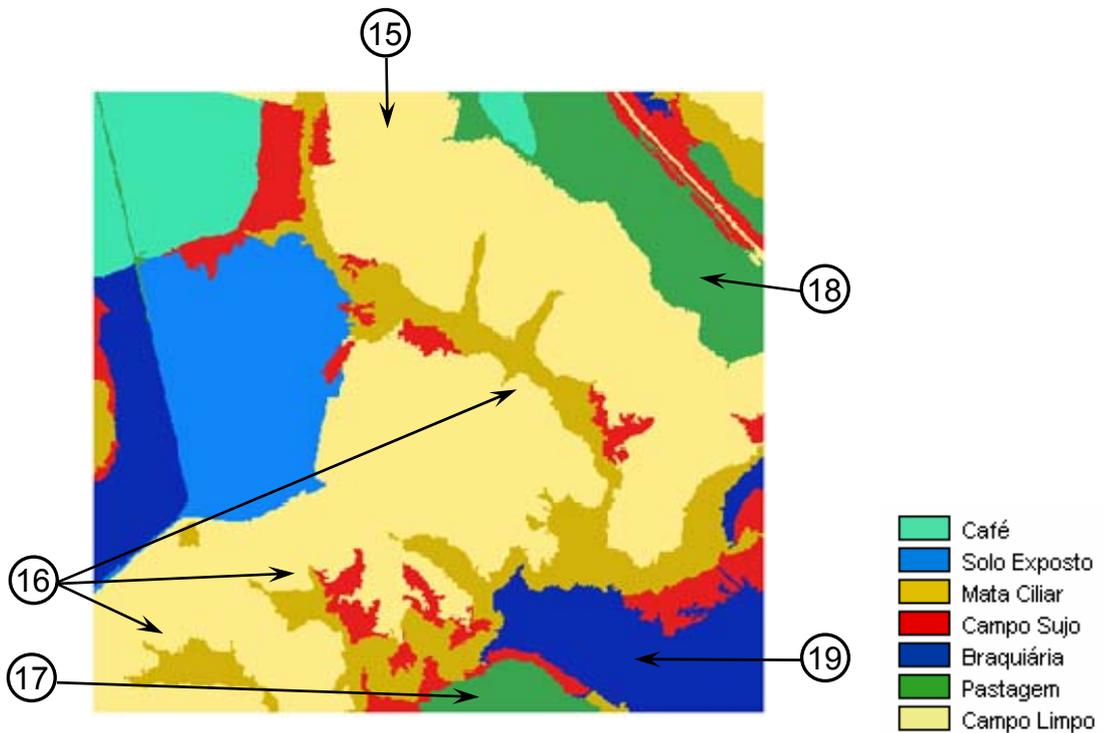


Figura 9 - Mapa Temático gerado pelo algoritmo Bhattacharya.

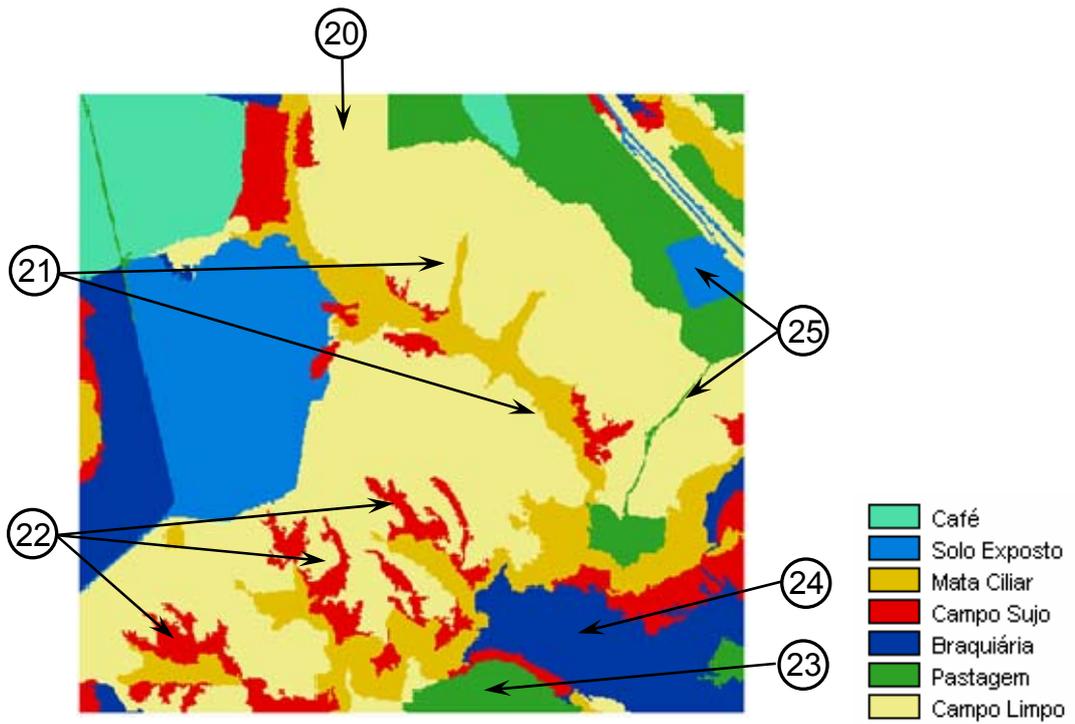


Figura 10 - Mapa Temático gerado pelo algoritmo Isoseg.

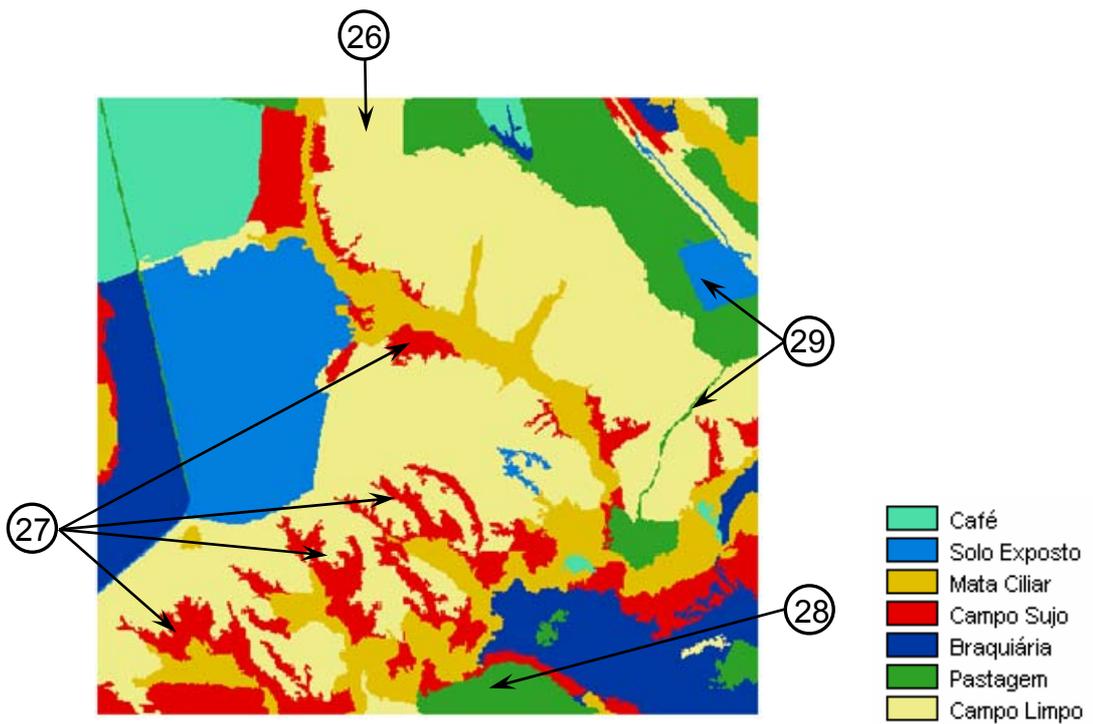


Figura 11 - Mapa Temático gerado pelo algoritmo Histograma.

Continuando a análise do algoritmo de Redes Neurais, observou-se uma confusão na porção sudeste da imagem (12), onde deveria ser classificado como Braquiária, mas foi classificado como Pastagem. Os maiores erros de comissão para a classe Café, dentre todos os algoritmos, foram constatados para o Redes Neurais, cujas confusões foram com a classe Mata Ciliar (13). Na parte oeste da imagem (14), aconteceram confusões entre as classes Campo Sujo e Pastagem.

As imagens classificadas por regiões apresentaram melhores resultados visuais que os algoritmos por *pixels*, com as classes bem definidas nas bordas e sem apresentar ruídos. Entretanto, o algoritmo Bhattacharya, similarmente ao da Distância Euclidiana Mínima, misturou pastagem com campo limpo na porção centro-superior da imagem (15), além de subestimar a classe campo sujo ao longo da classe mata ciliar (16). Houve uma confusão também entre as classes braquiária e pastagem, assim como no algoritmo Distância Euclidiana Mínima, na parte centro-inferior da imagem (17). Uma área com solo exposto também não foi caracterizada, sendo classificado como pastagem, na parte leste da imagem (18). As classes café e mata ciliar, assim como nas classificações por *pixels*, tiveram bom desempenho. A classe solo exposto, na parte oeste da imagem, teve uma ótima delimitação. A classe braquiária foi classificado corretamente na parte sudeste da imagem (19).

O algoritmo Iseseg apresentou algumas melhoras visuais em relação ao Bhattacharya, apesar de manter a confusão entre campo limpo e pastagem na parte superior-central da imagem (20). Outro erro continuou acontecendo com o algoritmo Iseseg, ao subestimar a classe campo sujo ao longo de mata ciliar, na parte central da imagem (21), o que não ocorreu na direção sudoeste (22), onde houve uma melhoria na classificação desta feição. Outro erro de classificação aconteceu na parte sul da imagem (23), com a confusão entre braquiária e pastagem. Na parte sudeste da imagem a classe braquiária foi bem caracterizada (24). Na porção oeste da imagem, as delimitações de contorno das áreas de solo exposto e pastagem foram boas, havendo apenas a confusão ao classificar o caminho como pastagem, sendo que o correto seria classificá-lo como solo exposto (25).

Os erros de confusão cometidos na parte superior-central (26) e inferior-central (28) continuaram ao se utilizar o algoritmo Histograma, havendo uma melhora da classe campo sujo ao longo da mata ciliar (27). Os erros e acertos registrados para o algoritmo Isoseg (25) na porção oeste se repetiram para o algoritmo histograma (29).

É importante ressaltar que, por ter uma alta resolução espacial, as variações dos níveis de cinza do sensor do IKONOS II, considerando a vizinhança, são bem maiores do que as de sensores com menor resolução. Os algoritmos classificadores por *pixels* Máxima Verossimilhança, Distância Euclidiana Mínima e Redes Neurais não consideram a textura da imagem, quando comparam o Valor Nominal de um *pixel* com a estatística de um grupo pré-selecionado (amostras de treinamento). RIBEIRO (2003) utilizou um filtro de mediana para melhorar a apresentação visual das imagens geradas por estes algoritmos, eliminando ruídos dentro das classes.

O processo de segmentação, feito anteriormente à classificação, eliminou os *pixels* isolados, levando em consideração sua vizinhança. Os pixels adjacentes são unidos a uma mesma região quando determinadas características estão estatisticamente dentro de limiares pré-estabelecidos. Com a imagem dividida por regiões, contendo *pixels* de características semelhantes, estas foram classificadas de acordo com o algoritmo utilizado. Esta técnica fez com que a imagem temática final apresentasse regiões mais bem definidas, o que não eliminou os erros de comissão e omissão, inerentes ao processo.

4.2 – Avaliação Quantitativa da classificação

As seis imagens classificadas pelos diferentes algoritmos foram submetidas ao teste de acurácia Kappa, a partir do uso de imagens de referências, com diferentes tamanhos de amostras testes. Os resultados encontram-se reportados no formato de matriz de erros, ou matriz de contingência, que podem ser visualizadas no Apêndice I. A seguir são discutidos os erros de omissão e comissão para todas as imagens classificadas e avaliadas, considerando os quatro diferentes tamanhos de amostras testes. O erro de comissão é um dado percentual relativo à

quantidade de *pixels* atribuídos, erroneamente à determinada classe. Já o erro de omissão é *pixel* omitido dessa determinada classe.

A Figura 12 mostra os erros de comissão e omissão para as imagens classificadas pelo algoritmo da Máxima Verossimilhança. Observa-se que as classes de Café (CAFE) e Solo Exposto (SOEX) apresentaram erros de comissão baixos, próximos de zero. A classe Mata Ciliar (MACI), seguida de Pastagem (PAST) e Campo Limpo (CALI), também apresentaram erros de comissão baixos, inferiores a 20%. Já as classes de Campo Sujo (CASU) e Braquiária (BRAQ) apresentaram erros de comissão próximos a 50%, indicando que essas classes tiveram uma quantidade de *pixels* assinalados erroneamente a elas, na imagem classificada. Ao considerar os erros de omissão, a classe Café (CAFE) apresentou pequenos erros, seguida das classes Mata Ciliar (MACI) e Braquiária (BRAQ), que demonstraram erros abaixo de 20%. As classes Campo Sujo (CASU) e Campo Limpo (CALI) apresentaram erros de omissão em torno de 30%. Já as classes Solo Exposto (SOEX) e Pastagem (PAST) apresentaram os maiores erros de omissão, ficando acima de 40% e 60%, respectivamente.

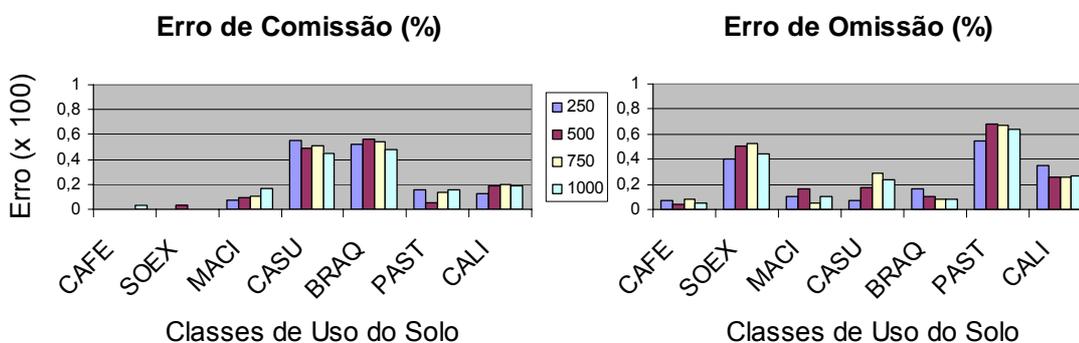


Figura 12 - Erros de comissão e omissão para as feições classificadas pelo algoritmo da Máxima Verossimilhança.

Comparando a imagem temática gerada pelo algoritmo da Máxima Verossimilhança (Figura 6), com os gráficos da Figura 12 e os Quadros 1 a 4 (Apêndice I), percebe-se que o maior erro de comissão da classe Campo Sujo foi por confusão com a classe Campo Limpo. Para a classe Braquiária, houve maiores confusões com as classes Pastagem, Campo Limpo e Solo Exposto. Já para os maiores erros de omissão, as classes Solo Exposto e Pastagem, ambas foram confundidas com as classes Braquiária e Campo

Limpo. Observa-se nos Quadros 1 a 4 que há uma quantidade considerável de *pixels* omitidos da classe Campo Limpo, principalmente por confusão com as classes Campo Sujo seguida de Braquiária, o que não é muito perceptível nos gráficos da Figura 12.

A Figura 13 mostra os erros de comissão e omissão para as imagens classificadas pelo algoritmo da Distância Euclidiana Mínima. Nota-se, novamente, que as classes Café, Solo Exposto e Mata Ciliar apresentaram os melhores resultados para os erros de comissão, inferiores a 20%, seguidas das classes Braquiária, Pastagem e Campo Limpo, em torno de 20%. A classe Campo Sujo foi a que apresentou os maiores erros de comissão, em torno de 60%. Em termos de erros de omissão, a classe de Café apresentou erros próximos de zero, seguida das classes Solo Exposto, Mata Ciliar e Campo Sujo, com erros de omissão girando em torno de 20%. Já as classes Braquiária, Pastagem e Campo Limpo apresentaram os maiores erros de omissão, em torno de 40%.

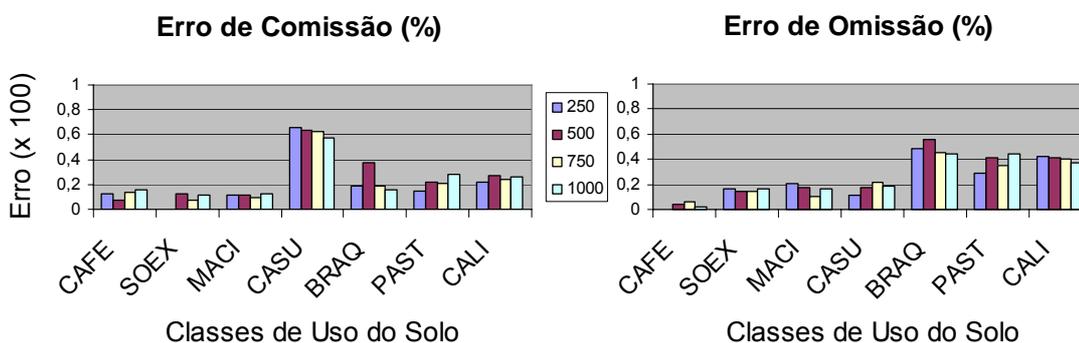


Figura 13 - Erros de comissão e omissão para as feições classificadas pelo algoritmo da Distância Euclidiana Mínima.

A análise dos gráficos da Figura 13 está detalhada nos Quadros 5 a 8 (Apêndice I), onde pode ser observado que o erro de comissão para a classe Campo Sujo é originário da classe Campo Limpo, cujo inverso acontece para os erros de omissão, em que os principais erros da classe Campo Limpo estão relacionados à classe Campo Sujo. Os erros de omissão das classes Braquiária e Pastagem são por confusão com a classe Campo Limpo. A Figura 13 não demonstra essa discrepância de erros da classe Campo Limpo por ela ser a mais representativa dentro da imagem, sendo

perceptível apenas quando analisamos os Quadros, já referenciados neste parágrafo.

A Figura 14 mostra os erros de comissão e omissão para as imagens classificadas pelo algoritmo de Redes Neurais. Analisando as classes em termos de erros de comissão, nota-se que houve um aumento dos erros para as classes de Café, Solo Exposto, Mata Ciliar e Pastagem, em comparação com os erros encontrados nos dois métodos por *pixels*, analisados anteriormente. A classe Café foi confundida com a classe Mata Ciliar. A classe de Campo Limpo teve seus erros de comissão reduzidos a menos de 20%, enquanto que a classe de Campo Sujo continuou com erros relativamente altos. Considerando os erros de omissão, as classes que apresentaram menores erros foram a Café, Pastagem, Solo Exposto e Mata Ciliar, enquanto que as classes Campo Sujo e Braquiária apresentaram os maiores erros, em torno de 60%.

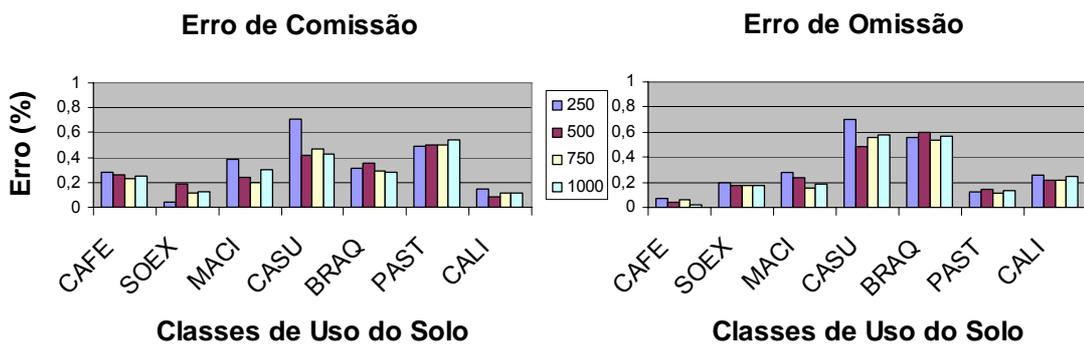


Figura 14 - Erros de comissão e omissão para as feições classificadas pelo algoritmo de Redes Neurais.

Os gráficos da Figura 15 mostram o comportamento dos erros de comissão e omissão para a imagem classificada pelo algoritmo Bhattacharya. Analisando os erros de comissão, nota-se que 06 (seis) das 07 (sete) classes apresentaram erros inferiores a 20%, permanecendo a classe Café com os melhores resultados. A classe Campo Sujo foi a única que obteve erros pouco superiores a 20%. Para os erros de omissão, as classes que apresentaram melhores resultados, nesta ordem, foram as classes Café, Braquiária, Campo Limpo, Mata Ciliar e Solo Exposto, todas com erros próximos a 0 (zero) e inferiores a 20%. A classe Pastagem

apresentou erros de omissão em torno de 30% e Campo Sujo em torno de 50%.

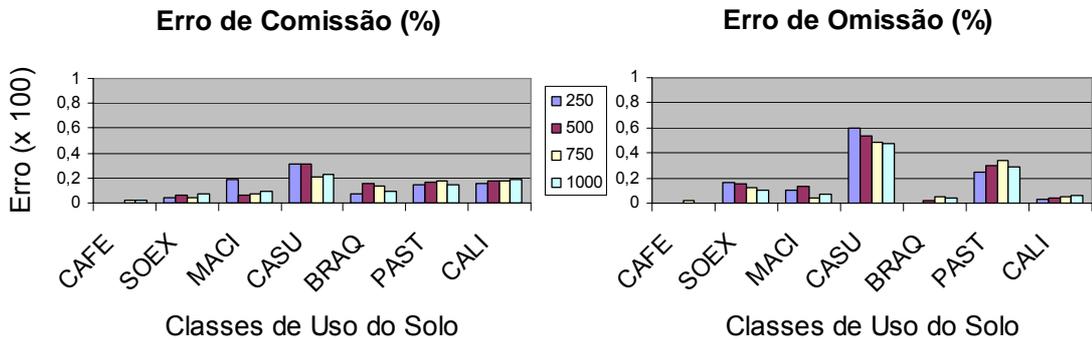


Figura 15 - Erros de comissão e omissão para as feições classificadas pelo algoritmo Bhattacharya.

Correlacionando as Figuras 9 e 15 com os Quadros 13 a 16 (Apêndice I), os principais erros de comissões da classe Campo Sujo foram para Campo Limpo, Mata Ciliar e Pastagem. Os principais erros de omissão das classes Campo Sujo e Pastagem foram por confusão com a classe Campo Limpo. A classe Campo Sujo teve alguns de seus *pixels* atribuídos com moderada intensidade à classe Mata Ciliar.

Os erros de comissão e omissão do algoritmo Iseseg estão exibidos nos gráficos da Figura 16. Analisando as classes em termos de erros de comissão, nota-se que apenas a classe Campo sujo, quando avaliada com 250 amostras, obteve erros acima de 20%, em que novamente a classe Café apresentou os melhores resultados. O desempenho da classe Café também foi melhor ao analisar os erros de omissão. Outra característica semelhante entre os erros de comissão e omissão é que apenas a classe Campo Sujo apresentou erros acima de 20%.

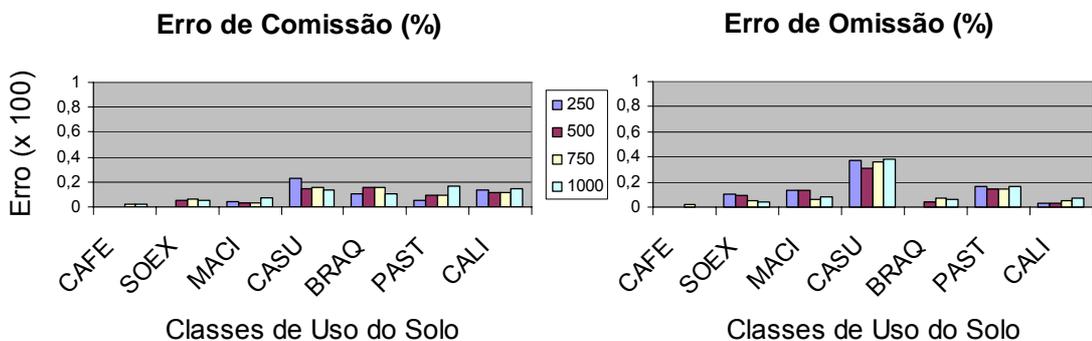


Figura 16 - Erros de comissão e omissão para as feições classificadas pelo algoritmo Iseseg.

A análise conjunta da imagem (Figura 10), gráficos (Figura 16) e matrizes de erros (Quadros 17 a 20 do Apêndice I) nos indica que a classes Campo Sujo teve seus maiores erros de comissão associados às classes Mata Ciliar e Campo Limpo. Os erros de omissão da classe Campo Sujo estão associados principalmente à classe Campo Limpo, cujos resultados nos gráficos da Figura 16, não se destaca mais como em vezes anteriores, por causa de sua área ocupada na imagem de estudo.

Assim como o algoritmo Iseseg, para o Histograma apenas a classe Campo Sujo apresentou resultados dos erros de comissão e omissão superiores a 20%, com a classe Café se destacando com os melhores resultados na classificação digital (Figura 17).

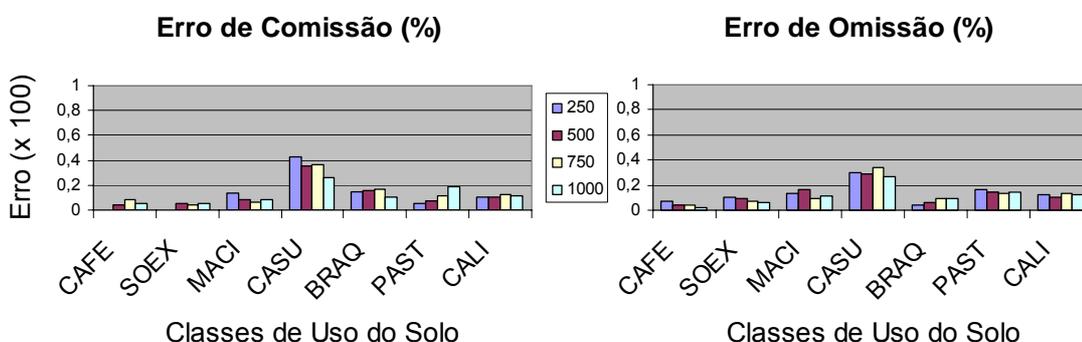


Figura 17 - Erros de comissão e omissão para as feições classificadas pelo algoritmo do Histograma.

Os Quadros 21 a 24 (Apêndice I) indicam outras semelhanças entre os algoritmos Iseseg e Histograma. O Histograma também apresentou maiores erros de comissão entre as classes Campo Sujo e Campo Limpo, bem como erros de omissão entre a classe Campo Sujo e as classes Campo Limpo e Mata Ciliar. O tamanho da classe Campo Limpo novamente impediu que os erros fossem verificados nos gráficos da Figura 17.

No geral, percebe-se que as imagens classificadas usando os algoritmos por regiões apresentaram erros de comissão e omissão bem inferiores às imagens que usaram os algoritmos classificadores por *pixels*, indicando que as imagens de alta resolução, por apresentarem uma alta variabilidade espectral entre os *pixels*, são melhores classificadas quando passam por um processo de segmentação antes da classificação final.

4.2 – A significância entre as diferentes quantidades amostrais

A análise do Quadro 7 mostra que os maiores valores de Kappa, para os algoritmos de classificação por *pixels*, ou seja, Máxima Verossimilhança, Distância Euclidiana Mínima e Redes Neurais foram: 0,667; 0,6449 e 0,6826, respectivamente. Para os algoritmos de classificação por regiões, que incluem os algoritmos Bhattacharya, Iseseg e Histograma, os maiores valores de kappa foram, respectivamente, 0,8219; 0,8788 e 0,8433. Os valores de Kappa para os três primeiros algoritmos citados neste parágrafo ficaram no intervalo considerado razoável (Quadro 3), de acordo com a tabela do índice de Kappa fornecido por CONGALTON e GREEN (1998), enquanto os valores de Kappa para os três últimos foram considerados excelentes, de acordo com o mesmo critério.

Na Figura 18 visualiza-se a distribuição dos valores de Kappa, considerando os diferentes tamanhos de amostras, para todos os algoritmos de classificação utilizados neste estudo. A segmentação, para a metodologia utilizada, foi fator preponderante para os resultados encontrados, melhorando tanto a qualidade visual da imagem classificada como também a acurácia do produto final.

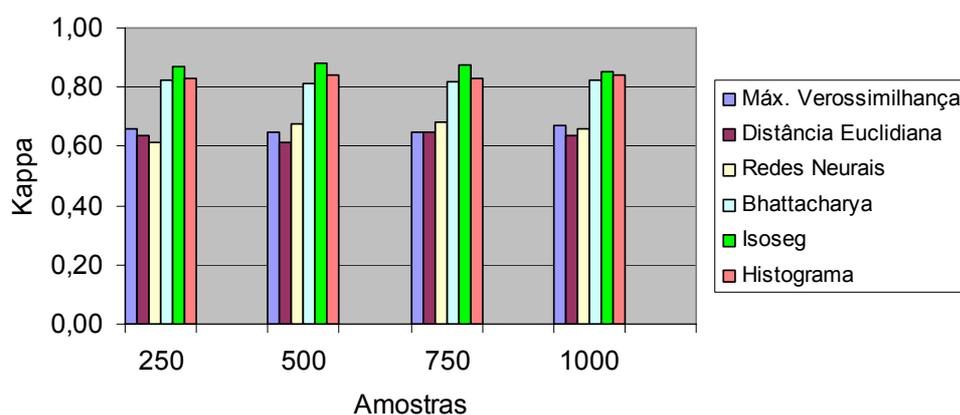


Figura 18 - Valores calculados de Kappa para cada quantidade amostral.

Quadro 7 - Valores dos índices de kappa e suas respectivas variâncias, considerando os diferentes tamanhos de amostras.

ALGORITMOS	ÍNDICE KAPPA E VARIÂNCIA POR QUANTIDADE DE AMOSTRA							
	250		500		750		1000	
	Kappa	Variância	Kappa	Variância	Kappa	Variância	Kappa	Variância
Máxima Verossimilhança	0,6608	0,001260	0,6486	0,000650	0,6488	0,000430	0,6677	0,000311
Distância Euclidiana	0,6355	0,001313	0,6164	0,000673	0,6449	0,000425	0,6391	0,000324
Redes Neurais	0,6145	0,001384	0,6759	0,000614	0,6826	0,000401	0,6569	0,000313
Bhattacharya	0,8217	0,000830	0,8112	0,000427	0,8210	0,000270	0,8219	0,000201
Isoseg	0,8707	0,000626	0,8788	0,000289	0,8732	0,000198	0,8543	0,000168
Histograma	0,8307	0,000776	0,8433	0,000358	0,8299	0,000253	0,8430	0,000177

O teste estatístico Z foi usado para verificar se existe diferença estatística significativa entre duas variâncias, para os diferentes tamanhos de amostras e algoritmos de classificação utilizados. Os resultados podem ser visualizados no Apêndice II (Quadros 1 a 3), sendo comparado no Quadro 1 os índices Kappa resultantes dos mapas temáticos classificados pelos algoritmos por *pixels*. No Quadro 2 foram comparados os índices Kappa das classificações por Regiões e no Quadro 3 os índices Kappa dos algoritmos classificadores por *Pixels versus* os algoritmos por Regiões.

Nos Quadros 1 e 2, pode ser constatado que o aumento do tamanho da amostra teste, para a avaliação da imagens classificadas, pelo menos neste trabalho, não trouxe melhoras significativas para os valores do índice Kappa. Ao se analisar o Quadro 3, a comparação dos resultados entre os algoritmos por *Pixel* e por Região apresentaram diferenças estatisticamente significativas, com os maiores valores para os algoritmos por Regiões.

Dentre os algoritmos por Regiões, o algoritmo de classificação Ioseg foi o que apresentou os maiores índices kappa (Figura 18), bem como os menores erros de omissão e comissão (Figura 16).

5 – CONCLUSÃO

- As classificações geradas pelos algoritmos que segmentam as imagens, antes de classificá-las, eliminando os pixels isolados, geraram imagens temáticas com qualidades visuais superiores àquelas geradas pelos algoritmos por *pixels*;
- Dentre as imagens classificadas e avaliadas pelos erros de comissão, a classe de café foi a que apresentou melhor resultado;
- As feições terrestres que possuíam uma das bandas da imagem com resposta espectral diferenciada das demais foram melhores classificadas;
- Não houve diferença estatística significativa, a 95% de probabilidade, entre os índices Kappa, para os algoritmos de classificação, ao se variar o tamanho das amostras de testes;
- Houve diferença estatística significativa, a 95% de probabilidade, entre os índices Kappa gerados, ao se comparar os algoritmos de classificação por *pixels* e por regiões.
- Os algoritmos de classificação por *pixels* apresentaram valores de índice Kappa considerados razoáveis, de acordo com a metodologia utilizada, enquanto que os algoritmos por regiões apresentaram resultados de Kappa considerados excelentes.
- O algoritmo por regiões Iseseg obteve os melhores desempenhos nas classificações automáticas;
- O aumento do tamanho das amostras de testes, para avaliação das imagens classificadas, não demonstrou, estatisticamente, melhorias no cálculo do índice de exatidão kappa.

6 – RECOMENDAÇÕES

As melhorias em processos e procedimentos devem fazer parte do aprendizado, que não deve se limitar a uma única explicação, buscando sempre novos argumentos para compreender os fenômenos naturais. Visando subsidiar estudos futuros, seguem algumas recomendações que foram observadas ao final do trabalho, sugerindo algumas idéias para sua continuidade:

- Repetir a metodologia utilizando um número maior de amostras de treinamento, visando um melhor desempenho dos algoritmos classificadores por *pixels*;
- Selecionar uma área maior para estudar a metodologia, contendo, inclusive, um número maior de classes temáticas;
- Variar a técnica de amostragem para validação das imagens temáticas, incluindo, além da utilizada (aleatória por *pixels*), uma amostragem por regiões, a fim de testar se há diferença estatística significativa entre os dois tipos de classificadores utilizados;
- Testar o algoritmo de Redes Neurais utilizando uma taxa de aprendizagem menor, sendo recomendado 0,2;
- Passar um filtro de médias nas imagens temáticas geradas pelos algoritmos classificadores por *pixels*, antes de validar a classificação.

REFERÊNCIAS BIBLIOGRÁFICAS

BINS, L. S.; FONSECA, L. M. G.; ERTHAL, G. J.; MITSUO, F. Satellite Imagery Segmentation: a region growing approach. In: Simpósio Brasileiro de Sensoriamento Remoto, 8. (SBSR), 14-19 abr. 1996, Salvador. **Anais...** São José dos Campos: INPE, 2004. Artigos, p. 677-680. CD-ROM. ISBN 85-17-00014-5.

BINS, L. S. Correspondência eletrônica enviada pelo autor em 05 de julho de 2006. Contato: leonardo@dpi.inpe.br (Tecnologista Sênior III – Instituto Nacional de Pesquisas Espaciais / Divisão de Processamento de Imagens).

BORGES, F. S. P.; ROTUNNO FILHO, O. C.; WILSON-JR, G. Sensoriamento remoto e geoestatística aplicados à estimativa de batimetria em reservatórios do setor elétrico brasileiro. In: Simpósio Brasileiro de Sensoriamento Remoto, 12. (SBSR), 16-21 abr. 2005, Goiânia. **Anais...** São José dos Campos: INPE, 2005. Artigos, p. 2471-2478. CD-ROM. ISBN 85-17-00018-8.

BRAGA, A. P.; CARVALHO, A. P. L. F.; LUDERMIR, T. B. **Redes Neurais Artificiais: Teoria e aplicação**. Rio de Janeiro, RJ. RTC, 2000. 262 p.

CAMPBELL, J. B. **Introduction to remote sensing**. New York, The Guilford Press, 1987. 551p.

CONGALTON, R.G.; A comparison of sampling schemes used in generating error matrices for assessing the accuracy of maps generated from remotely sensed data. **Photogrammetric Engineering and Remote Sensing**, Bethesda, v. 54, p.593-600, 1998. New York: Lewis Publishers, 1998. 137 p.

CONGALTON, R.G.; GREEN, K. **Assessing the accuracy of remotely sensed data: principles and practices**. New York: Lewis Publishers, 1998. 137 p.

COUTINHO, A. C.; MIRANDA, E. E.; MIRANDO, J. R. Mapeamento da Superfície Terrestre Através da Utilização do Método de Segmentação por Crescimento de Regiões e Classificação Supervisionada de Imagens de Satélite. In: Simpósio Brasileiro de Sensoriamento Remoto, 9. (SBSR), 11-18 set. 1998, Santos. **Anais...** São José dos Campos: INPE, 2004. P. 525-535. CD-ROM. ISBN 85-17-00015-3.

CURRAN, P. J. **Principles of remote sensing**. London: Longman Group Limited, 1992. 282 p.

FITZPATRICK-LINS, K. **The accuracy of selected land use and land cover maps at scales of 1:250.000 and 1:100.000**. Journal of research, U. S. Geological Survey, vol. 6, p.169-173, 1980.

FLORENZANO, T. G. **Imagens de satélite para estudos ambientais**. São Paulo: Oficina de Textos, 2002.

GONG, P.; HOWART, PL J. An assessment of some factors influencing multispectral land-cover classification. **Photogrammetric Engineering and Remote Sensing**, Bethesda. V.56, n.5, p.597-603, 1990.

IBGE. **Unidades da Federação**. Disponível em:
<<http://www.ibge.com.br/cidadesat/default.php>> Acesso em 08 de julho de 2006.

INPE. SPRING: **Sistema de Processamento de Informações Georreferenciadas**. Manual de ajuda, 2004.

JENSEN, J.R. **Introductory digital image processing: A remote sensing perspective**. New Jersey: Prentice-Hall, 1996. 316 p.

KURASZ, G; OLIVEIRA, Y. M. M.; ROSOT, N. C.; ROSOT, M. A. D. In: Simpósio Brasileiro de Sensoriamento Remoto, 12. (SBSR), 16-21 abr. 2005, Goiânia. **Anais...** São José dos Campos: INPE, 2005. Artigos, p. 1585-1592. CD-ROM. ISBN 85-17-00018-8.

LILLESAND, T. M.; KIEFER, R. W.; CHIPMAN, J. W. **Remote sensing and image interpretation**. New York: John Miley & Sons. 5. ed. 2004. 763 p.

MATHER, P.; TSO, BRANDT. **Classification Methods for remotely sensed data**. Taylor e Francis, 322p. 2001.

MEDEIROS, C. N. de; PETTA, R. A. Exploração de imagens de satélite de alta resolução visando o mapeamento do uso e ocupação do solo. In: Simpósio Brasileiro de Sensoriamento Remoto, 12. (SBSR), 16-21 abr. 2005, Goiânia. **Anais...** São José dos Campos: INPE, 2005. Artigos, p. 2709-2716. CD-ROM. ISBN 85-17-00018-8.

MOREIRA, M.A. **Fundamentos do sensoriamento remoto e metodologias de aplicação**. 2.ed. Viçosa: UFV, 2005. 307P.

NASCIMENTO, M. C.; SOARES, V. P.; RIBEIRO, C. A. A. S.; SILVA, E. **Mapeamento dos fragmentos de vegetação florestal nativa da bacia hidrográfica do Rio Alegre, Espírito Santo, a partir de imagens do satélite IKONOS II**. Revista *Árvore*, Viçosa-MG, v.30, n.3, p.389-398, 2006.

NOVO, E. M. L. M. **Sensoriamento remoto: princípios e aplicações**, Ed. Edgard Blucher LTDA, 307 p, 1992.

RAMIREZ, G. M. *et al.* Comparison between Ikonos-II and Landsat/ETM+ satellites data in the study of coffee areas. **Pesq. agropec. bras.**, Brasília, v. 41, n. 4, 2006. Disponível em:
<http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-

204X2006000400017&lng=en&nrm=iso>. Acesso em: 10 Feb 2007. Pré-publicação. doi: 10.1590/S0100-204X2006000400017

RIBEIRO, R. M. P. **Avaliação de métodos de classificação de imagens IKONOS II para o mapeamento da cobertura terrestre**. 2003. 53 f. Tese (Mestrado em Ciência Florestal) – Universidade Federal de Viçosa, Viçosa, 2003.

RUMELHART, D. *et. al.* **Learning Internal representation by error propagation parallel distributed processing: exploration in microstructure of cognition, 1: foundation**. MIT Press, Cambridge, Massachusetts, 1986. P. 318-362.

SPACE IMAGING. **IKONOS**. Disponível em:
<<http://www.spaceimaging.com/products/ikonos>> Acesso em 15 de julho de 2006.

VIEIRA, C.A.O. **Accuracy of remotely sensing classification of agricultural crops: a comparative study**. 2000. 323f. Thesis (Ph.D.) – University of Nottingham, Nottingham, 2000.

VILELA, M. F. **Integração de técnicas de geoprocessamento e levantamento participativo de informações sócio-ambientais: um subsídio para a reforma agrária**. 2002. 135 f. Tese (Doutorado em Ciência Florestal) – Universidade Federal de Viçosa, Viçosa, 2002.

APÊNDICE I

Matrizes de Erros geradas, para as imagens temáticas classificadas pelos algoritmos da Máxima Verossimilhança, Distância Euclidiana Mínima, Redes Neurais, Bhattacharya, Isepeg e Histograma, com base em visitas a campo e conhecimento da área.

Quadro 1 – Matriz de Erros do classificador Máxima Verossimilhança utilizando 250 amostras.

	CAFE	SOEX	MACI	CASU	BRAQ	PAST	CALI	Total	Erro Com
CAFE	13	0	0	0	0	0	0	13	0,000
SOEX	0	18	0	0	0	0	0	18	0,000
MACI	0	0	26	2	0	0	0	28	0,071
CASU	0	0	3	25	2	0	26	56	0,554
BRAQ	1	5	0	0	21	12	5	44	0,523
PAST	0	2	0	0	0	11	0	13	0,154
CALI	0	5	0	0	2	1	58	66	0,121
Total	14	30	29	27	25	24	89	238	
Erro Omis	0,071	0,400	0,103	0,074	0,160	0,542	0,348		

Índice Kappa = 0,6608

Variância = 0,001260

Quadro 2 – Matriz de Erros do classificador Máxima Verossimilhança utilizando 500 amostras.

	CAFE	SOEX	MACI	CASU	BRAQ	PAST	CALI	Total	Erro Com
CAFE	25	0	0	0	0	0	0	25	0,000
SOEX	0	28	0	1	0	0	0	29	0,035
MACI	0	0	57	4	0	2	0	63	0,095
CASU	0	0	8	43	1	0	32	84	0,488
BRAQ	1	14	2	3	45	27	10	102	0,559
PAST	0	0	0	0	0	18	1	19	0,053
CALI	0	15	1	1	4	9	126	156	0,192
Total	26	57	68	52	50	56	169	478	
Erro Omis	0,039	0,509	0,162	0,173	0,100	0,679	0,254		

Índice Kappa = 0,6486

Variância = 0,000650

Nome	Nomenclatura
Café -----	CAFE
Solo Exposto -----	SLEX
Mata Ciliar -----	MACI
Campo Sujo -----	CASU
Braquiária -----	BRQU
Pastagem -----	PSTO
Campo Limpo -----	CALI

Quadro 3 – Matriz de Erros do classificador Máxima Verossimilhança utilizando 750 amostras.

	CAFÉ	SOEX	MACI	CASU	BRAQ	PAST	CALI	Total	Erro Com
CAFE	45	0	0	0	0	0	0	45	0,000
SOEX	0	40	0	0	0	0	0	40	0,000
MACI	1	0	84	9	0	0	0	94	0,106
CASU	0	1	4	63	1	0	59	128	0,508
BRAQ	3	17	1	12	67	35	11	146	0,541
PAST	0	3	0	0	1	25	0	29	0,138
CALI	0	23	0	4	4	17	201	249	0,193
Total	49	84	89	88	73	77	271	731	
Erro Omis	0,082	0,524	0,056	0,284	0,082	0,675	0,258		

Índice Kappa = 0,6488

Variância = 0,000430

Quadro 4 – Matriz de Erros do classificador Máxima Verossimilhança utilizando 1000 amostras.

	CAFE	SOEX	MACI	CASU	BRAQ	PAST	CALI	Total	Erro Com
CAFE	59	0	1	0	0	0	1	61	0,033
SOEX	0	58	0	0	0	0	0	58	0,000
MACI	3	0	107	16	1	1	0	128	0,164
CASU	0	1	8	96	2	0	68	175	0,451
BRAQ	0	21	2	7	96	39	21	186	0,484
PAST	0	4	0	0	0	38	3	45	0,156
CALI	0	21	1	6	6	27	256	317	0,192
Total	62	105	119	125	105	105	349	970	
Erro Omis	0,048	0,448	0,101	0,232	0,086	0,638	0,267		

Índice Kappa = 0,6677

Variância = 0,000311

Nome	Nomenclatura
Café -----	CAFE
Solo Exposto -----	SLEX
Mata Ciliar -----	MACI
Campo Sujo -----	CASU
Braquiária -----	BRQU
Pastagem -----	PSTO
Campo Limpo -----	CALI

Quadro 5 – Matriz de Erros do classificador Distância Euclidiana Mínima utilizando 250 amostras.

	CAFE	SOEX	MACI	CASU	BRAQ	PAST	CALI	Total	Erro Com
CAFE	14	0	2	0	0	0	0	16	0,125
SOEX	0	25	0	0	0	0	0	25	0,000
MACI	0	0	23	3	0	0	0	26	0,115
CASU	0	0	4	24	4	0	38	70	0,657
BRAQ	0	1	0	0	13	2	0	16	0,188
PAST	0	3	0	0	0	17	0	20	0,150
CALI	0	1	0	0	8	5	51	65	0,215
Total	14	30	29	27	25	24	89	238	
Erro Omis	0,000	0,167	0,207	0,111	0,480	0,292	0,427		

Índice Kappa = 0,6355

Variância = 0,001313

Quadro 6 – Matriz de Erros do classificador Distância Euclidiana Mínima utilizando 500 amostras.

	CAFE	SOEX	MACI	CASU	BRAQ	PAST	CALI	Total	Erro Com
CAFE	25	0	2	0	0	0	0	27	0,074
SOEX	0	49	0	1	0	2	4	56	0,125
MACI	0	0	56	5	0	2	0	63	0,111
CASU	0	0	7	43	5	1	62	118	0,636
BRAQ	1	3	2	3	22	2	2	35	0,371
PAST	0	2	0	0	6	33	1	42	0,214
CALI	0	3	1	0	17	16	100	137	0,270
Total	26	57	68	52	50	56	169	478	
Erro Omis	0,039	0,140	0,177	0,173	0,560	0,411	0,408		

Índice Kappa = 0,6164

Variância = 0,000673

Nome	Nomenclatura
Café -----	CAFE
Solo Exposto -----	SLEX
Mata Ciliar -----	MACI
Campo Sujo -----	CASU
Braquiária -----	BRQU
Pastagem -----	PSTO
Campo Limpo -----	CALI

Quadro 7 – Matriz de Erros do classificador Distância Euclidiana Mínima utilizando 750 amostras.

	CAFE	SOEX	MACI	CASU	BRAQ	PAST	CALI	Total	Erro Com
CAFE	46	0	4	2	1	0	0	53	0,132
SOEX	0	72	0	0	0	5	1	78	0,077
MACI	0	0	80	8	0	0	0	88	0,091
CASU	0	0	5	69	6	0	105	185	0,627
BRAQ	2	1	0	5	40	0	1	49	0,184
PAST	1	3	0	0	8	50	1	63	0,206
CALI	0	8	0	4	18	22	163	215	0,242
Total	49	84	89	88	73	77	271	731	
Erro Omis	0,061	0,143	0,101	0,216	0,452	0,351	0,399		

Índice Kappa = 0,6449

Variância = 0,000425

Quadro 8 – Matriz de Erros do classificador Distância Euclidiana Mínima usando 1000 amostras.

	CAFE	SOEX	MACI	CASU	BRAQ	PAST	CALI	Total	Erro Com
CAFE	61	0	9	1	0	0	1	72	0,153
SOEX	0	88	0	0	0	7	4	99	0,111
MACI	0	0	100	12	1	1	0	114	0,123
CASU	0	1	9	102	9	1	116	238	0,571
BRAQ	1	1	0	4	58	3	2	69	0,159
PAST	0	7	0	0	9	58	7	81	0,284
CALI	0	8	1	6	28	35	219	297	0,263
Total	62	105	119	125	105	105	349	970	
Erro Omis	0,016	0,162	0,160	0,184	0,448	0,448	0,373		

Índice Kappa = 0,6391

Variância = 0,000324

Nome	Nomenclatura
Café -----	CAFE
Solo Exposto -----	SLEX
Mata Ciliar -----	MACI
Campo Sujo -----	CASU
Braquiária -----	BRQU
Pastagem -----	PSTO
Campo Limpo -----	CALI

Quadro 9 – Matriz de Erros do classificador Redes Neurais utilizando 250 amostras.

	CAFÉ	SOEX	MACI	CASU	BRAQ	PAST	CALI	Total	Erro Com
CAFE	13	0	4	1	0	0	0	18	0,278
SOEX	0	24	0	0	0	0	1	25	0,040
MACI	0	0	21	12	0	0	1	34	0,382
CASU	0	0	2	8	1	0	16	27	0,704
BRAQ	1	1	1	0	11	2	0	16	0,313
PAST	0	5	0	0	10	21	5	41	0,488
CALI	0	0	1	6	3	1	66	77	0,143
Total	14	30	29	27	25	24	89	238	
Erro Omiss	0,071	0,200	0,276	0,704	0,560	0,125	0,258		

Índice Kappa = 0,6145

Variância = 0,001384

Quadro 10 – Matriz de Erros do classificador Redes Neurais utilizando 500 amostras.

	CAFE	SOEX	MACI	CASU	BRAQ	PAST	CALI	Total	Erro Com
CAFE	25	0	9	0	0	0	0	34	0,265
SOEX	0	47	0	1	0	0	10	58	0,190
MACI	0	0	52	15	0	1	0	68	0,235
CASU	0	0	4	27	1	0	14	46	0,413
BRAQ	1	2	2	4	20	2	0	31	0,355
PAST	0	8	0	0	28	48	12	96	0,500
CALI	0	0	1	5	1	5	133	145	0,083
Total	26	57	68	52	50	56	169	478	
Erro Omiss	0,039	0,175	0,235	0,481	0,600	0,143	0,213		

Índice Kappa = 0,6759

Variância = 0,000614

Nome	Nomenclatura
Café -----	CAFE
Solo Exposto -----	SLEX
Mata Ciliar -----	MACI
Campo Sujo -----	CASU
Braquiária -----	BRQU
Pastagem -----	PSTO
Campo Limpo -----	CALI

Quadro 11 – Matriz de Erros do classificador Redes Neurais utilizando 750 amostras.

	CAFE	SOEX	MACI	CASU	BRAQ	PAST	CALI	Total	Erro Com
CAFE	46	0	11	2	1	0	0	60	0,233
SOEX	0	69	0	0	0	1	8	78	0,115
MACI	0	0	75	18	1	0	0	94	0,202
CASU	0	0	1	39	0	0	34	74	0,473
BRAQ	2	0	1	9	34	0	2	48	0,292
PAST	1	13	0	5	35	68	14	136	0,500
CALI	0	2	1	15	2	8	213	241	0,116
Total	49	84	89	88	73	77	271	731	
Erro Omis	0,061	0,179	0,157	0,557	0,534	0,117	0,214		

Índice Kappa = 0,6826

Variância = 0,000401

Quadro 12 – Matriz de Erros do classificador Redes Neurais utilizando 1000 amostras.

	CAFÉ	SOEX	MACI	CASU	BRAQ	PAST	CALI	Total	Erro Com
CAFE	61	0	15	3	1	0	1	81	0,247
SOEX	0	87	0	1	0	1	10	99	0,121
MACI	0	0	97	40	1	0	0	138	0,297
CASU	0	0	4	53	1	0	35	93	0,430
BRAQ	1	2	0	9	46	3	3	64	0,281
PAST	0	13	2	2	54	91	36	198	0,540
CALI	0	3	1	17	2	10	264	297	0,111
Total	62	105	119	125	105	105	349	970	
Erro Omis	0,016	0,171	0,185	0,576	0,562	0,133	0,244		

Índice Kappa = 0,6569

Variância = 0,000313

Nome	Nomenclatura
Café -----	CAFE
Solo Exposto -----	SLEX
Mata Ciliar -----	MACI
Campo Sujo -----	CASU
Braquiária -----	BRQU
Pastagem -----	PSTO
Campo Limpo -----	CALI

Quadro 13 – Matriz de Erros do classificador Bhattacharya utilizando 250 amostras.

	CAFE	SOEX	MACI	CASU	BRAQ	PAST	CALI	Total	Erro Com
CAFE	14	0	0	0	0	0	0	14	0,000
SOEX	0	25	0	0	0	0	1	26	0,039
MACI	0	0	26	5	0	1	0	32	0,188
CASU	0	1	2	11	0	0	2	16	0,313
BRAQ	0	1	0	1	25	0	0	27	0,074
PAST	0	3	0	0	0	18	0	21	0,143
CALI	0	0	1	10	0	5	86	102	0,157
Total	14	30	29	27	25	24	89	238	
Erro Omiss	0,000	0,167	0,103	0,593	0,000	0,250	0,034		

Índice Kappa = 0,8217

Variância = 0,000830

Quadro 14 – Matriz de Erros do classificador Bhattacharya utilizando 500 amostras.

	CAFE	SOEX	MACI	CASU	BRAQ	PAST	CALI	Total	Erro Com
CAFE	26	0	0	0	0	0	0	26	0,000
SOEX	0	48	2	0	0	0	1	51	0,059
MACI	0	0	59	4	0	0	0	63	0,064
CASU	0	0	5	24	0	3	3	35	0,314
BRAQ	0	1	1	4	49	0	3	58	0,155
PAST	0	7	0	1	0	39	0	47	0,170
CALI	0	1	1	19	1	14	162	198	0,182
Total	26	57	68	52	50	56	169	478	
Erro Omis	0,000	0,158	0,132	0,539	0,020	0,304	0,041		

Índice Kappa = 0,8112

Variância = 0,000427

Nome	Nomenclatura
Café -----	CAFE
Solo Exposto -----	SLEX
Mata Ciliar -----	MACI
Campo Sujo -----	CASU
Braquiária -----	BRQU
Pastagem -----	PSTO
Campo Limpo -----	CALI

Quadro 15 – Matriz de Erros do classificador Bhattacharya utilizando 750 amostras.

	CAFE	SOEX	MACI	CASU	BRAQ	PAST	CALI	Total	Erro Com
CAFE	48	0	0	1	0	0	0	49	0,020
SOEX	0	74	0	0	0	0	3	77	0,039
MACI	0	0	85	6	1	0	0	92	0,076
CASU	0	0	3	45	1	2	6	57	0,211
BRAQ	0	0	0	6	69	0	5	80	0,138
PAST	1	10	0	0	0	51	0	62	0,177
CALI	0	0	1	30	2	24	257	314	0,182
Total	49	84	89	88	73	77	271	731	
Erro Omiss	0,020	0,119	0,045	0,489	0,055	0,338	0,052		

Índice Kappa = 0,8210

Variância = 0,000270

Quadro 16 – Matriz de Erros do classificador Bhattacharya utilizando 1000 amostras.

	CAFE	SOEX	MACI	CASU	BRAQ	PAST	CALI	Total	Erro Com
CAFE	62	0	0	0	0	0	1	63	0,016
SOEX	0	94	0	5	0	0	3	102	0,078
MACI	0	0	110	10	1	0	0	121	0,091
CASU	0	0	3	66	1	4	11	85	0,224
BRAQ	0	1	0	3	101	0	6	111	0,090
PAST	0	9	1	0	1	75	2	88	0,148
CALI	0	1	5	41	1	26	326	400	0,185
Total	62	105	119	125	105	105	349	970	
Erro Omiss	0,000	0,105	0,076	0,472	0,038	0,286	0,066		

Índice Kappa = 0,8219

Variância = 0,000201

Nome	Nomenclatura
Café -----	CAFE
Solo Exposto -----	SLEX
Mata Ciliar -----	MACI
Campo Sujo -----	CASU
Braquiária -----	BRQU
Pastagem -----	PSTO
Campo Limpo -----	CALI

Quadro 17 – Matriz de Erros do classificador Iseseg utilizando 250 amostras.

	CAFÉ	SOEX	MACI	CASU	BRAQ	PAST	CALI	Total	Erro Com
CAFE	14	0	0	0	0	0	0	14	0,000
SOEX	0	27	0	0	0	0	0	27	0,000
MACI	0	0	25	1	0	0	0	26	0,039
CASU	0	0	2	17	0	0	3	22	0,227
BRAQ	0	1	0	1	25	1	0	28	0,107
PAST	0	1	0	0	0	20	0	21	0,048
CALI	0	1	2	8	0	3	86	100	0,140
Total	14	30	29	27	25	24	89	238	
Erro Omis	0,000	0,100	0,138	0,370	0,000	0,167	0,034		

Índice Kappa = 0,8707

Variância = 0,000626

Quadro 18 – Matriz de Erros do classificador Iseseg utilizando 500 amostras.

	CAFE	SOEX	MACI	CASU	BRAQ	PAST	CALI	Total	Erro Com
CAFE	26	0	0	0	0	0	0	26	0,000
SOEX	0	52	2	1	0	0	0	55	0,055
MACI	0	0	59	2	0	0	0	61	0,033
CASU	0	0	3	36	0	0	3	42	0,143
BRAQ	0	2	1	3	48	0	3	57	0,158
PAST	0	3	0	0	2	48	0	53	0,094
CALI	0	0	3	10	0	8	163	184	0,114
Total	26	57	68	52	50	56	169	478	
Erro Omis	0,000	0,088	0,132	0,308	0,040	0,143	0,036		

Índice Kappa = 0,8788

Variância = 0,000289

Nome	Nomenclatura
Café -----	CAFE
Solo Exposto -----	SLEX
Mata Ciliar -----	MACI
Campo Sujo -----	CASU
Braquiária -----	BRQU
Pastagem -----	PSTO
Campo Limpo -----	CALI

Quadro 19 – Matriz de Erros do classificador Iseseg utilizando 750 amostras.

	CAFE	SOEX	MACI	CASU	BRAQ	PAST	CALI	Total	Erro Com
CAFE	48	0	0	1	0	0	0	49	0,0204
SOEX	0	80	0	1	0	0	4	85	0,059
MACI	0	0	83	2	1	0	0	86	0,035
CASU	0	0	5	56	2	0	3	66	0,152
BRAQ	0	2	0	5	68	1	5	81	0,161
PAST	1	2	0	0	2	66	2	73	0,096
CALI	0	0	1	23	0	10	257	291	0,117
Total	49	84	89	88	73	77	271	731	
Erro Omis	0,020	0,048	0,067	0,364	0,069	0,143	0,052		

Índice Kappa = 0,8732

Variância = 0,000198

Quadro 20 – Matriz de Erros do classificador Iseseg utilizando 1000 amostras.

	CAFÉ	SOEX	MACI	CASU	BRAQ	PAST	CALI	Total	Erro Com
CAFE	62	0	0	0	0	0	1	63	0,016
SOEX	0	101	0	4	0	0	1	106	0,047
MACI	0	0	109	6	1	0	1	117	0,068
CASU	0	0	4	77	1	1	6	89	0,135
BRAQ	0	2	0	1	98	1	8	110	0,109
PAST	0	2	1	0	5	88	9	105	0,162
CALI	0	0	5	37	0	15	323	380	0,150
Total	62	105	119	125	105	105	349	970	
Erro Omis	0,000	0,038	0,084	0,384	0,067	0,162	0,075		

Índice Kappa = 0,8543

Variância = 0,000168

Nome

Café -----
 Solo Exposto -----
 Mata Ciliar -----
 Campo Sujo -----
 Braquiária -----
 Pastagem -----
 Campo Limpo -----

Nomenclatura

CAFE
 SLEX
 MACI
 CASU
 BRQU
 PSTO
 CALI

Quadro 21 – Matriz de Erros do classificador Histograma utilizando 250 amostras.

	CAFÉ	SOEX	MACI	CASU	BRAQ	PAST	CALI	Total	Erro Com
CAFE	13	0	0	0	0	0	0	13	0,000
SOEX	0	27	0	0	0	0	0	27	0,000
MACI	0	0	25	3	0	0	1	29	0,138
CASU	0	0	3	19	1	0	10	33	0,424
BRAQ	1	1	0	1	24	1	0	28	0,143
PAST	0	1	0	0	0	20	0	21	0,048
CALI	0	1	1	4	0	3	78	87	0,103
Total	14	30	29	27	25	24	89	238	
Erro Omiss	0,071	0,100	0,138	0,296	0,040	0,167	0,124		

Índice Kappa = 0,8307

Variância = 0,000776

Quadro 22 – Matriz de Erros do classificador Histograma utilizando 500 amostras.

	CAFE	SOEX	MACI	CASU	BRAQ	PAST	CALI	Total	Erro Com
CAFE	25	1	0	0	0	0	0	26	0,039
SOEX	0	52	1	1	0	0	1	55	0,055
MACI	0	0	57	5	0	0	0	62	0,081
CASU	0	0	6	37	0	1	13	57	0,351
BRAQ	1	1	1	3	47	0	3	56	0,161
PAST	0	2	0	0	2	48	0	52	0,077
CALI	0	1	3	6	1	7	152	170	0,106
Total	26	57	68	52	50	56	169	478	
Erro Omis	0,039	0,088	0,162	0,289	0,060	0,143	0,101		

Índice Kappa = 0,8433

Variância = 0,000358

Nome	Nomenclatura
Café -----	CAFE
Solo Exposto -----	SLEX
Mata Ciliar -----	MACI
Campo Sujo -----	CASU
Braquiária -----	BRQU
Pastagem -----	PSTO
Campo Limpo -----	CALI

Quadro 23 – Matriz de Erros do classificador Histograma utilizando 750 amostras.

	CAFE	SOEX	MACI	CASU	BRAQ	PAST	CALI	Total	Erro Com
CAFE	47	0	2	1	1	0	0	51	0,079
SOEX	0	78	0	0	0	0	3	81	0,037
MACI	0	0	81	4	0	0	1	86	0,058
CASU	0	0	5	58	3	0	25	91	0,363
BRAQ	1	1	0	6	66	0	5	79	0,165
PAST	1	3	0	0	3	67	2	76	0,118
CALI	0	2	1	19	0	10	235	267	0,120
Total	49	84	89	88	73	77	271	731	
Erro Omiss	0,041	0,071	0,090	0,341	0,096	0,130	0,133		

Índice Kappa = 0,8299 Variância = 0,000253

Quadro 24 – Matriz de Erros do classificador Histograma utilizando 1000 amostras.

	CAFÉ	SOEX	MACI	CASU	BRAQ	PAST	CALI	Total	Erro Com
CAFE	61	0	2	0	0	0	1	64	0,047
SOEX	0	99	0	3	0	0	3	105	0,057
MACI	0	0	106	8	1	0	1	116	0,086
CASU	0	0	6	92	2	0	24	124	0,258
BRAQ	1	1	0	3	95	0	6	106	0,104
PAST	0	4	1	0	7	90	9	111	0,189
CALI	0	1	4	19	0	15	305	344	0,113
Total	62	105	119	125	105	105	349	970	
Erro Omiss	0,016	0,057	0,109	0,2640	0,095	0,143	0,126		

Índice Kappa = 0,8430 Variância = 0,000177

Nome	Nomenclatura
Café -----	CAFE
Solo Exposto -----	SLEX
Mata Ciliar -----	MACI
Campo Sujo -----	CASU
Braquiária -----	BRQU
Pastagem -----	PSTO
Campo Limpo -----	CALI

APÊNDICE II

Matriz comparativa dos resultados de Kappa e suas Variâncias, cujos resultados indicam se existem diferenças estatisticamente significativas.

Os algoritmos classificadores foram comparados da seguinte forma:

- Pixel x Pixel → Quadro 1
- Região x Região → Quadro 2
- Pixel x Região → Quadro 3

Seguem as nomenclaturas dos algoritmos, utilizadas nos Quadros 1, 2 e 3 deste Apêndice:

- MV – Máxima Verossimilhança
- DE – Distância Euclidiana Mínima
- RN – Redes Neurais
- BT – Bhattacharya
- IS – Isoseg
- HI – Histograma

Quadro 1 - Comparação dos resultados de Kappa e suas variâncias entre os algoritmos classificadores por *Pixel*.

Classe	MV250	MV500	MV750	MV1000	DE250	DE500	DE750	DE1000	RN250	RN500	RN750	RN1000
MV250	18.616											
MV500	0.279	25.440										
MV750	0.292	0.006	31.288									
MV1000	0.174	0.616	0.694	37.862								
DE250	0.499	0.296	0.319	0.799	17.538							
DE500	1.010	0.885	0.976	1.635	0.429	23.760						
DE750	0.387	0.113	0.133	0.840	0.225	0.860	31.282					
DE1000	0.545	0.304	0.353	1.135	0.089	0.719	0.212	35.506				
RN250	0.900	0.756	0.805	1.292	0.404	0.042	0.715	0.595	16.518			
RN500	0.349	0.768	0.839	0.270	0.920	1.659	0.962	1.202	1.374	27.277		
RN750	0.535	1.049	1.173	0.558	1.138	2.020	1.312	1.616	1.612	0.210	34.087	
RN1000	0.098	0.267	0.297	0.432	0.531	1.290	0.442	0.705	1.029	0.624	0.962	37.130

Quadro 2 - Comparação dos resultados de Kappa e suas variâncias entre os algoritmos classificadores por Regiões.

Classe	BT250	BT500	BT750	BT1000	IS250	IS500	IS750	IS1000	HI250	HI500	HI750	HI1000
BT250	28.522											
BT500	0.296	39.257										
BT750	0.021	0.371	49.964									
BT1000	0.006	0.427	0.041	57.972								
IS250	1.284	1.834	1.660	1.697	34.800							
IS500	1.707	2.526	2.445	2.570	0.268	51.694						
IS750	1.606	2.480	2.413	2.568	0.087	0.254	62.056					
IS1000	1.032	1.767	1.591	1.687	0.582	1.146	0.988	65.911				
HI250	0.225	0.562	0.300	0.282	1.068	1.474	1.362	0.768	29.820			
HI500	0.627	1.146	0.890	0.905	0.873	1.396	1.268	0.480	0.374	44.570		
HI750	0.249	0.717	0.389	0.375	1.376	2.100	2.039	1.189	0.025	0.542	52.175	
HI1000	0.671	1.294	1.041	1.085	0.978	1.658	1.560	0.608	0.398	0.013	0.632	63.364

Quadro 3 - Comparação dos resultados de Kappa e suas variâncias entre os algoritmos classificadores por *Pixel* e por Região.

Classe	MV250	MV500	MV750	MV1000	DE250	DE500	DE750	DE1000	RN250	RN500	RN750	RN1000
BT250	3.520	4.500	4.871	4.559	4.022	5.296	4.991	5.375	4.404	3.837	3.965	4.875
BT500	3.662	4.955	5.547	5.282	4.212	5.873	5.697	6.280	4.622	4.193	4.469	5.672
BT750	4.096	5.684	6.509	6.360	4.662	6.663	6.680	7.463	5.078	4.880	5.343	6.796
BT1000	4.215	5.941	6.891	6.815	4.791	6.951	7.074	7.978	5.209	5.114	5.677	7.278
IS250	4.833	6.218	6.829	6.632	5.341	7.056	6.965	7.514	5.715	5.532	5.870	6.977
IS500	5.539	7.512	8.578	8.618	6.079	8.460	8.753	9.681	6.462	6.752	7.469	9.044
IS750	5.563	7.713	8.955	9.109	6.115	8.701	9.147	10.246	6.504	6.924	7.788	9.569
IS1000	5.121	7.192	8.404	8.526	5.686	8.203	8.599	9.702	6.087	6.380	7.198	9.001
HI250	3.765	4.822	5.238	4.944	4.271	5.630	5.361	5.777	4.652	4.152	4.317	5.267
HI500	4.537	6.132	6.929	6.789	5.083	7.067	7.090	7.819	5.482	5.369	5.833	7.196
HI750	4.347	6.033	6.930	6.830	4.912	7.016	7.105	7.943	5.324	5.230	5.760	7.272
HI1000	4.806	6.760	7.882	7.935	5.376	7.772	8.074	9.110	5.783	5.941	6.672	8.407