

UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE CIÊNCIAS AGRÁRIAS
DEPARTAMENTO DE ENGENHARIA AGRÍCOLA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA AGRÍCOLA

BEATRIZ FERNANDES SIMPLICIO SOUSA

**APRENDIZADO DE MÁQUINA NA DETECÇÃO DO USO DO SOLO NO BIOMA
CAATINGA VIA SENSORIAMENTO REMOTO**

FORTALEZA

2009

BEATRIZ FERNANDES SIMPLICIO SOUSA

APRENDIZADO DE MÁQUINA NA DETECÇÃO DO USO DO SOLO NO BIOMA
CAATINGA VIA SENSORIAMENTO REMOTO

Dissertação submetida à Coordenação do Programa de Pós-graduação em Engenharia Agrícola, da Universidade Federal do Ceará, como requisito parcial para a obtenção do grau de Mestre em Engenharia Agrícola.

Área de concentração: Manejo e Conservação de Bacias Hidrográficas no Semiárido.

Orientador: Professor Adunias dos Santos Teixeira, Ph.D.

Co-orientador: Francisco de Assis Tavares Ferreira da Silva, Dr.

FORTALEZA

2009

BEATRIZ FERNANDES SIMPLICIO SOUSA

APRENDIZADO DE MÁQUINA NA DETECÇÃO DO USO DO SOLO NO BIOMA
CAATINGA VIA SENSORIAMENTO REMOTO

Dissertação submetida à Coordenação do Programa de Pós-Graduação em Engenharia Agrícola, da Universidade Federal do Ceará, como requisito parcial para a obtenção do grau de Mestre em Engenharia Agrícola. Área de concentração em Manejo e Conservação de Bacias Hidrográficas no Semiárido.

Aprovado em 06 de março de 2009

BANCA EXAMINADORA

Prof^o. Adunias dos Santos Teixeira (Orientador), Ph.D.
Universidade Federal do Ceará - UFC

Dr. Francisco de Assis Tavares Ferreira da Silva (Co-Orientador), Dr.
Instituto Nacional de Pesquisas Espaciais - INPE

Prof^o. Arthur Plínio de Souza Braga (Conselheiro), Dr.
Universidade Federal do Ceará - UFC

Prof^a. Eunice Maia de Andrade (Conselheira), Ph.D.
Universidade Federal do Ceará - UFC

Dedico à Deus, meu mantenedor.

AGRADECIMENTOS

Começo agradecendo a quem me concedeu a vida: Deus e àqueles a quem Ele permitiu que me educassem com tanta dedicação: Meus pais Ademar e Célia.

À minha família pelo apoio oferecido.

Ao meu esposo Thiago Handerson. Obrigada por estar ao meu lado, por ter tido muita paciência e ter me dado forças em todos os momentos.

Ao professor Adunias pela oportunidade e orientação proporcionada não só no período de mestrado, mas também na graduação.

Ao professor Tavares por estar sempre disposto a me orientar quer fosse pessoalmente ou pela internet e sempre me animar para melhorar cada vez mais a pesquisa. Aproveito ainda para agradecer também ao Instituto Nacional de Pesquisas Espaciais (INPE), o qual cedeu as imagens de satélite gratuitamente.

Ao amigo Davi. Sua ajuda foi essencial! Obrigada pela parceria.

Aos amigos e amigas de mestrado. Meu coração tem um lugarzinho para cada um de vocês! Em especial agradeço ao Clemilson e Firmino pela ajuda na programação e também ao Carlos Alexandre, ao Clênio Jario e ao Amaury pela ajuda na coleta dos pontos em Iguatu.

Ao CNPq pelo apoio financeiro e à Universidade Federal do Ceará.

À Professora Eunice Maia do departamento de Engenharia Agrícola e ao Professor Arthur Braga do departamento de Engenharia Elétrica da Universidade Federal do Ceará. Muito obrigada por contribuir para a melhoria do trabalho.

O temor do Senhor é o princípio da sabedoria...

Provérbios 9:10

RESUMO

APRENDIZADO DE MÁQUINA NA DETECÇÃO DO USO DO SOLO NO BIOMA CAATINGA VIA SENSORIAMENTO REMOTO

O manejo adequado dos recursos naturais em ambientes frágeis, como o da Caatinga, requer o conhecimento de suas propriedades e distribuição espacial. Desta forma, o presente trabalho propõe uma abordagem para a classificação de imagens do satélite LANDSAT-5, correspondente a uma região semiárida localizada no município de Iguatu no Estado do Ceará, objetivando detectar o bioma da Caatinga por meio de dois tipos de classificadores baseados em aprendizado de máquina: o método baseado em Perceptrons de Múltiplas Camadas-MLP (do inglês *Multi Layer Perceptron*) e o método Máquinas de Vetores de Suporte-SVM (do inglês *Support Vector Machine*). O classificador estatístico da máxima verossimilhança, por ser amplamente utilizado na literatura, também foi aplicado à área em estudo para que o desempenho dos métodos propostos fosse comparado aos destes. Cinco classes foram definidas para a classificação, a saber: agricultura, antropizada, água, caatinga herbácea arbustiva (CHA) e caatinga arbórea densa (CAD). Para o método MLP, foram realizados testes variando a quantidade de neurônios na camada intermediária. Já os testes para o método SVM consistiram em variar o parâmetro σ da função gaussiana e o parâmetro de penalização (C). A eficiência dos métodos foi analisada por meio dos coeficientes de Exatidão Global, Exatidão Específica e de Kappa calculados por meio dos dados da matriz de confusão. Esta, por sua vez, foi gerada para cada método a partir da comparação entre a classificação e os pontos georreferenciados com aparelho GPS (correspondentes à verdade terrestre). O método MLP apresentou melhor desempenho para o teste em que 12 neurônios foram atribuídos à camada intermediária, com valores de Exatidão Global e de Kappa de 82,14% e 0,76, respectivamente. Já o método SVM apresentou melhor performance para o teste com $C=1000$ e $\sigma=2$ no qual se obteve valores de 86,03% e 0,77 para os coeficientes de Exatidão Global e Kappa, respectivamente. O valor de Exatidão Global para o classificador estatístico da máxima verossimilhança permitiu concluir que 81,2% dos pixels foram classificados corretamente e o coeficiente de Kappa para este método foi de 0,73. Os valores dos coeficientes de Exatidão Específica, que proporcionam analisar o desempenho dos métodos em cada classe, foram superiores a 70%. A área total classificada foi de 576 km² e, dentre as duas classes consideradas para o bioma Caatinga, a predominante é a do tipo caatinga herbácea arbustiva (CHA). Assim, por meio dos resultados experimentais obtidos, pode-se afirmar que os métodos SVM e MLP, baseados em aprendizado de máquina, apresentaram desempenho satisfatório para a classificação do bioma Caatinga.

Palavras-chave: Inteligência artificial, semiárido, classificação de imagens de satélite.

ABSTRACT

REMOTE SENSING AND MACHINE LEARNING APPLIED TO SOIL USE DETECTION IN CAATINGA BIOMA

In order to manage adequately natural resources inside a fragile environment, just like *Caatinga*, one should know its properties and spatial distribution. This work proposes an approach to classify LANDSAT-5 satellite images. These images, corresponding to a semi-arid environment located in Iguatu country, Ceara, Brazil, were classified aiming at detecting the *Caatinga* biome by two type of classifiers based on machinery learning: Multi Layer Perceptron (MLP) and Support Vector Machine (SVM). The static classifier of Maximum Likelihood was also used as comparison to the other two methods. Agriculture, water, anthropical, herbaceous shrub *Caatinga* (CHA) and dense high *Caatinga* (CAD) are the five classes defined for classifying. MLP method tests were carried out changing neurons quantity in the intermediate layer. SVM method tests were carried out changing σ , from Gauss function, and penalization parameter (C). Performance of the tests was analyzed by Global Accuracy, Specific Accuracy and Kappa coefficient. The last one calculated by confusion matrix, which has been generated by comparison of classification data and ground control points GPS georeferenced (true points). MLP method presented best performance for tests in which 12 neurons have been attributed to the intermediate layer resulting in Global Accuracy and Kappa values of 82.14% and 0.76, respectively. On the other hand, SVM method presented best performance for tests carried out with C=1000 and $\sigma=2$, resulting in Global Accuracy and Kappa values of 86.03% and 0.77, respectively. The Maximum Likelihood classifier presented 81.2% of its pixels correctly classified (Global Accuracy) and K coefficient value of 0.73. The values of Specific Accuracy, which makes it possible to analyze the performance of each individual class, were above 70% in each class. A total 576 km² area was classified. Between the two types of *Caatinga* biome considered, herbaceous shrub *Caatinga* (CHA) comes to be the most common. Therefore, taking into account experimental results, it is possible to conclude that both SVM and MLP methods, which are based on machine learning, show satisfactory performance for classifying *Caatinga* biome.

Key-words: artificial intelligence, semi-arid, satellite image classification.

LISTA DE FIGURAS

Figura 1 – Modelo do neurônio artificial	21
Figura 2 – Funções de ativação	22
Figura 3 – Topologia de uma RNA	23
Figura 4 – Redes feedforward com uma única camada.....	24
Figura 5 – Redes feedforward com única camada oculta e uma camada de saída.....	25
Figura 6 - Desenho esquemático ilustrando o hiperplano ótimo, a margem de separação e os vetores de suporte	31
Figura 7 – Ilustração das posições das variáveis de folga	33
Figura 8 – Ilustração de mapeamento para o espaço característico.....	35
Figura 9 – Exemplo de uma matriz de confusão	40
Figura 10 – Localização da bacia hidrográfica do Alto Jaguaribe	43
Figura 11 – Precipitação mensal do ano de 2008 para o município de Iguatu-CE.....	44
Figura 12 – Município de Iguatu e área de estudo representada pela imagem do satélite LANDSAT-5 de 20/08/08	45
Figura 13 – Classe água entre a Caatinga.....	46
Figura 14 – Classe CAD no município de Iguatu-CE	47
Figura 15 – Classe CHA no município de Iguatu-CE	47
Figura 16 – Exemplo de área antropizada	48
Figura 17 – Área de agricultura no município de Iguatu-CE	48
Figura 18 – Arquitetura da MLP correspondente ao teste 1	51
Figura 19 – Fluxograma das atividades relacionadas ao método MLP.....	52
Figura 20 – Fluxograma das atividades relacionadas ao método SVM	54
Figura 21 – Georreferenciamento de ponto correspondente à verdade terrestre.....	55
Figura 22 – (A) Pontos georreferenciados na imagem original; (B) Pontos georreferenciados na imagem classificada pelo método da máxima verossimilhança	56
Figura 23 – Valores de RGB em plixal da classe água	56
Figura 24 – Dados de entrada RGB e suas respectivas classes após processamento no MATLAB	56
Figura 25 – Planilha com o teste e a saída desejada.....	57
Figura 26 – Produção da matriz de confusão	57
Figura 27 – Etapas e ordem de execução	58

Figura 28 – Resposta espectral média das diferentes classes de cobertura do solo das amostras de treinamento	60
Figura 29 – Classificação pelo método da máxima verossimilhança.....	62
Figura 30 – Classificação por meio de uma rede MLP	64
Figura 31 – Comportamento do coeficiente de Kappa com C mantido constante e variando o fator ξ	66
Figura 32 – Comportamento do coeficiente de Kappa com ξ mantido constante e variando o fator C.....	67
Figura 33 – Classificação utilizando o método SVM.....	68
Figura 34 – Área total por cada classe (km ²).....	70

LISTA DE TABELAS

Tabela 1 – Qualidade da classificação associada aos valores da estatística de Kappa	42
Tabela 2 – Classes de interesse na imagem do satélite LANDSAT-5 na combinação 432 do espaço RGB	49
Tabela 3 – Testes realizados na rede MLP	51
Tabela 4 – Índice de separabilidade entre as classes	61
Tabela 5 – Matriz de confusão para o método da máxima verossimilhança.....	62
Tabela 6 – Resultados dos testes pelo método MLP	63
Tabela 7 – Matriz de confusão para o método MLP	64
Tabela 8 – Desempenho do classificador SVM com variação de ξ	65
Tabela 9 – Desempenho do classificador SVM com variação de do fator C	66
Tabela 10 – Matriz de confusão para o método SVM.....	67
Tabela 11 – Exatidão específica das classes para cada método	69

LISTA DE ABREVIATURAS E SIGLAS

IA	Inteligência Artificial
RNA	Rede Neural Artificial
MLP	Multi Layer Perceptron
LANDSAT	Land Remote Sensing Satellite
TM	Thematic Mapper
SVM	Support Vector Machine
VC	Vapnik-Chervonenkis
ATM	Airborne Thematic Mapper
MISR	Multi-angle Imaging Spectro-Radiometer
OA0	One-Against- One
OAA	One-Against-All
AVRIS	Airbone Visible Infrared Imaging Spectrometer
RBF	Radial Basis Function
JM	Jeffries-Matusita
EG	Exatidão Global
EE	Exatidão Específica
K	Coeficiente de Kappa
CAD	Caatinga Arbórea Densa
CHA	Caatinga Herbácea Arbustiva
IA	Inteligência Artificial
RNA	Rede Neural Artificial
SVM	Support Vector Machine
MLP	Multi Layer Perceptron

LISTA DE SÍMBOLOS

ω_i	Classe i
$P(\omega_i/X)$	Probabilidade de ω_i ocorrer na posição X;
$P(\omega_i)$	Probabilidade da classe i ocorrer na imagem;
$P(X/\omega_i)$	Distribuição de probabilidade de X dado uma classe ω_i ;
$P(X)$	Probabilidade do pixel X ocorrer na imagem.
exp	Função exponencial
ln	Logaritmo neperiano
μ_i	Vetor médio da classe i
$g_i(X)$	Função discriminante da classe i
x_m	Elementos do vetor de entrada da MLP
w_{kj}	Pesos entre o neurônio j e a entrada i
u_k	Combinador linear
y_k	Sinal de saída do neurônio
v_k	Campo local induzido
$\varphi(\cdot)$	Função de ativação
d_k	Saída desejada
e_k	Termo de erro do neurônio de saída
t	Iteração da rede MLP
η	Coefficiente de aprendizado para o algoritmo <i>backpropagation</i>
δ_j	Gradiente local
α	Taxa de momento
\mathbf{x}_i	Objeto, padrão de entrada ou atributo da SVM
d_i	Saída desejada
\mathbf{w}	Vetor de peso da SVM
b	Bias
\mathbf{w}_o	Vetor de peso ótimo da SVM
b_o	Bias ótimo
ρ	Margem de separação
α_i	Multiplicador de Lagrange
$f(x)$	Função de separação do hiperplano
ξ_i	Magnitude do erro de classificação
Φ	Hiperplano de separação ótimo

C	Parâmetro de penalização
$K(\mathbf{x}_i, \mathbf{x}_j)$	Função kernel
σ	Largura da gaussiana
B_{ji}	Distância de Bhattacharyya
M_i	Vetor de média da classe i
Σ_i	Matriz de covariância da classe i
r	Número de linhas da matriz de erro
x_{ii}	Valor da linha i e coluna i
N	Quantidade total de pontos amostrais
n	Quantidade total de pontos amostrais da classe r
x_{i+}	Soma da linha i
x_{+i}	Soma da coluna i .

SUMÁRIO

1. INTRODUÇÃO.....	14
2. REVISÃO DE LITERATURA	16
2.1 O bioma Caatinga	16
2.2 Sensoriamento remoto e classificação de imagens de satélite.....	17
2.2.1 Classificação supervisionada.....	18
2.2.2 Classificação não-supervisionada.....	20
2.3 Aprendizado de máquina	20
2.3.1 Definição de Rede Neural Artificial (RNA) e o modelo do neurônio artificial	21
2.3.2 Arquiteturas das RNA	23
2.3.2.1. Redes <i>feedforward</i> (única camada)	24
2.3.2.2.Redes <i>feedforward</i> (múltiplas camada)	25
2.3.3.Treinamento e Teste	26
2.3.4 Máquinas de Vetores de Suporte (SVM).....	29
2.3.4.1 Casos linearmente separáveis	30
2.3.4.2 Casos não linearmente separáveis	33
2.3.4.3 Classificadores binários e classificadores multiclases	38
2.4 Distância de Jeffries-Matusita	39
2.5 Validação da exatidão do mapeamento	40
2.6 Análise da separabilidade das classes.....	42
3. MATERIAL E MÉTODOS.....	43
3.1 Localização, aspectos fisiográficos e climáticos da área de estudo	43
3.2 Configuração da máquina.....	45
3.3 Aquisição da imagem e georreferenciamento	45
3.4 Escolha das classes e processo de amostragem	46
3.5 Análise da separabilidade das amostras de treinamento.....	49
3.6 Classificação no software ENVI 4.3.....	50
3.7 Classificação no software MATLAB	50
3.7.1 Método Perceptron de Múltiplas Camadas (MLP).....	50
3.7.2 Método Máquina de Vetor de Suporte	53
3.8 Validação da classificação.....	55
3.9 Fluxograma.....	58
3.10 Análise da separabilidade das classes.....	59

4 RESULTADOS E DISCUSSÃO	60
4.1 Análise da separabilidade das classes.....	60
4.2 Classificação pelo método estatístico da máxima verossimilhança	61
4.3 Classificação por meio do método Perceptron de Múltiplas Camadas (MLP).....	63
4.4 Classificação por meio do método Máquina de Vetor de Suporte (SVM).....	65
4.5 Avaliação dos coeficientes de exatidão específica (EE)	69
4.6 Quantificação das áreas	69
5. Conclusões.....	71
6. Perspectivas	72
REFERÊNCIAS	73
APÊNDICE 1	80
APÊNDICE 2	84

1. INTRODUÇÃO

Ações de degradação à natureza, provocadas pelo homem, fazem com que esta tenha reações que prejudicam ao próprio agente causador. O que vem se notando é que o homem está adquirindo novo posicionamento com relação aos recursos naturais, passando a apresentar maior preocupação com a questão da degradação ambiental (OLIVEIRA et al., 2003) e com práticas não-sustentáveis de uso dos recursos naturais.

A Caatinga, um dos mais ricos biomas brasileiros, encontra-se ameaçada. O que agrava a situação é que este bioma, que é exclusivamente brasileiro e ocupa 11% do território nacional, é um dos menos estudados e protegidos do Brasil (WANDSCHEER, 2008).

Ribeiro et al. (2001), afirmam que após séculos de gestão desastrosa dos recursos naturais, atingimos estágios onde a manutenção do nosso padrão de vida tornou-se incompatível com o de muitas outras espécies e que uma possibilidade de reverter esta situação talvez esteja no conhecimento sobre a dinâmica dos ecossistemas. Segundo Wandscheer (2008), na Caatinga a maior agressão decorre da extração de lenha. No entanto, o conhecimento mais aprofundado sobre este bioma pode tornar possível realizar esta atividade sem destruí-la, pela adoção de um planejamento da exploração sustentável, de forma que a Caatinga possa se recompor.

Assim, a busca pelo conhecimento das características de cobertura do solo através de classificadores de imagens de satélite tem sido objeto de estudo e pesquisa pela comunidade científica (SOUSA et al., 2007). É de extrema importância que a classificação da cobertura do solo seja realizada de forma rápida e exata, ou seja, que exista uma boa correspondência entre o mapa e a realidade terrestre.

O mapeamento da cobertura do solo é uma informação essencial em estudos de gestão ambiental, em avaliações de biodiversidade e como apoio para decisões de ações ambientais, sociais e políticas econômicas. Este tipo de mapeamento também pode ser considerado fundamental para realizar análises de erosão por meio de modelagens, já que serve como dado de entrada para diversos modelos utilizados. As modelagens podem gerar informações importantes para áreas onde qualidade de água e erosão são problemas, como é o caso de semiárido.

Na literatura é possível encontrar desde métodos tradicionais para classificação de imagens de satélite como, por exemplo, o da máxima verossimilhança (QUEIROZ et al., 2004 e SOUSA et al., 2007) e o da mínima distância (LEDUC et al., 2007) até métodos mais avançados como, por exemplo, os que utilizam aprendizado de máquina (CARVALHO et al.,

2004). Diversos métodos contidos no grupo do aprendizado de máquina têm atraído a atenção da comunidade de sensoriamento remoto (GIGANDET et al., 2005). As Support Vector Machines, por exemplo, estão presentes na nova geração de sistemas de aprendizado supervisionado baseados na teoria de aprendizagem estatística. Uma SVM é basicamente uma máquina linear, cuja idéia principal é construir um hiperplano como superfície de decisão de tal forma que a margem de separação entre exemplos positivos e negativos seja máxima (HAYKIN, 1999). As SVMs também têm sido aplicadas em problemas de classificação de dados provenientes do sensoriamento remoto (MERCIER et al., 2003; GONÇALVES et al., 2006) e os resultados mostram que estas podem obter uma melhor performance na classificação e ter melhor capacidade de generalização do que classificadores estatísticos tradicionais.

Desta forma, o objetivo do presente trabalho foi avaliar o desempenho de algoritmos baseados em aprendizado de máquina na detecção do uso do solo no bioma Caatinga utilizando imagens do satélite LANDSAT-5. Foram escolhidos os métodos Perceptrons de Múltiplas Camadas (MLP) e Support Vector Machines (SVM). Os resultados destes foram ainda comparados com o método estatístico da máxima verossimilhança.

Esta dissertação está organizada em seis capítulos. O corrente capítulo contém uma introdução com uma contextualização sobre a importância da classificação de imagens de satélite. No segundo capítulo são apresentadas as fundamentações teóricas sobre o sensoriamento remoto, bem como as técnicas computacionais tradicionais e inteligentes para classificação de imagens de satélite. Ainda neste capítulo é feita uma revisão de trabalhos que abordam a utilização dos métodos de classificação propostos. No terceiro capítulo estão apresentados os materiais e métodos. Já os resultados do experimento são detalhados no quarto capítulo. E, finalmente, o quinto capítulo apresenta conclusões e o sexto as perspectivas para trabalhos futuros.

2. REVISÃO DE LITERATURA

Na primeira subsecção deste capítulo é apresentada uma breve descrição do bioma Caatinga. Em seguida, a base teórica para o processo de classificação de imagens de satélite por meio de métodos tradicionais e de métodos baseados em aprendizado de máquina é apresentada. Assim como a fundamentação dos métodos de validação para classificações. Ainda neste capítulo são apresentados os trabalhos que foram utilizados como referência ao desenvolvimento deste.

2.1 O bioma Caatinga

A Caatinga, um dos maiores biomas brasileiros, ocupa uma área de aproximadamente 935.000 km², correspondendo à aproximadamente 54% da região Nordeste e 11% do território brasileiro. Está presente em vários Estados como Piauí, Ceará, Rio Grande do Norte, Paraíba, Pernambuco, Alagoas, Sergipe, Bahia e Minas Gerais. Trata-se do único ecossistema exclusivamente brasileiro, sendo composto por um mosaico de florestas secas e vegetação arbustiva. Este bioma encontra-se presente no Estado do Ceará com uma área de 126.926 km² o que corresponde a quase 85% da área total. Infelizmente, o Ceará é o Estado nordestino em que a Caatinga encontra-se em maior estado de devastação (MMA, 2002).

Segundo Maia (2004), a presença deste bioma pode proporcionar proteção contra erosão (já que o solo fica protegido pelas copas da vegetação); conservação da fertilidade do solo; conservação da água; conservação do clima (a vegetação ameniza os extremos do clima) e proteção natural contra pragas (pois abriga todos os organismos responsáveis pela manutenção do equilíbrio natural).

O clima da Caatinga se caracteriza por ser semi-árido, quente e apresentar duas estações distintas: a estação chuvosa, que dura de 3 a 5 meses, e a estação seca, que pode prolongar-se de 7 a 9 meses (MAIA, 2004).

A Caatinga é um bioma muito rico apresentando fauna e flora com características peculiares que os permitem viver nestas condições. Já foram registradas 932 espécies de plantas vasculares, 185 espécies de peixes, 154 répteis e anfíbios, 348 espécies de aves e 148 espécies de mamíferos. A caatinga possui três estratos de formação vegetal: arbóreo (8 a 12 metros), arbustivo (2 a 5 metros) e herbáceo (abaixo de 2 metros) (MMA, 2002).

Algumas características deste bioma podem ser citadas: Alta diversidade e heterogeneidade das espécies; proteção do solo; proteção da água; adaptação das espécies às condições semi-áridas (como a perda da folhagem na estação seca).

De acordo com LEAL et al. (2005), a pecuária bovina associada a práticas agrícolas degradáveis, iniciadas no processo de colonização, resultou em um bioma empobrecido, devastado e com alterações nas espécies nativas (como árvores mais baixas, caule mais fino, menor número de espécies). Como exemplo de práticas degradáveis pode-se citar a retirada de madeira para diversos fins, agricultura itinerante, aumento dos rebanhos e desmatamento de terrenos de alta declividade (favorecendo a erosão). De acordo com o autor, o desmatamento reduz a diversidade florística, resultando na quebra do equilíbrio entre espécies, exposição do solo e perda do banco de sementes. Práticas errôneas como estas resultaram em 28 espécies ameaçadas de extinção, nacional ou globalmente.

2.2 Sensoriamento remoto e classificação de imagens de satélite

Erbert (2001) afirma que o sensoriamento remoto tem por objetivo estudar objetos ou regiões na superfície da Terra sem entrar diretamente em contato com os mesmos. Estes objetos de interesse, como: vegetação, culturas agrícolas, solos, formações rochosas, corpos de água, entre outros, são denominados de alvos. As informações sobre os alvos são obtidas basicamente através do estudo da interação da radiação eletromagnética emitida por fontes naturais (ex: Sol) ou artificiais (ex: Radar) com estes alvos.

As imagens provenientes do sensoriamento remoto vêm servindo como base para estudos e levantamentos geológicos, ambientais, agrícolas, cartográficos, florestais, urbanos, entre outros. Além disso, essas imagens passaram a representar uma das únicas formas viáveis para realizar monitoramentos ambientais tanto em escalas locais como em globais já que são produzidas com rapidez, eficiência e periodicidade

Técnicas de classificação de imagens implicam na implementação de um processo de decisão para que o computador possa atribuir certo pixel a uma classe (NOVO, 1989). Segundo Iwai (2003), são gerados neste processo, mapas temáticos ou cartas temáticas correspondentes aos temas de interesse, de acordo com critérios bem definidos. Cada pixel da imagem original é classificado em alguma das classes ou temas definidos e representado por símbolos gráficos ou cores. A maioria dos classificadores baseia-se exclusivamente na detecção da assinatura espectral (isto é, padrões de respostas espectrais) de classes de

cobertura do solo. Neste caso, para que a classificação seja bem sucedida, é necessário que as assinaturas espectrais dos alvos sejam bem definidas e distintas entre si.

2.2.1 Classificação supervisionada

Nos métodos de classificação supervisionada, as classes são previamente definidas pelo analista, isto é, definidas ou caracterizadas através das amostras de treinamento. Os métodos pertencentes a este tipo de classificação utilizam algoritmos que realizam o reconhecimento dos padrões espectrais na imagem com base nessas amostras (MOREIRA, 2005). Como exemplo, pode-se citar alguns métodos estatísticos de classificação supervisionada: da distância mínima, do paralelepípedo e da máxima verossimilhança.

O classificador estatístico da máxima verossimilhança rotula os pixels em uma determinada classe conforme a distância estatística entre cada pixel e a média da resposta espectral de classe (ASSAD e SANO, 1998). É um dos classificadores mais aplicados no tratamento de dados provenientes do sensoriamento remoto como pode ser visto nos trabalhos de Queiroz et al. (2004) e Sousa et al. (2007). Em comparação aos demais classificadores supervisionados tradicionais, este método vem apresentando resultados superiores como mostra os estudos de Oruc et al. (2004), de Silva e Pereira (2007) e de Figueiredo e Carvalho (2007).

Uma função de probabilidade $P(x)$ define a que classe pertencerá determinado pixel. O cálculo é feito para as várias classes pré-definidas e o pixel é atribuído à classe cuja probabilidade de pertencer a esta seja maior. É considerado um método paramétrico, pois envolve parâmetros (vetor média e matriz de covariância) da distribuição gaussiana multivariada, e é supervisionado, pois estima estes parâmetros através das amostras de treinamento (ERBERT, 2001).

As classes na imagem serão representadas por ω_i , $i = 1, \dots, M$, onde M é o número total de classes. Um pixel X pertencerá à classe ω_i se a probabilidade que associa este pixel à classe ω_i for maior do que a probabilidade que associa o pixel a qualquer outra classe (RICHARDS, 1993):

$$X_i \in \omega_i, \text{ se } P(\omega_i / X) > P(\omega_j / X), \text{ para todo } \omega_i \neq \omega_j \quad (1)$$

Onde $P(\omega_i / X)$ é a probabilidade de ω_i ocorrer na posição X .

Para determinar $P(\omega_i / X)$ têm-se o teorema de Bayes:

$$P(\omega_i / X) = \frac{P(X / \omega_i)P(\omega_i)}{P(X)} \quad (2)$$

Onde:

$P(\omega_i)$ – Probabilidade da classe i ocorrer na imagem;

$P(X/\omega_i)$ - Distribuição de probabilidade de X dado uma classe ω_i ;

$P(X)$ - Probabilidade do pixel X ocorrer na imagem.

Aplicando a Equação 2 na condição apresentada em 1, tem-se que:

$$X_i \in \omega_i, \text{ se } P(X/\omega_i) P(\omega_i) > P(X/\omega_j) P(\omega_j), \forall j \neq i \quad (3)$$

A probabilidade de determinada classe ocorrer na imagem, neste exemplo dado por $P(\omega_i)$ e $P(\omega_j)$, é quase sempre negligenciada no processo de classificação já que se supõe que todas as classes têm igual chance de ocorrência. Assim:

$$X_i \in \omega_i, \text{ se } P(X/\omega_i) > P(X/\omega_j), \forall j \neq i \quad (4)$$

O valor de $P(X/\omega_i)$ é obtido assumindo-se a distribuição normal para as classes:

$$P(X/\omega_i) = (2\pi)^{-\frac{N}{2}} |\Sigma_i|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(X - \mu_i)^T \Sigma_i^{-1}(X - \mu_i)\right\} \quad (5)$$

Onde N é o número de bandas; Σ_i a matriz de covariância da classe ω_i para as N bandas; μ_i vetor que contém o valor médio da resposta espectral da classe ω_i para as N bandas.

O valor médio da resposta espectral (μ_i) e a matriz de covariância (Σ_i) estão apresentados a seguir:

$$\mu_i = \frac{1}{K} \sum_{j=1}^K (X_j - \mu_i)(X_j - \mu_i)^T \quad (6)$$

$$\Sigma_i = \left(\frac{1}{K-1}\right) \sum_{j=1}^K X_j \quad (7)$$

Onde K é o número de pixels no conjunto de treinamento da classe ω_i .

Aplicando logaritmo natural à Equação 5, tem-se:

$$\ln P(X/\omega_i) = -\frac{N}{2} \ln(2\pi) - \frac{1}{2} |\Sigma_i| - \frac{1}{2} (X - \mu_i)^T \Sigma_i^{-1} (X - \mu_i) \quad (8)$$

A partir da simplificação matemática da Equação 8, a seguinte função de discriminação é obtida:

$$g_i(X) = -\ln|\Sigma_i| - (X - \mu_i)^T \Sigma_i^{-1} (X - \mu_i) \quad (9)$$

Desta forma, a regra de decisão para este método é dada por:

$$X_i \in \omega_i, \text{ se } g_i(X) > g_j(X), \forall j \neq i \quad (10)$$

2.2.2 Classificação não-supervisionada

Para realizar este tipo de classificação, os algoritmos deste método devem buscar encontrar as classes na imagem por agrupamento, ou seja, procurando agrupar pixels com padrões de reflectância semelhantes entre si ou separando grupos de pixels diferentes. Neste caso, estes agrupamentos não assumem qualquer conhecimento prévio da distribuição de densidade de probabilidade dos temas, como ocorre nos métodos supervisionados. O resultado deste tipo de método é uma imagem sem rótulos, ou seja, não se conhece previamente a identificação das classes resultantes e o próprio usuário deve, ao final do processo, rotular essas classes (MOREIRA, 2005). Alguns métodos comuns de “cluster” são os algoritmos Isodata e K-Médias.

Figueiredo e Carvalho (2007) realizaram uma análise comparativa entre técnicas de classificação digital em imagens do satélite LANDSAT para o mapeamento do uso e ocupação do solo em Capixaba, Acre. Por meio do experimento foi possível notar que, para as três imagens analisadas, o método supervisionado da máxima verossimilhança apresentou maiores valores de Kappa quando comparado ao método não supervisionado Isodata.

2.3 Aprendizado de máquina

O aprendizado de máquina está inserido na área de conhecimento denominada Inteligência Artificial (IA) e têm por objetivo desenvolver métodos, técnicas e ferramentas para construir máquinas inteligentes capazes de realizar as mais diversas tarefas. De acordo com Mitchell (1997), essas máquinas devem treinar sobre determinado conjunto de amostras e adquirir conhecimento de forma automática. As técnicas de aprendizado podem ser classificadas em supervisionadas (em que os rótulos das classes são previamente conhecidos e o algoritmo de aprendizado recebe um conjunto de exemplos de treinamento para cada classe) e não supervisionadas (em que os rótulos das classes não são conhecidos e o algoritmo de aprendizado procura formar agrupamentos ou clusters dos dados de entrada). Como exemplo de algoritmos de aprendizado de máquina pode-se citar os que têm como inspiração os sistemas biológicos (Redes Neurais Artificiais (RNAs) e Algoritmos Genéticos), o aprendizado simbólico (Árvores de Decisão) e as teorias estatísticas (Support Vector Machines (SVMs)) (MITCHELL, 1997). Os algoritmos RNA e SVM, utilizados nas experiências práticas deste trabalho, serão detalhados nesta seção.

2.3.1 Definição de Rede Neural Artificial (RNA) e o modelo do neurônio artificial

Segundo HAYKIN (1999), uma RNA pode ser definida como:

Um processador maciçamente paralelo distribuído constituído de unidades de processamento simples, que têm a propensão natural para armazenar conhecimento experimental e torná-lo disponível para uso. Ela se assemelha ao cérebro em dois aspectos:

1. O conhecimento é adquirido pela rede a partir de seu ambiente através de um processo de aprendizagem.
2. Forças de conexão entre neurônios, conhecidas como pesos sinápticos, são utilizadas para armazenar o conhecimento adquirido.

As unidades básicas das redes neurais, os neurônios artificiais, simulam as funções dos neurônios naturais e processam todas as informações das RNA.

A modelagem do neurônio biológico visando gerar o modelo do neurônio artificial foi iniciada por McCulloch e Pitts (1943) e seu modelo mais simples está exposto na Figura 1.

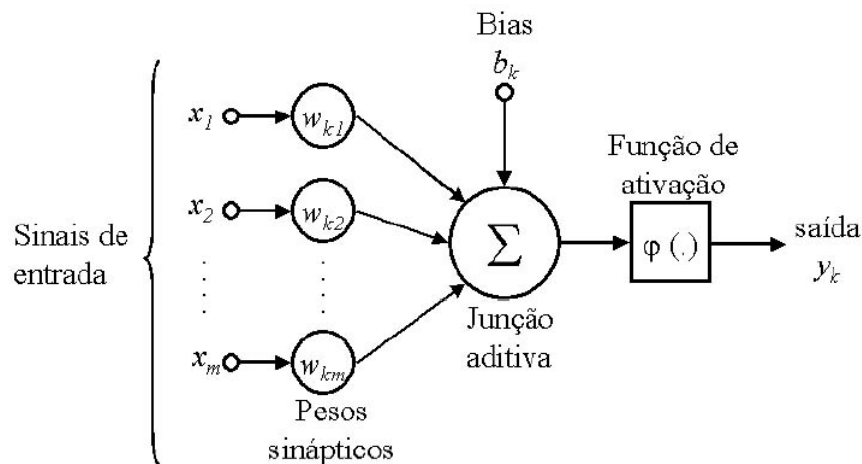


Figura 1 – Modelo do neurônio artificial

Fonte: Haykin, 1999

Este modelo é constituído de um conjunto de sinapses, sendo cada uma dessas caracterizada por um peso (w); um somador (Σ) para somar os sinais de entrada, ponderado pelas respectivas sinapses do neurônio; e uma função de ativação (φ) para restringir a amplitude de saída de um neurônio, ou seja, limitar o intervalo permissível de amplitude do sinal de saída a um valor finito (HAYKIN, 1999).

Observando-se a Figura 1, tem-se que: x_1, x_2, \dots, x_m são os elementos do vetor de entrada; w_1, w_2, \dots, w_{kj} são os pesos (ou conexões sinápticas); b_k : limiar de ativação do neurônio; $\varphi(\cdot)$ é a função de ativação e y_k é o sinal de saída do neurônio. É importante notar a

forma como são escritos os índices do peso sináptico w_{kj} . O índice k se refere ao neurônio em questão e o índice j ao terminal de entrada da sinapse à qual o peso se refere.

Os elementos do vetor de entrada x_1, x_2, \dots, x_m são multiplicados pelos pesos w_1, w_2, \dots, w_{km} . Desta forma, o neurônio k é descrito pelo seguinte par de Equações. Assim:

$$u_k = \sum_{j=1}^m w_{kj} \cdot x_j \quad (11)$$

e

$$y_k = \varphi(u_k + b_k) \quad (12)$$

O campo local induzido (v_k) é dado por:

$$v_k = u_k + b_k \quad (13)$$

A função de ativação deve processar o conjunto de entradas recebidas e o transformar em estado de ativação. Os neurônios podem assumir os seguintes estados de ativação: binários (0 e 1), bipolares (-1 e 1) e reais. Algumas das funções de ativação mais utilizadas estão ilustradas na Figura 2, que são: função linear, função tangente hiperbólica, função degrau e função sigmoideal.

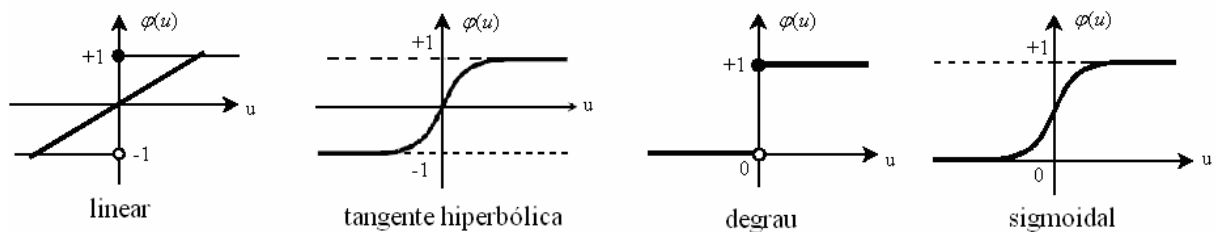


Figura 2 – Funções de ativação

A Equação 14 corresponde à função de ativação do tipo tangente hiperbólica, usada no experimento com rede Perceptron Múltiplas Camadas (MLP).

$$\varphi(u) = \frac{1 - e^{-u}}{1 + e^{-u}} \quad (14)$$

A topologia de uma RNA está ilustrada na Figura 3. Nesta, estão expostas as nomenclaturas das camadas que compõe uma RNA.

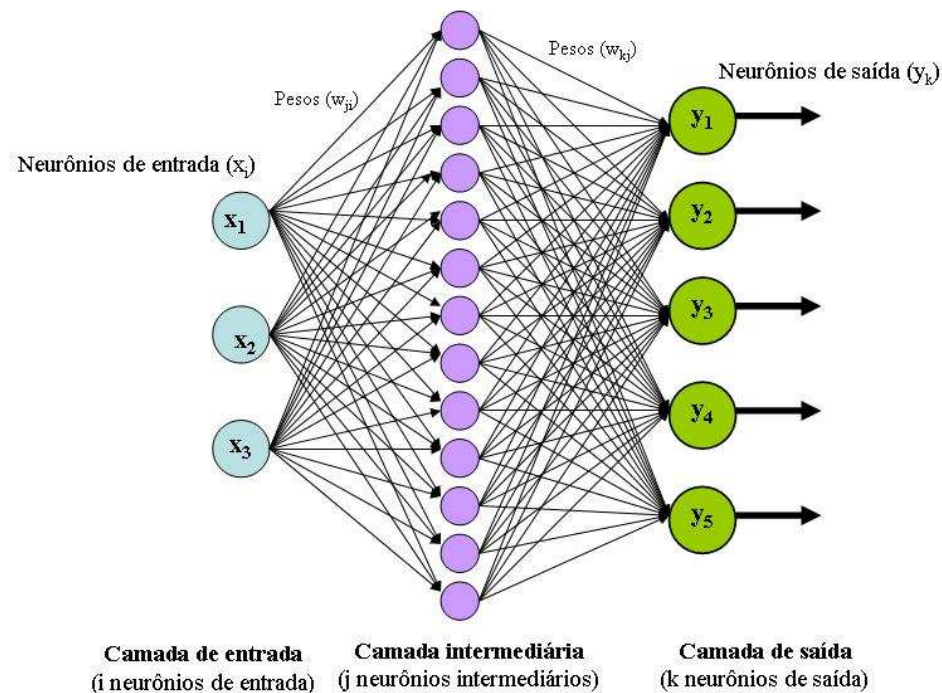


Figura 3 - Topologia de uma RNA
Fonte: Haykin, 1999

- 1- Camada de Entrada: onde os padrões são apresentados à rede (x_1, \dots, x_i).
- 2- Camadas Intermediárias ou Escondidas: Através das conexões ponderadas, estas camadas podem ser consideradas como extratoras de características.
- 3- Saída: onde o resultado final é concluído e apresentado (y_1, \dots, y_k).

O número de neurônios e camadas em uma dada estrutura é função do problema em questão. Na classificação de imagens de satélite, por exemplo, o número de neurônios na camada de entrada poderá corresponder ao número de bandas utilizadas; na camada intermediária o número de neurônios e de camadas ficará a critério do usuário e na camada de saída, a quantidade de neurônios corresponde ao número de classes para as quais os padrões poderão ser designados.

A primeira camada de neurônios usa os padrões de entrada da rede e, por meio do cálculo executado pela aplicação da função de transferência, gera a resposta que alimenta a camada seguinte, e assim sucessivamente até que a última camada seja atingida.

2.3.2 Arquiteturas das RNA

As arquiteturas de redes se classificam em função do número de camadas de neurônios e de sua estruturação. Abaixo serão apresentadas as mesmas:

2.3.2.1. Redes *feedforward* (única camada)

As redes *feedforward* com única camada correspondem ao tipo de RNA onde se tem uma camada de entrada e uma única camada de neurônios que é a própria camada de saída. O Perceptron com uma camada de neurônios foi desenvolvido por Rosenblatt em 1958 (ROSENBLATT, 1958). Esse tipo de arquitetura é definida como uma rede linear que classifica apenas padrões linearmente separáveis. A Figura 4 apresenta um exemplo deste tipo de arquitetura.

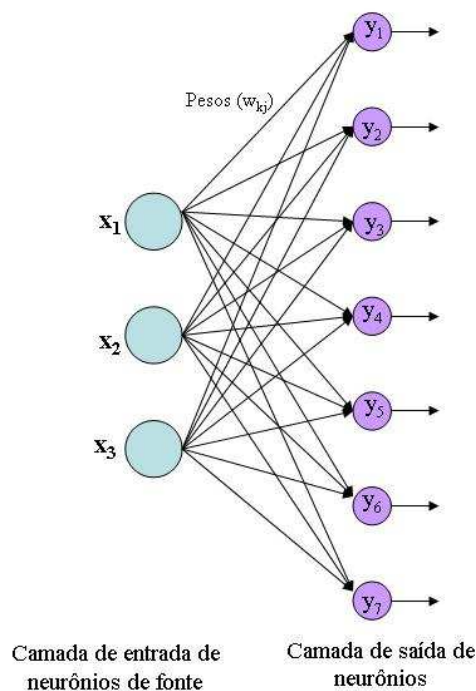


Figura 4 – Redes Feedforward com uma única camada
 Fonte: Adaptada de Haykin, 1999

Na Figura 4, apresentou-se um exemplo deste tipo de arquitetura em que uma camada de entrada de nós de fonte que se projeta sobre a camada de saída de neurônios, porém não ocorre o inverso (HAYKIN, 1999). Ou seja, a característica *feedforward* atribuída ao Perceptron significa que os sinais calculados apenas passam adiante na rede e não retornam nunca.

2.3.2.2 Redes *feedforward* (múltiplas camadas)

Normalmente, esta extensão do Perceptron de Rosenblatt, possui uma arquitetura constituída por uma camada de entrada, uma camada de saída e uma ou mais camadas escondidas com neurônios ocultos (Figura 5).

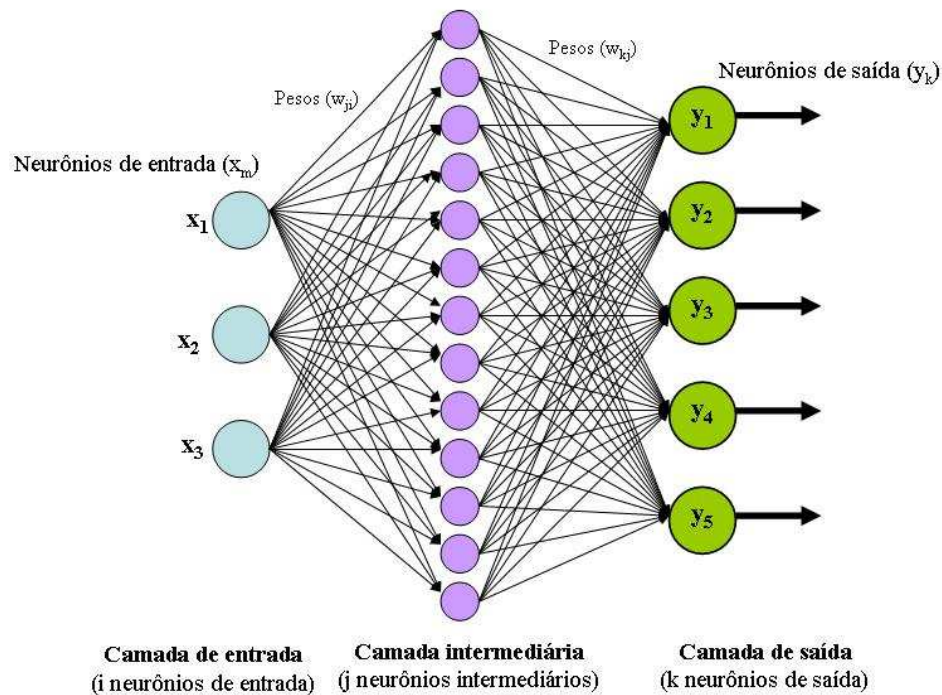


Figura 5 - Redes Feedforward com única camada oculta e uma camada de saída
 FONTE: Adaptada de Haykin, 1999

A função dos neurônios ocultos é intervir entre a camada externa e a saída da rede extraíndo as informações necessárias à classificação. Os neurônios da camada de entrada da rede fornecem os respectivos elementos do padrão de ativação (vetor de entrada), que constituem os sinais de entrada aplicados aos neurônios da segunda oculta. Na existência de mais de uma camada oculta os sinais de saída da segunda camada serão utilizados como entrada para esta, e assim por diante. O conjunto de sinais de saída dos neurônios da camada de saída (final) da rede constitui a resposta global da rede para o padrão de ativação fornecido pelos nós de fonte da camada de entrada (HAYKIN, 1999).

Para o exemplo ilustrado na Figura 5, a rede alimentada adiante com i neurônios na camada de entrada, j neurônios na camada oculta e k neurônios na camada de saída é escrita da seguinte forma: $i-j-k$. Por exemplo, uma rede com 3 neurônios na camada de entrada, 8 na intermediária e 5 na de saída é escrita como: 3-8-5.

Um dos problemas em relação à definição de uma topologia de uma MLP consiste em determinar o número de camadas e a quantidade de neurônios em cada uma. Se o número de camadas e neurônios por camada for muito maior que o necessário, haverá uma tendência de sobre-ajuste (“*overfitting*”), isto é, a rede treinará bem, podendo atingir um valor baixo para o erro, mas apresentará desempenho ruim no conjunto de testes. Desta forma, busca-se configurar uma rede de menor tamanho que cumpra a sua função. Entretanto, se o tamanho da rede for muito pequeno, a mesma poderá ser incapaz de atingir o desempenho desejado mesmo durante o treinamento. Neste caso ocorre *underfitting*, ou seja, poucos neurônios para armazenar muitas informações e com pouca capacidade de realizar um mapeamento não-linear da rede (VONK, 1997).

O Perceptron de Múltiplas Camadas é utilizado para a classificação de padrões não linearmente separáveis, ou seja, padrões que caem em determinadas posições de um hiperespaço nas quais não podem ser separados por um hiperplano. Esta capacidade de classificar dados não lineares a posiciona como uma rede adequada a solucionar problemas mais próximos da realidade, tais como: no processamento de imagens (LIN, 1991), no mapeamento de florestas tropicais (CARVALHO et al., 2004) e na classificação de terras para irrigação por meio da análise dos parâmetros de produtividade do solo (BUCENE e RODRIGUES, 2004), dentre outros.

2.3.3 Treinamento e Teste

O objetivo do treinamento em RNAs é que esta aprenda por meio de exemplos apresentados. Haykin (1999) define aprendizagem como: “Um processo pelo qual os parâmetros livres de uma rede neural são adaptados através de um processo de estimulação pelo ambiente na qual a rede está inserida. O tipo de aprendizagem é determinado pela maneira como a modificação dos parâmetros ocorre”.

Um algoritmo de aprendizagem largamente empregado nas redes neurais artificiais do tipo MLP é o *backpropagation*, relatado no trabalho de Rumelhart et al. (1986). Neste mesmo ano foi lançado o livro *Parallel Distributed Processing: Explorations in the Microstructures of Cognition* de Rumelhart e McClelland. O processo de aprendizagem realizado neste algoritmo pode ser resumido nos seguintes passos:

- 1 - Inicializar os pesos (\mathbf{w}), bias (b) e parâmetros da rede (número de camadas escondidas, número de neurônios na camada escondida, taxa de aprendizagem, etc);
- 2 - Apresentar as entradas e saídas desejadas da rede. As entradas serão os vetores x_0, x_1, \dots, x_i de valores contínuos e o vetor d_0, d_1, \dots, d_k serão as saídas desejadas;
- 3 - Calcular os valores internos (ativação) dos neurônios j da camada escondida h .

$$u_j = \sum_{i=1}^m w_{ji} x_m + b_j \quad (15)$$

- 4 - Calcular a saída dos neurônios z das camadas escondidas por meio da função de ativação (φ).

$$y_j = \varphi(u_j) \quad (16)$$

- 5 - Calcular os valores internos (ativação) dos neurônios k da camada de saída.

$$u_k = \sum_{i=1}^m w_{ki} y_j + b_k \quad (17)$$

- 6 - Calcular a saída dos neurônios k da camada de saída por meio da função de ativação (φ).

$$y_k = \varphi(u_k) \quad (18)$$

- 7 - Calcular os termos de erro para as unidades de saída:

$$e_k = d_k - y_k \quad (19)$$

- 8 - Adaptar os pesos por meio da seguinte Equação:

$$w_{ji}(t+1) = w_{ji}(t) + \eta \delta_j y_i' \quad (20)$$

onde w_{ji} é o peso do elemento de processamento oculto j no tempo t ; y_i saída produzida pela unidade i ; η corresponde à taxa de aprendizagem e; δ_j corresponde ao gradiente local para o elemento de processamento j ; Caso j seja um elemento da camada de saída (L) tem-se que:

$$\delta_j = e_j^{(L)} \varphi_j'(v_j^{(L)}) \quad (21)$$

Caso o elemento j seja um elemento oculto da camada l , então:

$$\delta_j = \varphi_j'(v_j^{(l)}) \sum_k \delta_k^{(l+1)} w_{kj}^{(l+1)} \quad (22)$$

Onde k denota todos os elementos acima dos elementos j ; os limiares delta dos elementos internos são ajustados de forma semelhante. A convergência poderá ser mais rápida pela adição de um termo de momento. Este termo faz com que os pesos sejam alterados de forma mais suave, pela Equação:

$$w_{ji}^{(l)}(t+1) = w_{ji}^{(l)}(t) + \alpha [w_{ji}^{(l)}(t-1)] + \eta \delta_j^{(l)} y_i^{(l-1)} \quad (23)$$

- 9 - Repetir retornando para o passo três até que seja satisfeito o critério de parada escolhido.

O momento de parar é uma decisão importante na etapa de treinamento e o critério que utiliza o número de iterações é bastante utilizado, porém, não garante que o erro desejado tenha sido obtido. Por exemplo, se a quantidade de ciclos escolhida for pequena o erro poderá estar muito elevado e a rede poderá não classificar os dados corretamente. Outro critério de parada é o que define o erro médio quadrático desejado. Ou seja, permite-se que a rede realize o número de ciclos necessários até que o erro inicialmente fixado seja obtido (GLEZAKOS e TSILIGIRIDIS, 2002).

Os classificadores baseados em aprendizado de máquina, no âmbito da classificação de imagens de satélite, vêm apresentando vantagens em relação aos classificadores supervisionados baseados em conceitos estatísticos, visto que para a sua aplicação nenhum conhecimento da forma da distribuição de probabilidade dos dados de entrada é exigido e face à maior flexibilidade que esses modelos possuem quanto à entrada de dados para a classificação (BISCHOF et al., 1992; RIBEIRO et al., 2005; GELETE E VOLATÃO, 2007).

Bischof et al. (1992) fizeram uma comparação entre os resultados da classificação de uma MLP com algoritmo *backpropagation* e o método da máxima verossimilhança em imagens do satélite LANDSAT . O estudo mostrou a capacidade da rede em obter resultado superior ao do classificador da máxima verossimilhança e que informações de textura podem ser agregadas possibilitando melhoria nos resultados. A exatidão global para a classificação utilizando o método estatístico da máxima verossimilhança, a rede MLP e MLP com textura foi de 84,7, 85,9, e 88,1%, respectivamente. A classificação visava gerar um mapa temático para as classes: antropizada, floresta, água e agricultura.

Ribeiro et al. (2005), comparou dois tipos de classificadores tradicionais (mínima distância, máxima verossimilhança) com uma RNA utilizando como dados de entrada as bandas azul, verde, vermelho e infravermelho próximo do satélite IKONOS II. Foram propostas para a classificação da área de estudo, localizada em uma área do município de Serra do Salitre (MG), quinze classes de uso do solo. Os valores de Kappa apontam para superioridade das RNAs, a saber: 0,68, 0,65 e 0,58 para RNA, máxima verossimilhança e mínima distância, respectivamente.

Gelete e Volatão (2007) usaram a Toolbox Neural Network do software MATLAB (The MathWorks Inc., Natick, MA, USA) para, através de uma rede do tipo MLP, distinguir as classes floresta, solo, água, culturas, nuvens e sombras na imagem do satélite LANDSAT (bandas 1 a 7) em uma área no Estado do Pará. A MLP que proporcionou melhores resultados possuía 11 neurônios na camada oculta. Os pesquisadores ainda compararam com o método

da máxima verossimilhança e obtiveram resultados demonstrando o melhor desempenho da rede MLP.

De acordo com Haykin (1999), no aprendizado não-supervisionado somente os padrões de entrada estão disponíveis para a rede, cujo conjunto de treinamento possui pares de entrada e saída. A partir do momento em que a rede estabelece uma harmonia com as regularidades estatísticas da entrada de dados, desenvolve-se nela uma habilidade de formar representações internas para codificar características da entrada e criar novas classes ou grupos automaticamente.

O modelo de Kohonen, também conhecido como Rede de Kohonen, é um tipo de rede com aprendizado não supervisionado e também vêm sendo utilizando na classificação de dados provenientes do sensoriamento remoto como mostram os trabalhos de Ji (2000), Jianwen et al. (2005).

2.3.4 Support Vector Machines (SVM)

Uma Máquina de Vetor de Suporte (SVM) é basicamente uma máquina linear, cuja idéia principal é construir um hiperplano como superfície de decisão de tal forma que a margem de separação entre exemplos positivos e negativos seja máxima (HAYKIN, 1999). Este método foi desenvolvido pelo pesquisador Vladimir Vapnik e colaboradores (BOSER, GUYON e VAPNIK, 1992; CORTES e VAPNIK, 1995; VAPNIK, 1995; VAPNIK, 1998) e é baseado na teoria de aprendizagem estatística, por meio do princípio de minimização estrutural de risco. De acordo com Osuna et al. (1997), a teoria de aprendizagem estatística estabelece condições matemáticas que auxiliam na escolha de um classificador a partir dos dados de treinamento. Essas condições levam em conta o desempenho deste classificador no conjunto de treinamento e a sua complexidade, visando adquirir desempenho satisfatório para novos dados de mesmo domínio. No princípio da minimização estrutural de risco, o erro de generalização ($R(f)$) será limitado pelo erro de treinamento ($R_{emp}(f)$) mais um termo que depende da dimensão Vapnik-Chervonenkis (VC), desenvolvida por Vapnik e Chervonenkis (1971), como mostrado na Equação 24.

$$R(f) \leq R_{emp}(f) = \sqrt{\frac{h(\ln\left(\frac{2n}{h}\right) + 1) - \ln\left(\frac{\theta}{4}\right)}{n}} \quad (24)$$

Onde, h é a dimensão (VC) da classe de funções F à qual f pertence, n é a quantidade de exemplos no conjunto de treinamento T .

Esta dimensão é uma medida da capacidade de um conjunto de funções classificadoras. Desta forma, um conjunto de hiperplanos é construído tendo como estratégia a variação da dimensão VC. Assim, o risco empírico (erro de treinamento) e a dimensão VC serão minimizados ao mesmo tempo. A minimização da dimensão VC ocorre quando é realizada a maximização da margem de separação, ou seja, no processo de construção do hiperplano ótimo. Maiores detalhes deste princípio podem ser encontrados em OSUNA et al. (1997).

O treinamento de uma SVM envolve a resolução de um problema quadrático, dependente dos vetores de treinamento e de parâmetros especificados pelo usuário. Neste método, a partir de um espaço de entrada de padrões não-linearmente separáveis é formado um novo espaço de características, em dimensão outra, onde os padrões serão linearmente separáveis. Assim, um hiperplano de separação ótimo entre os exemplos é construído (VAPNIK, 1995).

Algumas das principais vantagens apresentadas pelas SVMs são que estas: Possuem teoria bem estabelecida dentro da matemática e estatística; apresentam robustez em dados de grandes dimensões, possibilitando a aplicação de SVM em vetores de características de grandes dimensões como, por exemplo, imagens de satélite; apresentam ótima capacidade de generalização, ou seja, apresentam boa eficiência para prever dados que não pertençam ao conjunto utilizado no treinamento (SMOLA et al., 2000).

As SVMs vêm sendo aplicadas nas mais diversas áreas de pesquisa como, por exemplo, no reconhecimento de faces e caracteres, classificação de impressões digitais, classificação de padrões, dentre outras. Wang e Yang (2006) utilizaram uma *toolbox* baseada na teoria SVM e desenvolvida por Canu et al. (2005) para reconhecimento de olhos em seres humanos e obtiveram precisão de 96,8% nos experimentos realizados. O formalismo do método SVM para o desenvolvimento desta *toolbox* foi baseado em Vapnik (1995), Osuna et al. (1997) e Cristianini e Taylor (2000). Uma breve introdução deste formalismo é apresentado, a seguir.

2.3.4.1 Casos linearmente separáveis

Considere que os dados de treinamento com k números de amostras são representados por $\{\mathbf{x}_i, d_i\}$, $i = 1, \dots, N$, onde $\mathbf{x} \in \mathbb{R}^n$ está em um espaço n -dimensional, e que $d \in \{+1, -1\}$ é a saída desejada. A Equação da superfície de decisão do hiperplano é

$$\mathbf{w}^T \mathbf{x} + b = 0 \quad (25)$$

Em que \mathbf{x} é um vetor de entrada, \mathbf{w} é um vetor peso ajustável e b é um bias. A Equação 25 divide os dados em duas regiões:

$$\mathbf{w}^T \mathbf{x}_i + b \geq 0 \text{ para } d_i = +1 \quad (26)$$

$$\mathbf{w}^T \mathbf{x}_i + b < 0 \text{ para } d_i = -1 \quad (27)$$

Esses padrões serão considerados linearmente separáveis se um vetor de entrada \mathbf{w}_o e um bias b_o poderem ser definidos de tal forma que satisfaçam as condições:

$$\mathbf{w}_o^T \mathbf{x}_i + b_o \geq +1 \text{ para } d_i = +1 \quad (28)$$

$$\mathbf{w}_o^T \mathbf{x}_i + b_o \leq -1 \text{ para } d_i = -1 \quad (29)$$

Caso as Equações 28 e 29 sejam satisfeitas com sinal de igualdade, os pontos (\mathbf{x}_i, d_i) serão chamados de vetores de suporte.

Desta forma, \mathbf{w} e b estimados por meio das Equações 28 e 29 devem satisfazer a seguinte condição:

$$d_i(\mathbf{w} \cdot \mathbf{x}_i + b) > 0 \quad (30)$$

O hiperplano ótimo é a região na qual a distância do ponto mais próximo da margem é máxima. O conceito de hiperplano ótimo foi proposto pelos pesquisadores Vapnik e Chervonenkis em 1974 (VAPNIK e CHERVONENKIS, 1974). Para \mathbf{w} e b encontrados, a separação entre o hiperplano ótimo e o ponto mais próximo é denominada margem de separação (ρ). Os dados de treinamento que se encontram à distância ρ do hiperplano são chamados vetores de suporte e correspondem aos dados em que as Equações 28 e 29 são satisfeitas com sinal de igualdade (HAYKIN, 1999). Na Figura 6, o desenho esquemático ilustra o hiperplano ótimo em um espaço bi-dimensional, a margem de separação e os vetores de suporte para a classificação das classes “estrelas” e “círculos”.

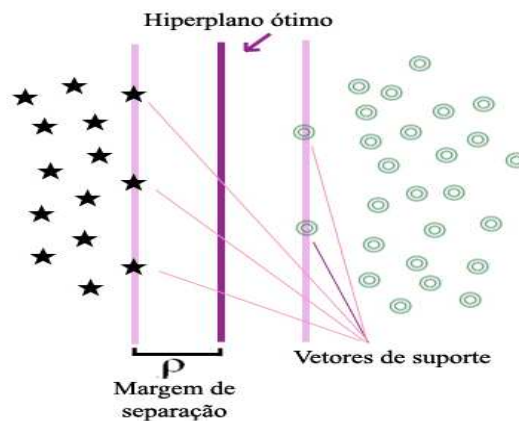


Figura 6 – Desenho esquemático ilustrando o hiperplano ótimo, a margem de separação e os vetores de suporte
Fonte: Semolini, 2002

A margem geométrica entre as duas classes é igual a $\frac{2}{\|\mathbf{w}\|}$. O conceito de margem é extremamente importante em SVM já que mede a capacidade de generalização (OSUNA et al., 1997). Para encontrar o hiperplano de separação ótimo, os valores para o vetor de peso \mathbf{w} e bias b devem minimizar a função de custo:

$$\frac{1}{2}\|\mathbf{w}\|^2 \quad (31)$$

Sujeito a

$$d_i(\mathbf{w}^T \mathbf{x}_i + b) \geq 1 \text{ para } i=1,2,\dots,N \quad (32)$$

O problema de otimização obtido é quadrático e podem ser solucionados com a introdução de uma função Lagrangeana, que engloba as restrições à função objetivo, associadas a parâmetros denominados multiplicadores de Lagrange α_i (VAPNIK, 1995). O problema de otimização é resolvido pela maximização da seguinte função:

$$Q(\alpha) = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j d_i d_j (\mathbf{x}_i \cdot \mathbf{x}_j) \quad (33)$$

Sujeito a:

$$\sum_{i=1}^N \alpha_i d_i = 0 \quad (34)$$

$$\alpha_i \geq 0 \quad (35)$$

Após a determinação dos multiplicadores de Lagrange ótimos ($\alpha_{o,i}$), o vetor de peso ótimo (\mathbf{w}_o) e o bias ótimo (b_o) são calculados usando as Equações 36 e 37, respectivamente:

$$\mathbf{w}_o = \sum_{i=1}^N \alpha_{o,i} d_i \mathbf{x}_i \quad (36)$$

$$b_o = 1 - \mathbf{w}_o^T \mathbf{x}^{(s)} \quad (37)$$

A função de separação, $f(x)$, que representa o hiperplano que separa os dados com maior margem é escrita como:

$$f(x) = \text{sign}\left(\sum_{i=1}^N d_i \alpha_i (\mathbf{x}_i \cdot \mathbf{x}) + b_o\right) \quad (38)$$

Em que sign é a função sinal.

2.3.4.2 Casos não linearmente separáveis

De acordo com VAPNIK (1995), para construir o hiperplano ótimo em casos onde os dados não são linearmente separáveis deve-se introduzir as variáveis de folga, $\{\xi_i\}_{i=1}^r$ na definição do hiperplano de separação e encontrar os valores ótimos do vetor de peso \mathbf{w} e do bias b de modo que satisfaçam a restrição:

$$d_i(\mathbf{w}^T \mathbf{x}_i + b) \geq 1 - \xi_i \quad (39)$$

$$\xi_i \geq 0, i = 1, \dots, N \quad (40)$$

As variáveis de folga ξ_i são responsáveis por medir a quantidade de violações das restrições (OSUNA et al., 1997). Estas violações podem ocorrer da seguinte forma: Para $0 \leq \xi_i \leq 1$, o ponto estará dentro da região de separação e do lado correto da superfície de decisão; Para $\xi_i > 1$ o ponto estará do lado incorreto do hiperplano separador e para $\xi_i = 0$ têm-se os vetores de suporte. A Figura 7 exemplifica os tipos de violações citados (HAYKIN, 1999).

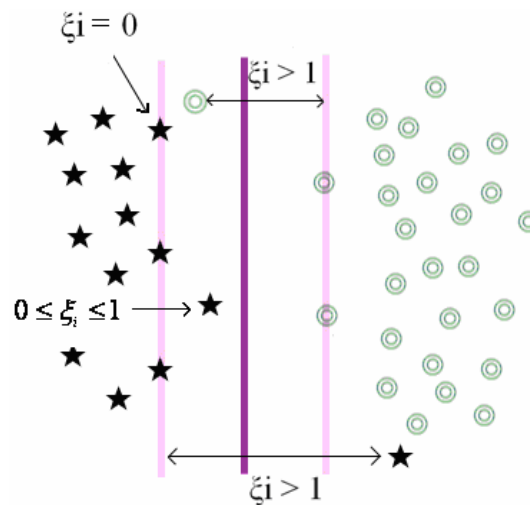


Figura 7 – Ilustração das posições das variáveis de folga
Fonte: Semolini, 2002

O vetor peso \mathbf{w} e as variáveis de folga ξ_i devem minimizar a seguinte função:

$$\Phi(\mathbf{w}, \xi) = \left[\frac{1}{2} \mathbf{w}^T \mathbf{w} + C \sum_{i=1}^N \xi_i \right] \quad (41)$$

Sujeito à:

$$d_i(\mathbf{w}^T \mathbf{x}_i + b) - 1 + \xi_i \geq 0 \quad (42)$$

$$\xi_i \geq 0 \quad i=1, \dots, N. \quad (43)$$

Onde C é o parâmetro de penalização.

O primeiro termo em 41 é o mesmo termo existente nos padrões linearmente separáveis e tem a função de controlar a capacidade de aprendizagem, enquanto que o segundo termo visa penalizar as amostras localizadas no lado incorreto do hiperplano. Este termo é visto como um parâmetro de regularização e deve ser definido pelo usuário. Quanto menor o valor de C mais rígida será a penalização dos pontos; caso contrário, a penalização será menos rígida e a classificação terá maior quantidade de erros (OSUNA et al., 1997).

Novamente o problema de otimização gerado é quadrático, com as restrições lineares apresentadas em 42 e 43. Esse problema de otimização é resolvido usando o método dos multiplicadores de Lagrange que maximizem a seguinte função objeto:

$$Q(\alpha) = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j d_i d_j (\mathbf{x}_i^T \mathbf{x}_j) \quad (44)$$

Sujeito às restrições:

$$\sum_{i=1}^N \alpha_i d_i = 0 \quad (45)$$

$$0 \leq \alpha_i \leq C \text{ para } i = 1, 2, \dots, N \quad (46)$$

A função objetivo $Q(\alpha)$ a ser maximizada é a mesma para os casos separáveis e não-separáveis. Porém, este último difere do caso separável pelo fato de que a restrição $\alpha_i \geq 0$ é substituída por uma mais rigorosa $0 \leq \alpha_i \leq C$. Os valores ótimos do vetor peso \mathbf{w} e do bias também são obtidos tal como nos casos linearmente separáveis (Equações 36 e 37) (HAYKIN, 1999).

A função de separação do hiperplano é escrita como (VAPNIK, 1995):

$$f(x) = \text{sign} \left(\sum_{i=1}^N d_i \alpha_i (\mathbf{x}_i \cdot \mathbf{x}) + b_o \right) \quad (47)$$

Uma máquina de vetor de suporte é construída, no caso de padrões não linearmente separáveis, projetando o conjunto de treinamento (\mathbf{x}) de seu espaço original, referenciado como de entradas, através de uma função de mapeamento (Φ), para o espaço característico de alta dimensionalidade (H) (Figura 8). Após essa projeção é construído o hiperplano ótimo que separa as classes (VAPNIK, 1995). A Figura 8 exemplifica um mapeamento não linear onde os dados de entrada são projetados no espaço característico e separados por um hiperplano.

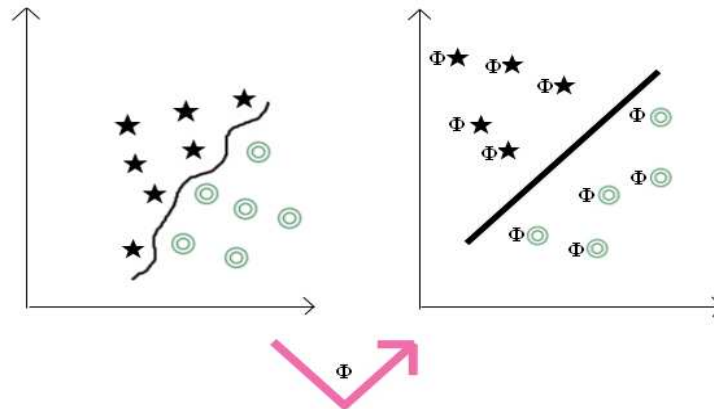


Figura 8 – Ilustração de mapeamento para o espaço característico
Fonte: Semolini, 2002

O teorema de Cover (1965) afirma que um espaço multidimensional pode ser transformado com alta probabilidade em um novo espaço de características onde os padrões são linearmente separáveis. Porém, duas condições devem ser satisfeitas: A transformação deve ser não linear e a dimensionalidade do espaço de características deve ser suficientemente alta.

Para realizar o mapeamento, aplica-se a função de mapeamento (Φ) aos exemplos presentes no problema de otimização apresentado na Equação 44, conforme ilustrado a seguir:

$$Q(\alpha) = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j d_i d_j (\Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x}_j)) \quad (48)$$

Sujeito às restrições apresentadas em (44) e (45).

Assim, têm-se a função de separação do hiperplano é reescrita como:

$$f(x) = \text{sign} \left(\sum_{i=1}^N d_i \alpha_i \Phi(\mathbf{x}_i) \cdot (\mathbf{x}_j) + b_o \right) \quad (49)$$

O algoritmo de aprendizagem por vetor de suporte pode ser usado para construir diversos tipos de máquina de aprendizagem, como por exemplo: máquinas de aprendizagem polinomial, redes de função de base radial e perceptrons de duas camadas. A construção destas dependerá de como os dados são projetados para o espaço característico (HAYKIN, 1999).

As funções usadas para projetar os dados do espaço de entrada para o espaço característico são chamadas de funções kernels. Em outras palavras, um kernel K é uma função que recebe dois pontos \mathbf{x}_i e \mathbf{x}_j do espaço de entradas computa o produto escalar desses dados no espaço de características (HAYKIN, 1999). Assim, têm-se a seguinte Equação:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x}_j) \quad (50)$$

De acordo com Chapelle et al. (1999), estas funções devem ser escolhidas cuidadosamente para evitar desempenhos ruins pelo classificador.

Algumas funções kernel comumente utilizadas (OSUNA et al., 1997), são: a função MLP:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \tanh(\mathbf{x}_i \mathbf{x}_j - \theta) \quad (51)$$

A função polinomial:

$$K(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i \mathbf{x}_j + 1)^d \quad (52)$$

Onde o parâmetro d (grau do polinômio) é especificado a priori pelo usuário.

E a função de base radial:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}\right) \quad (53)$$

Onde σ (interpretado como a largura da gaussiana) é especificado a priori pelo usuário.

Esta última função é uma das mais comumente aplicadas em dados provenientes do sensoriamento remoto, como pode ser observado nos trabalhos de: Pal e Mather (2004), Melgani (2004), Gonçalves et al. (2006) e Foody e Mathur (2004).

Com relação às funções kernel, as referências utilizadas por Canu et al. (2005) no desenvolvimento da SVM-KMToolbox estão em Vapnik (1995), Chapelle (1999), Scholkopf e Smola (2001), Rakotomamonjy e Canu (2002) e Canu et al. (2003).

O percentual de acertos de uma classificação pelo método SVM é dependente da magnitude do C (Equação 41) e, quando se faz uso da função kernel do tipo gaussiana, do parâmetro σ . Um ponto chave no sensoriamento remoto é que esta classificação seja feita a partir de uma determinada amostra de treinamento e que o aprendizado desta possa ser facilmente aplicado em outro conjunto de dados. A capacidade de generalização do método classificador escolhido é de fundamental importância, e só será obtida caso os valores destes parâmetros sejam determinados cuidadosamente (BELOUSOV et al., 2002).

Pal e Mather (2004) testaram o método SVM na classificação de uma imagem do satélite DAIS correspondente a uma região de La Mancha, Espanha. Foram utilizadas duas quantidades de amostras de treinamento diferentes: uma pequena (com 100 pontos) e uma média (com 500 pontos) para cada uma das 9 classes propostas. Essas amostras foram aplicadas à 13 combinações de número de bandas (de 5 à 65). O método testado teve seu desempenho comparado aos classificadores: máxima verossimilhança, árvore de decisão e rede neural tipo MLP. Para a SVM, foi utilizada função kernel gaussiana, e o melhor

resultado foi obtido quando os parâmetros C e σ foram definidos em 5000 e 2, respectivamente. A SVM apresentou resultados superiores tanto quando foram utilizados 100 e 500 pixels para todas as combinações de número de bandas propostas. Os autores ainda concluíram que, com o incremento do número de bandas, o classificador SVM tende a aumentar quase continuamente o valor da acurácia, o que não ocorreu nos demais classificadores avaliados. Um excelente valor de exatidão global obtido foi de 93,55%, 85,8%, 88,4% e 81,2% para SVM, máxima verossimilhança, MLP e árvore de decisão, respectivamente, para o teste realizado com 100 amostras e 65 bandas.

O método SVM tende a proporcionar melhores classificações do que a rede MLP e o da máxima verossimilhança. De acordo com Pal e Mather (2004), este método resulta em boas performances mesmo quando são apresentadas poucas amostras para realizar aprendizagem. Esta seria uma boa vantagem com relação a outros métodos, visto que o usuário não precisa gastar muito tempo definindo uma grande quantidade de amostras, tal como é necessário para o método da máxima verossimilhança.

Foody e Mathur (2004) classificaram uma imagem do satélite Airborne Thematic Mapper (ATM) correspondente à região de Feltweel, Inglaterra. Foram testados quatro métodos: um classificador probabilístico, uma árvore de decisão, uma MLP e uma SVM e avaliados o efeito do aumento do número de amostras de treinamento na acurácia. A maior exatidão global foi obtida para o método SVM, com 93,75% dos pixels classificados corretamente. A SVM não foi a mais rápida entre os métodos, porém, mais rápida do que a árvore de decisão e a MLP. Os testes também permitiram concluir o aumento do valor de exatidão global com o aumento do número de amostras. Os autores ainda concluíram que as SVMs possuem excelente potencial na classificação de imagens de satélite.

Su et al. (2007) avaliou a classificação de imagens provenientes do MISR (*Multi-angle Imaging Spectro-Radiometer*) através dos métodos máxima verossimilhança e SVM em uma região semiárida do Novo México, Estados Unidos. A exatidão global obtida para o método SVM foi de 76,7% e o de 68,1% para o método da máxima verossimilhança. Os valores do coeficiente de Kappa obtidos foram de 0,73 e 0,64 para o método SVM e máxima verossimilhança, respectivamente. Os autores consideram bons os resultados obtidos, já que os 19 tipos de vegetações propostos para a classificação exibiam características bastante diversificadas.

2.3.4.3 Classificadores binários e classificadores multiclases

As SVMs são originalmente classificadores binários, porém, a maioria dos dados reais envolve grande quantidade de informação e pode necessitar ser rotulada em mais de duas classes. Basicamente, duas estratégias são usadas para solucionar este tipo de problema: a estratégia OAO (*One-Against- One*) e a estratégia OAA (*One-Against-All*) (WESTON et al., 1999).

A estratégia OAO constrói $N(N-1)/2$ funções discriminantes, em que N é o número de classes. Um processo de escolha será aplicado para definir em que classe o dado x deverá ser classificado e a decisão será tomada pelo número de vezes que este foi associado à determinada classe pela função discriminante apresentada na Equação 27 (WESTON et al., 1999).

Na estratégia OAA cada SVM existente resolve um problema de separação entre duas classes. Por exemplo: em um determinado problema de classificação de padrões com N classes, serão criados N classificadores binários e cada um desses será treinado para distinguir uma classe das restantes, ou seja, das $N-1$ classes (WESTON et al. 1999). A decisão final sobre qual classe o dado irá pertencer é dada pela regra “o vencedor leva tudo”, ou seja, a classe vencedora é aquela em que o SVM possui maior saída, ou seja, possui maior valor de função discriminante, $f(x)$ (MELGANI, 2004). Maiores detalhes sobre a teoria das estratégias apresentadas podem ser encontrados nos trabalhos de Weston et al. (1999) e Hsu et al. (2002).

Gidudu et al. (2004), avaliaram o uso de estratégias multiclases na classificação de imagens da região de Jinja, Uganda. Os dados de entrada corresponderam às 7 bandas do satélite LANDSAT e as classes de saída foram água, vegetação e antropizada. Os testes realizados envolveram as estratégias one-against-one (OAO) e one-against-all (OAA) em quatro tipos de classificadores SVM, a saber: linear, quadrática, polinomial e gaussiana. A estratégia indicada foi a OAO já que a OAA teve maior propensão de gerar pixels não classificados ou erroneamente classificados. Para a SVM com classificador gaussiana, por exemplo, foram 103 pixels classificados erroneamente para a estratégia OAO e 4645 para a OAA.

Melgani (2004) utilizou uma imagem do sensor AVRIS correspondente uma região da Índia para a classificação por meio dos métodos SVM linear, SVM não-linear (com kernel tipo RBF), mínima distância e rede neural tipo Função de Base Radial (RBF). O estudo mostrou que a SVM não-linear apresentou valor de exatidão global de 92,64%, superando os

demais métodos. Para o método SVM não-linear foram realizados testes objetivando comparar as estratégias *multiclass* OAO e OAA. Os resultados apontaram melhor desempenho da estratégia OAO com exatidão global de 93,96% e menor custo computacional. Já a estratégia OAA obteve valor de exatidão global igual a 93,42%.

2.4 Distância de Jeffries-Matusita

Uma forma de avaliar uma classificação se dá por meio de medidas estatísticas. De acordo com Richards (1993), a distância de Jeffries-Matusita (JM) entre duas classes i e j que apresentem dados com distribuição normal é definida por:

$$JM_{ij} = 2 * (1 - e^{-B}) \quad (54)$$

Onde B é a distância de Bhattacharyya (B_{ij}) e calculada como:

$$B_{ij} = \frac{1}{8} (M_i - M_j)^T \left(\frac{\Sigma_i + \Sigma_j}{2} \right)^{-1} (M_i - M_j) + \frac{1}{2} \left(\frac{\left| \frac{\Sigma_i + \Sigma_j}{2} \right|}{\sqrt{|\Sigma_i| |\Sigma_j|}} \right) \quad (55)$$

Onde M_i e M_j correspondem aos vetores de média das classes i e j , respectivamente. Σ_i e Σ_j são as matrizes de covariância das classes i e j , respectivamente. Quanto maior o valor estimado por B , maior a separabilidade entre classes.

O valor de JM entre duas classes poderá variar entre 0 e 2. Caso o valor obtido seja igual a 2 significa que ocorreu uma separação completa entre os exemplos. Valores maiores que 1,8 indicam que determinado par de classes obteve boa separabilidade.

Kraisch et al. (2007), fizeram uma avaliação temporal do uso e ocupação do solo usando imagens do satélite LANDSAT dos anos de 1986, 2000 e 2004 da região de Itajaí, Santa Catarina. Foram definidas quatro classes: área urbana, floresta, água, campos e não classificada. A separabilidade entre as classes propostas foi analisada por meio do índice de Jeffries-Matusita que, para todas as combinações, foi superior a 1,8. Os autores concluíram que as amostras escolhidas contribuíram para a eficiente estratificação temática e acompanhamento do uso e ocupação do solo proposto no estudo.

2.5 Validação da exatidão do mapeamento

Um mapeamento do uso do solo, apoiado em sensoriamento remoto, necessita ser realizado com exatidão quando se deseja que as informações obtidas neste sejam confiáveis. Assim, parâmetros que indiquem que uma classificação quantificou determinada área próxima à realidade de seu uso são essenciais. A avaliação da acurácia pode ser obtida por meio dos coeficientes de concordância obtidos através de uma matriz de confusão.

A matriz de erro ou matriz de confusão é uma matriz quadrada de números que expressam a quantidade de unidades amostrais, associada a uma dada categoria durante o processo de classificação efetuado, e à categoria real a que pertencem essas unidades. (MANGABEIRA, 2003 apud CONGALTON 1991).

Um exemplo de uma matriz de confusão está exposto na Figura 9, em que se observam três classes de uso do solo X, Y e Z, confrontados com seus dados de referência (ou seja, dados da verdade de campo) e seus dados de classificação.

	Classes	Dados de referência			Total das Linhas
		X	Y	Z	
Classificação	X	15	2	4	21
	Y	3	12	2	17
	Z	1	3	14	18
Total das colunas		19	17	20	56

Figura 9 - Exemplo de uma matriz de confusão

A matriz de confusão restringe-se às linhas e colunas referentes às classes X, Y e Z. Os componentes da diagonal principal da matriz de confusão fornecem o número de pixels corretamente classificados para cada classe correspondente. Por exemplo, para a classe Y, nos dados de classificação foram corretamente classificados 12 pixels. No entanto, para os 3 pixels dessa mesma classe Y no mapa temático, analisando os dados de referência, constatou-se que na realidade eles pertenciam a classe X. Bem como outros 2 pixels que na realidade são da classe Z.

A partir da elaboração da matriz de confusão as seguintes medidas podem ser extraídas: Exatidão Global (EG), Exatidão Específica (EE), coeficiente de Kappa, dentre outros.

A EG é obtida pela divisão das somas das entradas que formam a diagonal principal da matriz, ou seja, o número de classificações corretas, pelo número total de amostras utilizadas como mostra a Equação 56 (WATZLAWICK, 2003).

$$EG = \frac{\sum_{i=1}^r x_{ii}}{N} * 100 \quad (56)$$

Onde, EG = Exatidão Global; r = número de linhas da matriz de erro; x_{ii} = Valor da linha i e coluna i ; N = Quantidade total de pontos amostrais.

A EE é obtida pela divisão no número total de amostras classificadas corretamente na classe pelo número total de amostras desta.

$$EE = \frac{x_{ii}}{n} * 100 \quad (57)$$

Onde, EE = Exatidão Específica; x_{ii} = valor na linha i e coluna i e n = quantidade total de pontos amostrais da classe r .

A análise de Kappa, diferentemente da EG, utiliza todos os elementos da matriz de confusão para seu cálculo. O coeficiente Kappa (K) é uma medida da concordância real (indicado pelos elementos diagonais da matriz de confusão) menos a concordância por chance (indicado pelo produto total da linha e coluna) (MOREIRA, 2005), conforme segue:

$$K = \frac{N \sum_{i=1}^r x_{ii} - \sum_{i=1}^r (x_{i+} x_{+i})}{N^2 - \sum_{i=1}^r (x_{i+} x_{+i})} \quad (58)$$

Onde, K = coeficiente Kappa de concordância; N = Quantidade total de pontos amostrais; r = número de linhas da matriz de erro; x_{ii} = valor na linha i e coluna i ; x_{i+} = soma da linha i ; x_{+i} = soma da coluna i .

A Tabela 1, desenvolvida por Landis e Koch (1977), é utilizada como referência para classificar mapas resultantes da utilização de imagens provenientes do sensoriamento remoto. Na referida Tabela, são apresentados níveis de desempenho da classificação que indicam a qualidade do mapa temático de acordo com o valor de Kappa obtido.

Tabela 1 – Qualidade da classificação associada aos valores da estatística de Kappa

Valor de Kappa	Qualidade do mapa temático
<0,00	Péssima
0,00-0,20	Ruim
0,20-0,40	Razoável
0,40-0,60	Boa
0,60-0,80	Muito boa
0,80-1,00	Excelente

FONTE: adaptada de Landis e Koch (1977)

Para realizar a validação da classificação dos dados provenientes do sensoriamento remoto a maioria dos autores utiliza o coeficiente de Kappa, como pode ser visto nos estudos de Nishida (1998) e Queiroz et al.(2004).

2.6 Comentários finais

Neste capítulo foi descrita toda a teoria de fundamental importância para o desenvolvimento deste trabalho. Os métodos relacionados à classificação de imagens de satélite situam o mesmo no contexto de reconhecimento de padrões. As teorias dos métodos máxima verossimilhança, MLP e SVM fornecem subsídios para os experimentos realizados neste trabalho e descritos no capítulo 3.

3. MATERIAL E MÉTODOS

Neste capítulo é detalhada a metodologia aplicada na abordagem proposta para classificação de imagens do satélite LANDSAT-5. É apresentada a definição das classes de interesse assim como a classificação pelo método da máxima verossimilhança e os testes realizados para os métodos baseados em aprendizado de máquina. Também são apresentados detalhes sobre o processo de validação.

3.1 Localização, aspectos fisiográficos e climáticos da área de estudo

A área de estudo está localizada em maior parte no município de Iguatu-CE. Este, por sua vez, encontra-se inserido na bacia hidrográfica do Alto Jaguaribe. A referida bacia hidrográfica seria, inicialmente, a área escolhida para a aplicação dos métodos propostos, porém, optou-se por limitá-la visando facilitar o andamento dos testes do ponto de vista de custo computacional, já que estes demandam grande quantidade de tempo e utilizar esta área seria inviável. A referida bacia possui uma área de 25.895,7 km² e importância significativa para o Estado do Ceará já que possui uma capacidade de armazenamento de 2.792.563,000 m³ de água e dois importantes açudes: o Orós e o Trussu (COGERH, 2005). Na Figura 10, o desenho esquemático expõe todas as bacias do Estado do Ceará destacando, na cor azul, a do Alto Jaguaribe. Na mesma Figura também está destacado, na cor cinza, o município de Iguatu, local de realização do experimento.

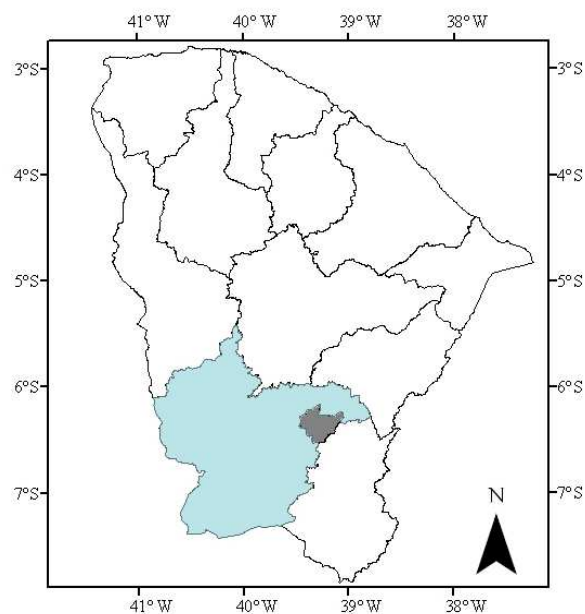


Figura 10 – Localização da bacia hidrográfica do Alto Jaguaribe

A bacia do Alto Jaguaribe possui vegetação típica da caatinga, relevo plano, suave ondulado e ondulado (JACOMINE et al., 1973) e está inserida nas condições de clima semi-árido predominante no Estado do Ceará. Pela classificação de Koppen (1948), no município de Iguatu, o clima é do tipo BSw'h': clima quente com chuvas de outono e temperaturas médias superiores a 18°C. A precipitação mensal no ano de 2008 deste município está apresentada na Figura 11 tendo, para o referido ano, precipitação acumulada de 1383,5 mm (FUNCEME, 2009). Por meio desta ainda é possível constatar que, para o ano em estudo, a maior precipitação ocorreu no mês de março com 525,9 mm.

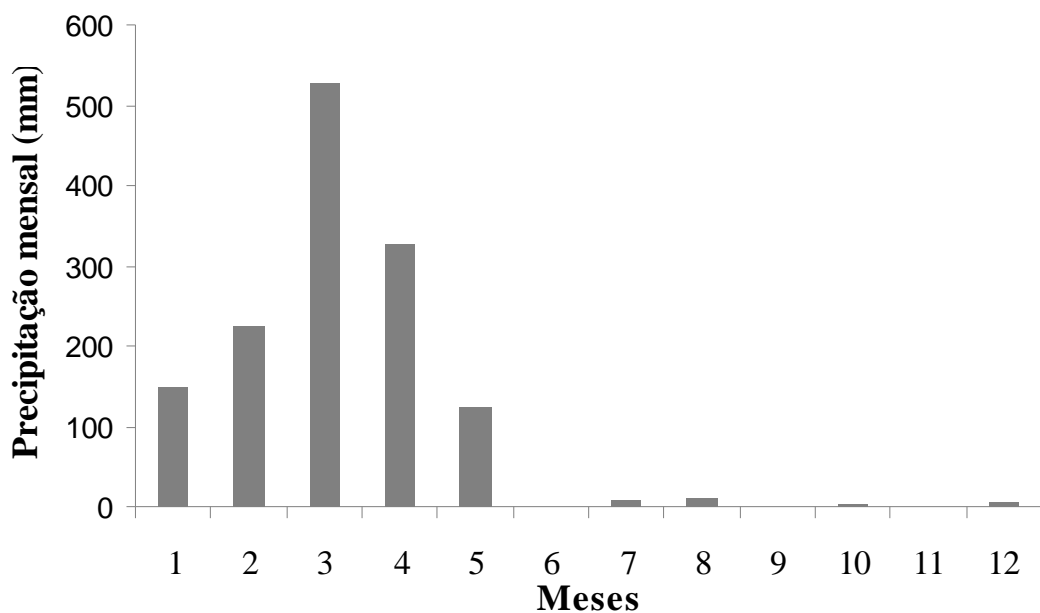


Figura 11 – Precipitação mensal do ano de 2008 para o município de Iguatu-CE.

O município de Iguatu está localizado entre os paralelos 6° 17' 42,33''S e 6° 33' 41,04''S e os meridianos 39° 3' 55,64'W e 39° 29' 28,52''W. Já a parcela escolhida para teste tem a maior parte de sua área inserida neste município e esta localizada entre os paralelos 6° 19' 29,91''S e 6° 25' 58,98''S e os meridianos 39° 14' 53,29''W e 39° 21' 22,89'W e exposta na Figura 12.

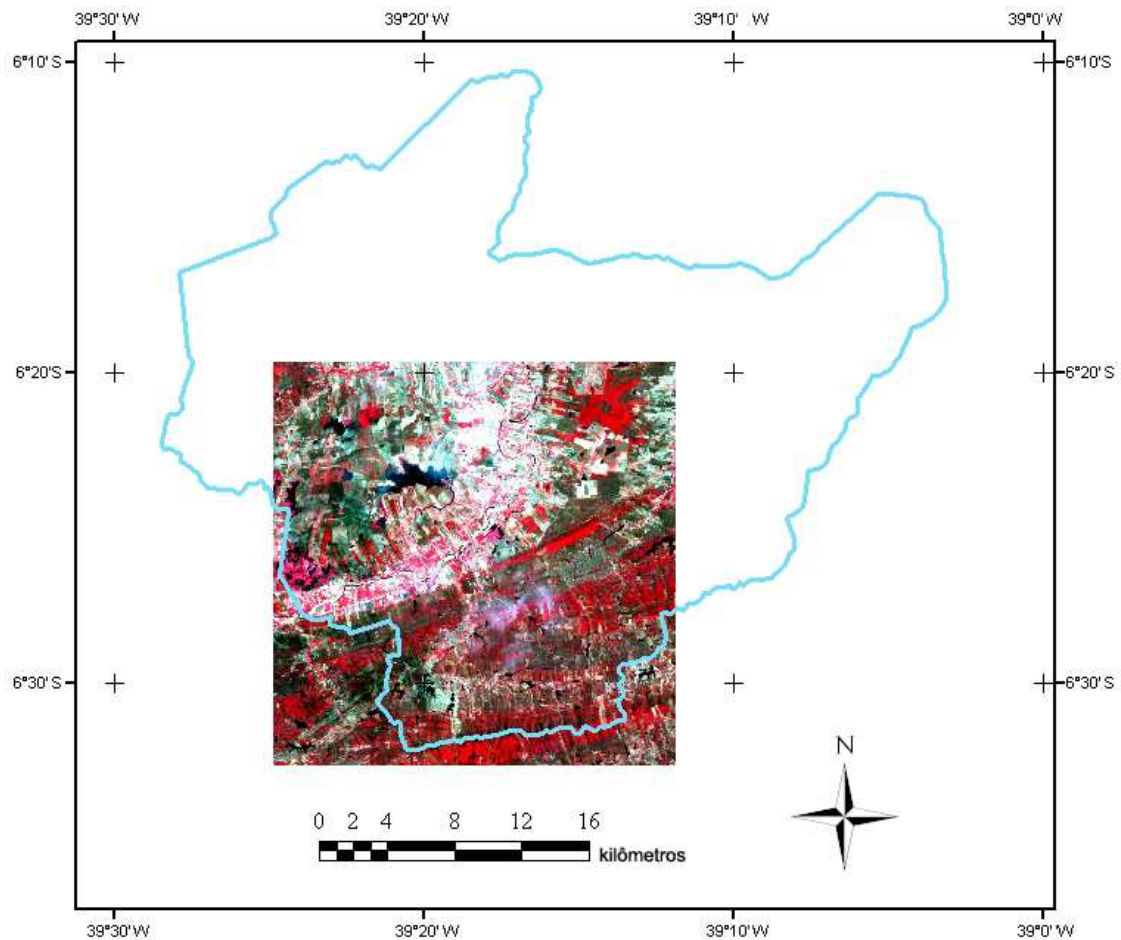


Figura 12 – Município de Iguatu e área de estudo representada pela imagem do satélite LANDSAT-5 de 20/08/08.

3.2 Configuração da máquina

O experimento foi realizado com o uso dos seguintes recursos de *hardware*: processador Intel *Pentium Duo*, 3 GHz, Placa mãe Intel Corporation D945GTP (Bus type: Intel NetBurst; bus width:64-bit; real clock: 200 MHz; effective clock: 800 MHz) , Memória RAM de 1 GB, Disco rígido de 80 GB e sistema Operacional *Windows XP*.

3.3 Aquisição da imagem e georreferenciamento

A área de estudo foi extraída de uma cena do satélite LANDSAT-5 (217/65) referente ao dia 20 de agosto de 2008 às 9 horas. A cena foi cedida gratuitamente pelo site do Instituto Nacional de Pesquisas Espaciais INPE (www.inpe.br). O LANDSAT-5 leva a bordo o sensor TM e contribui para o mapeamento temático da superfície terrestre. Este sensor

possui sete bandas espectrais entre 0,45 a 2,35 μm . Optou-se por utilizar a combinação das bandas 2 (0,50 - 0,60 μm), 3 (0,63 - 0,69 μm) e 4 (0,76 - 0,90 μm) que corresponde ao verde, vermelho e infravermelho, respectivamente. Estas bandas foram apresentadas na combinação 432 do espaço RGB. O satélite funciona a uma altitude de 705 km e possui uma resolução temporal de 16 dias.

Os pontos georreferenciados em campo por meio do aparelho Leica GPS1200 foram utilizados no software ENVI 4.3® para reajustar o georreferenciamento da imagem do satélite LANDSAT-5. Para este processo foram coletados pontos conhecidos e facilmente visualizados na imagem, como: cruzamentos de rodovias e paredes de açudes.

3.4 Escolha das classes e processo de amostragem

A área escolhida para estudo foi classificada em cinco classes genéricas, a saber:

1 - Classe água: Abrange rios, açudes, lagoas, etc. A ilustração desta classe está na Figura 13;



Figura 13 – Classe água entre a Caatinga

2 - Classe Caatinga Arbórea Densa (CAD): Engloba a vegetação arbórea densa, de porte mais elevado (8 - 12m). Nas regiões de serra observa-se uma vegetação com característica mais exuberante, onde as condições climáticas favorecem maior vigor na vegetação. Também está presente nas regiões do interior mais planas e mais secas apresentando uma leve diferença de tonalidade. A Figura 14 mostra uma região no município de Iguatu com CAD.



Figura 14 – Classe CAD no município de Iguatu-CE

3 - Classe Caatinga Herbácea Arbustiva (CHA): Esta classe engloba a vegetação herbácea arbustiva (porte baixo a médio) aberta a densa. Esta é a classe de maior vulnerabilidade à atuação antrópica, e sofrendo este processo pode passar a pertencer à classe antropizada ou agricultura. A Figura 15 mostra a classe CHA.



Figura 15 – Classe CHA no município de Iguatu-CE

4 - Classe antropizada: abrange as áreas onde não há cobertura vegetal. São exemplos: áreas desmatadas, estradas, aglomerados urbanos e área de expansão urbana, áreas de culturas abandonadas ou desestruturadas, áreas para uso da pecuária, solo exposto, etc. A Figura 16 mostra uma área apresentando processo de antropização.



Figura 16 – Exemplo de área antropizada

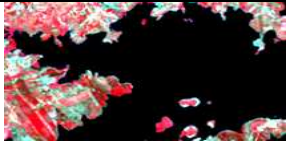




5- Classe agricultura: engloba áreas utilizadas na atividade agrícola, sendo caracterizada na imagem por polígonos de lados bem definidos. A Figura 17 expõe uma área plantada com arroz irrigado no município de Iguatu-CE.



Figura 17 – Área de agricultura no município de Iguatu-CE

As classes escolhidas para classificação apresentam-se na imagem, na composição de bandas escolhida, padrões de cores característicos. A Tabela 2 sintetiza as informações de cor e o padrão que cada classe se apresenta na imagem.

Tabela 2 – Classes de interesse na imagem do satélite LANDSAT-5 na combinação 432 do espaço RGB

Classe	Cor (composição colorida)	Padrão na imagem
Água	Coloração Preta ao azul escuro.	
CAD	Coloração vermelho intenso.	
CHA	Coloração verde claro a escuro.	
Antropizada	Coloração branca amarelada e azulada.	
Agricultura	Coloração rosa claro.	

Para o treinamento dos métodos testados (máxima verossimilhança, MLP e SVMs), realizou-se o processo de amostragem no software ENVI 4.3® e posteriormente os dados foram exportados para um formato legível do MATLAB 7.0 (*.txt), objetivando utilizá-los nos métodos baseados em aprendizado de máquina e, desta forma, evitar tendenciosidade nas comparações. Foram selecionadas 2000 amostras para cada classe. É importante salientar que estas amostras foram obtidas na mesma cena, mas fora da área de estudo, o que possibilita testar a capacidade de generalização dos classificadores.

3.5 Análise da separabilidade das amostras de treinamento

Objetivando analisar a resposta espectral pelas bandas 2, 3 e 4 de cada classe e a separabilidade entre estas, foi calculado o índice de Jeffries-Matusita pela Equação 54. Para o cálculo deste índice, os dados de entrada necessários são M_i e M_j correspondem aos vetores de média das classes i e j , respectivamente e C_i e C_j que correspondem as matrizes de covariância

das classes i e j . O software ENVI 4.3® dispõe de uma ferramenta que calcula este índice, restando ao usuário fornecer as amostras de treinamento utilizadas na classificação. O valor 1,8 considerado o limiar entre a separabilidade e não separabilidade entre as classes. Um detalhamento em português sobre o procedimento deste passo pode ser encontrado em SULSOFT (2009).

3.6 Classificação no software ENVI 4.3®

A partir das amostras de treinamento previamente definidas, foi realizada a classificação supervisionada pelo método estatístico da máxima verossimilhança no software ENVI 4.3®. Este é um procedimento simples e rápido, já que o software utilizado apresenta uma interface amigável ao usuário. Da mesma forma que o processo para o cálculo do índice de Jeffires-Matusita, um passo a passo do procedimento de classificação neste software pode ser encontrado em SULSOFT (2009).

3.7 Classificação no software MATLAB

O **MATLAB** é um sistema iterativo baseado em matrizes, para cálculos científicos e de engenharia. O nome MATLAB em si é a abreviação de **MATrix LABoratory** (Laboratório de Matrizes). Neste experimento, esse software foi utilizado para realizar a classificação das imagens usando os métodos MLP e SVM.

3.7.1 Método Perceptron de Múltiplas Camadas (MLP)

A rede Perceptron com uma única camada escondida e alimentação *feedforward*, treinada com algoritmo *backpropagation*, foi implementada em ambiente MATLAB, usando a Toolbox Neural Network (The MathWorks Inc., Natick, MA, USA) versão 6.0.

Para definir o número de neurônios na camada escondida foram realizadas simulações, durante a etapa de treinamento, e posterior análise do desempenho da rede treinada sobre a imagem da área de estudo, com a finalidade de encontrar uma arquitetura de rede que representasse melhor as classes de saídas pré-definidas.

Como os dados a serem classificados (dados de entrada) e as classes de interesse (saída) foram definidos inicialmente (número de neurônios na camada de entrada igual a 3 e na camada de saída igual a 5), os testes realizados objetivaram encontrar o número de

neurônios na camada escondida que proporcionassem melhor desempenho na classificação. A Tabela 3 apresenta as arquiteturas testadas na rede MLP.

Tabela 3 – Testes realizados na rede MLP

TESTE	BANDAS	Arquitetura da rede
1	234	3-12-5
2	234	3-15-5
3	234	3-18-5

Quanto ao item arquitetura da rede apresentado na Tabela 3 tem-se que: a quantidade de neurônios na camada de entrada (três) corresponde ao número de bandas consideradas na classificação visto que, para cada neurônio de entrada é atribuído o valor da resposta espectral do *pixel*, do conjunto de treinamento, em cada banda. Na camada intermediária foram testados de 12, 15 e 18 neurônios. Já o número de neurônios na camada de saída (cinco) é correspondente ao número de classes definidas. A Figura 18 exemplifica a arquitetura da rede neural correspondente ao teste 1.

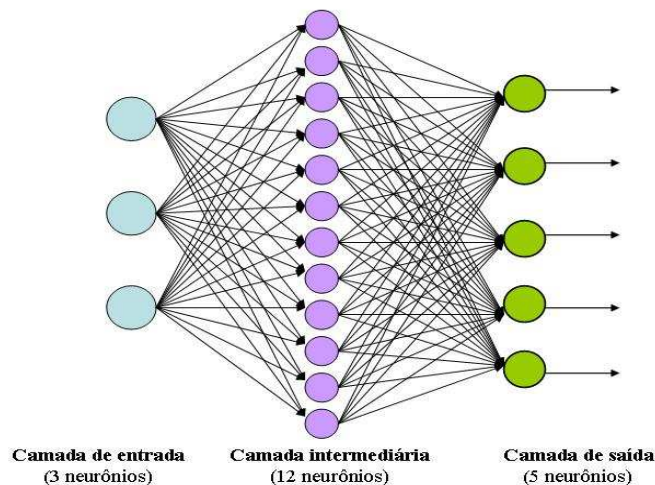


Figura 18 – Arquitetura da MLP correspondente ao teste 1

Para a fase de treinamento foram utilizadas 2000 amostras no formato *.txt para cada classe, definidos 3000 ciclos, erro final desejado de 0,001, taxa de aprendizagem (η) de 0,005 e função de ativação do tipo tangente hiperbólica (Equação 14) a qual limita os valores de saída de -1 à 1. Após a etapa de treinamento, a imagem correspondente à área de estudo foi transformada para o vetor e , em seguida, submetida ao processo de classificação.

O fluxograma apresentado na Figura 19 ilustra as etapas realizadas neste método em ordem cronológica e a listagem do algoritmo utilizado para o método MLP está no apêndice 1.

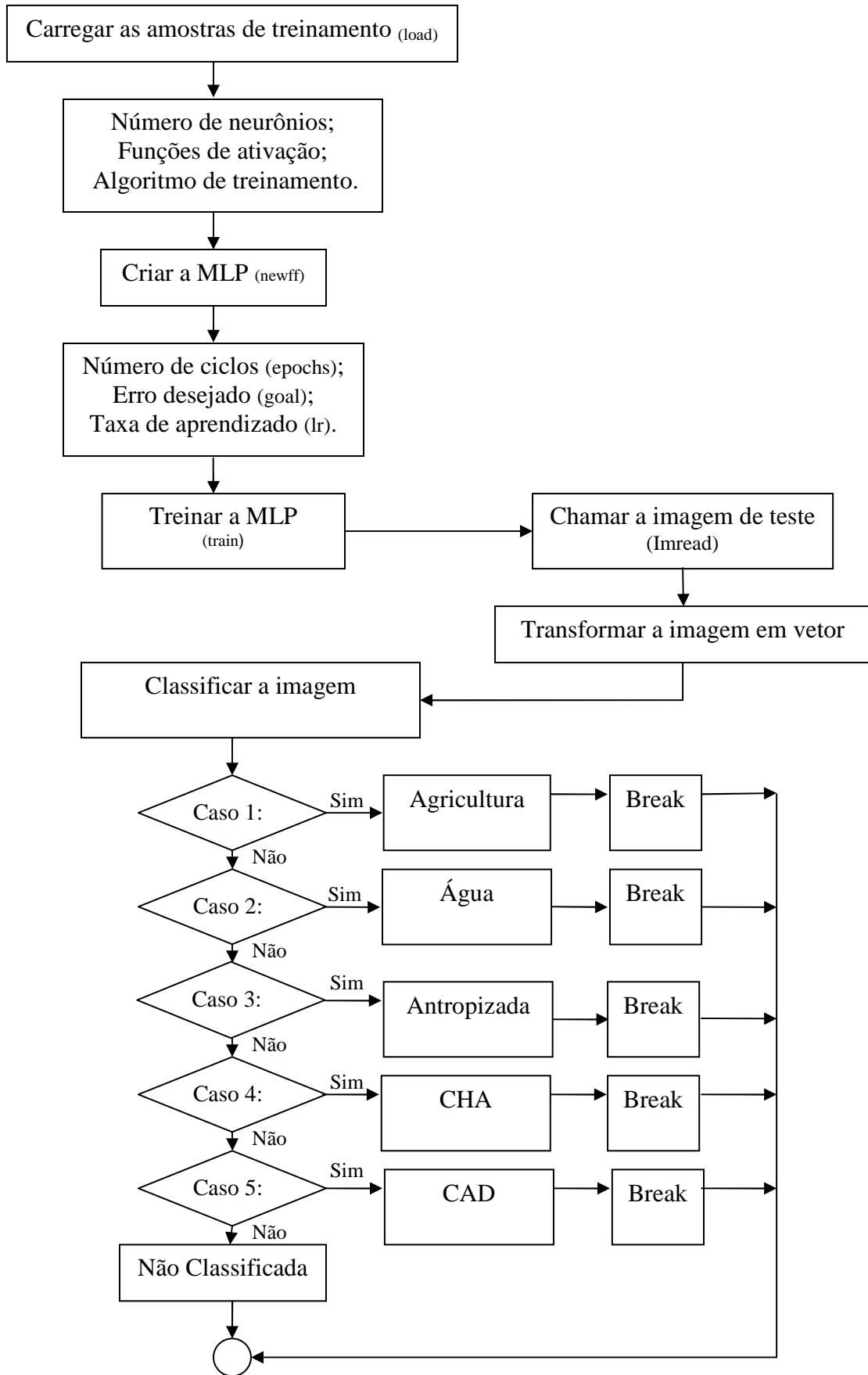


Figura 19 – Fluxograma das atividades relacionadas ao método MLP

3.7.2 Método Máquina de Vetor de Suporte (SVM)

Para a implementação da SVM no software MATLAB 7.0 foi utilizada a SVM-KMToolbox (*SVM and Kernel Methods MATLAB Toolbox*) desenvolvida por Canu et al. (2005).

Para a classificação das cinco classes propostas adotou-se a estratégia multiclasse one-against-one (OAO) pelos resultados superiores desta estratégia obtidos nos estudos de Gidudu et al. (2004) e Melgani (2004) na classificação de imagens de satélite e a função kernel do tipo gaussiana (Equação 53).

Os 22 testes foram realizados por meio combinações do parâmetro de penalização C nos valores de 10; 50; 100; 500; 1000; 1300; 1400; 1500 e 3000 com o parâmetro σ da função gaussiana nos valores de 1,5; 1,8; 2,2; 2,0; 2,5; 5 e 10. A quantidade de amostras disponível ao treinamento da rede foi de 2000 pixels para cada classe.

Os tempos de treinamento e teste, tanto quando foi realizada a classificação pelo método MLP como pelo método SVM, foram cronometrados visando observar o custo computacional de cada um dos métodos. Utilizou-se o comando tic e toc do MATLAB, o qual permite obter o tempo computacional gasto em cada ciclo. Os resultados estão apresentados nas sessões 4.4 e 4.5.

O fluxograma apresentado na Figura 20 ilustra as etapas realizadas neste método em ordem cronológica e a listagem do algoritmo utilizado para o método MLP está no apêndice 2.

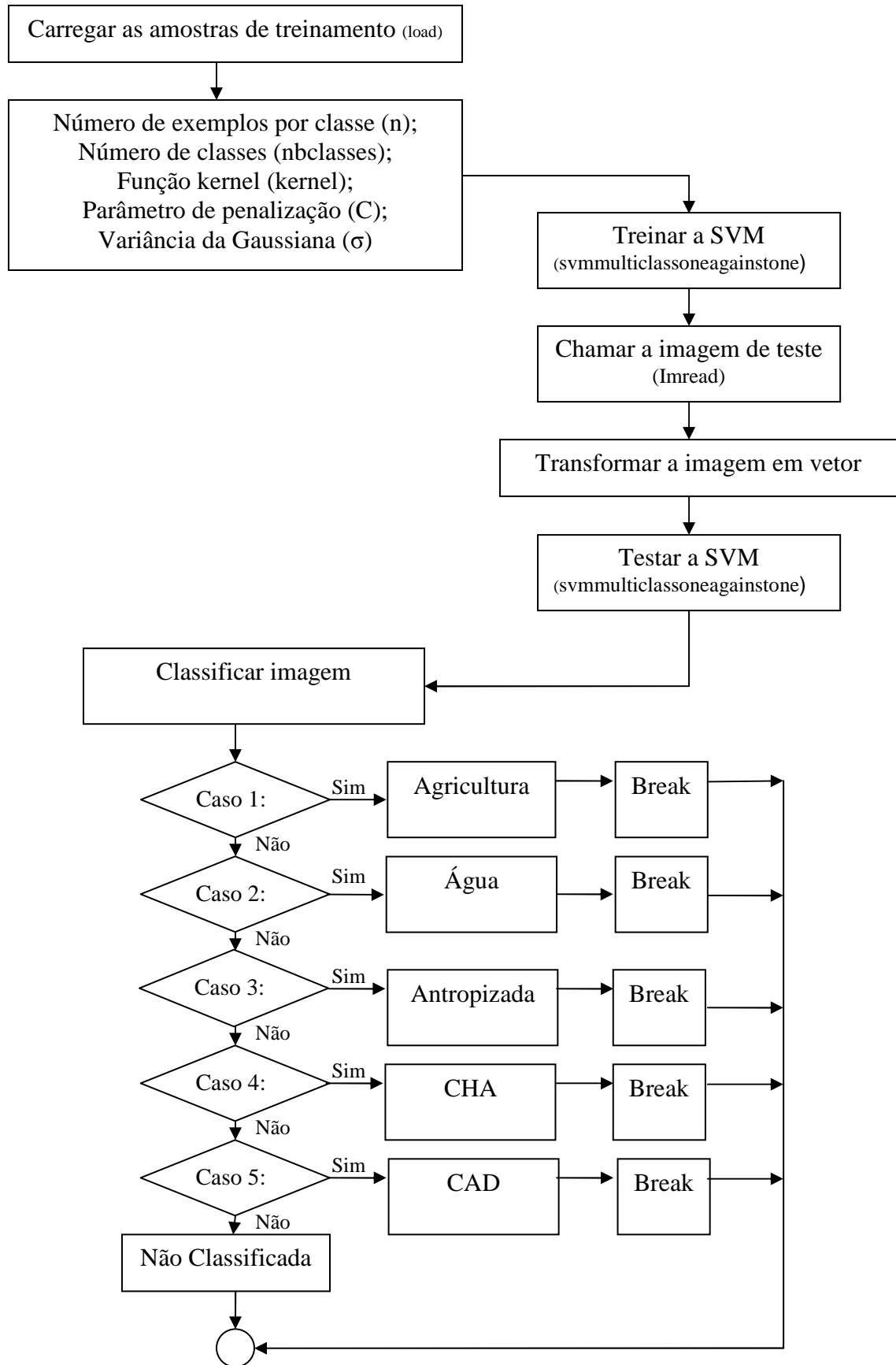


Figura 20 – Fluxograma das atividades relacionadas ao método SVM

3.8 Validação da classificação

Objetivando estabelecer uma análise quantitativa da acurácia obtida pelos métodos utilizados neste experimento, foi realizado o processo de validação por meio da construção de uma matriz de confusão utilizando-se 112 pontos de testes provenientes do georreferenciamento em campo, sendo estes diferentes dos 2000 pontos amostrais que foram utilizados para o treinamento da rede. O mesmo pode ser resumido nos passos a seguir:

- 1) Georreferenciamento dos pontos correspondente à verdade terrestre por meio de aparelho GPS Garmim em uma viagem à área de estudo entre os dias 29 à 31 de julho de 2008 (Figura 21).



Figura 21 – Georreferenciamento de ponto correspondente à verdade terrestre

- 2) Transferência dos 112 pontos correspondentes a verdade terrestre e georreferenciados com GPS Garmim para o software ArcGis 9.3® (Figura 22a) e comparação destes à classificação realizada pelo método da máxima verossimilhança (Figura 22b).

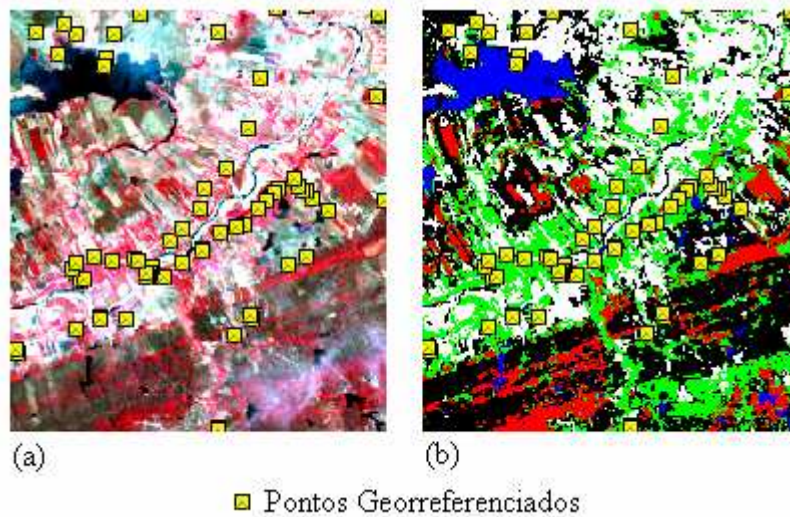


Figura 22 – (a) Pontos georreferenciados na imagem original; (b) Pontos georreferenciados na imagem classificada pelo método da máxima verossimilhança

3) Coleta dos valores de RGB, no software ArcGis 9.3, correspondente ao pixel georreferenciado em campo e armazenamento destes em formato *.txt. A Figura 23 mostra os valores de RGB para um pixel da classe água.

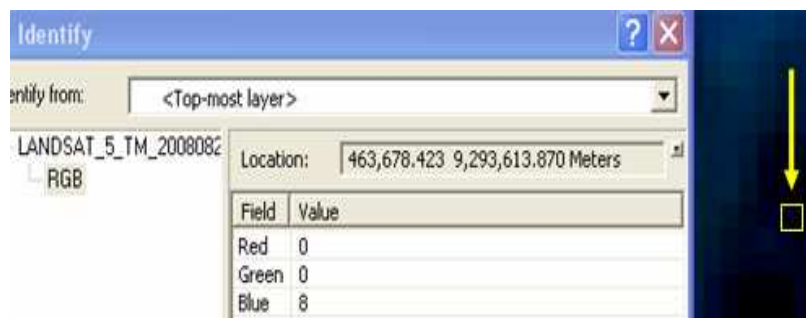


Figura 23 – Valores de RGB em pixel da classe água

4) Teste dos métodos MLP e SVM no software MATLAB, usando como dado de entrada os valores RGB coletados no passo 3. Na Figura 24, estão expostos os dados de entrada correspondentes aos valores de RGB e suas saídas após processamento no software MATLAB.



Figura 24 – Dados de entrada RGB e suas respectivas classes após processamento no MATLAB

5) Comparação das saídas resultantes da fase de teste ao resultado desejado, ou seja, a verdade terrestre para os métodos MLP e SVM (Figura 25).

		R	G	B	SAIDA DESEJADA	Teste 2	
pt 40	AGRI	255	63	100	1	1	
pt 41	AGRI	245	73	79	1	1	
pt 42	AGRI	255	14	98	1	1	
pt 43	AGRI	207	29	46	1	5	Pixel classificado errado! Classe 1 classificada como 5.
pt 56	AGRI	255	76	217	1	1	
pt 57	AGRI	255	52	212	1	1	
pt 58	AGRI	255	127	173	1	1	
pt 59	AGRI	255	78	92	1	1	
pt 60	AGRI	255	126	143	1	1	Pixel classificado correto! Classe 1 classificada como 1.
pt 61	AGRI	255	144	226	1	1	

Figura 25 – Planilha com o teste e a saída desejada

6) Quantificação dos erros e acertos gerados pelo processamento produzindo, desta forma, a matriz de confusão para os métodos máxima verossimilhança, MLP e SVM (Figura 26).

Ponto água contabilizado:	Verdade de Campo					Total Linhas
	Água	Agricultura	Antropizada	CHA	CAD	
Água	1					1
Agricultura						0
Antropizada						0
CHA						0
CAD						0
Classificação						
Total Colunas	1	0	0	0	0	1

Figura 26 – Produção da matriz de confusão

7) Calcular os coeficientes de exatidão global (EG), exatidão específica (EE) e Kappa (K) através das Equações 54, 55 e 56, respectivamente.

3.9 Fluxograma

A Figura 27 ilustra as principais etapas deste experimento com a condição cronológica em que o experimento foi executado.

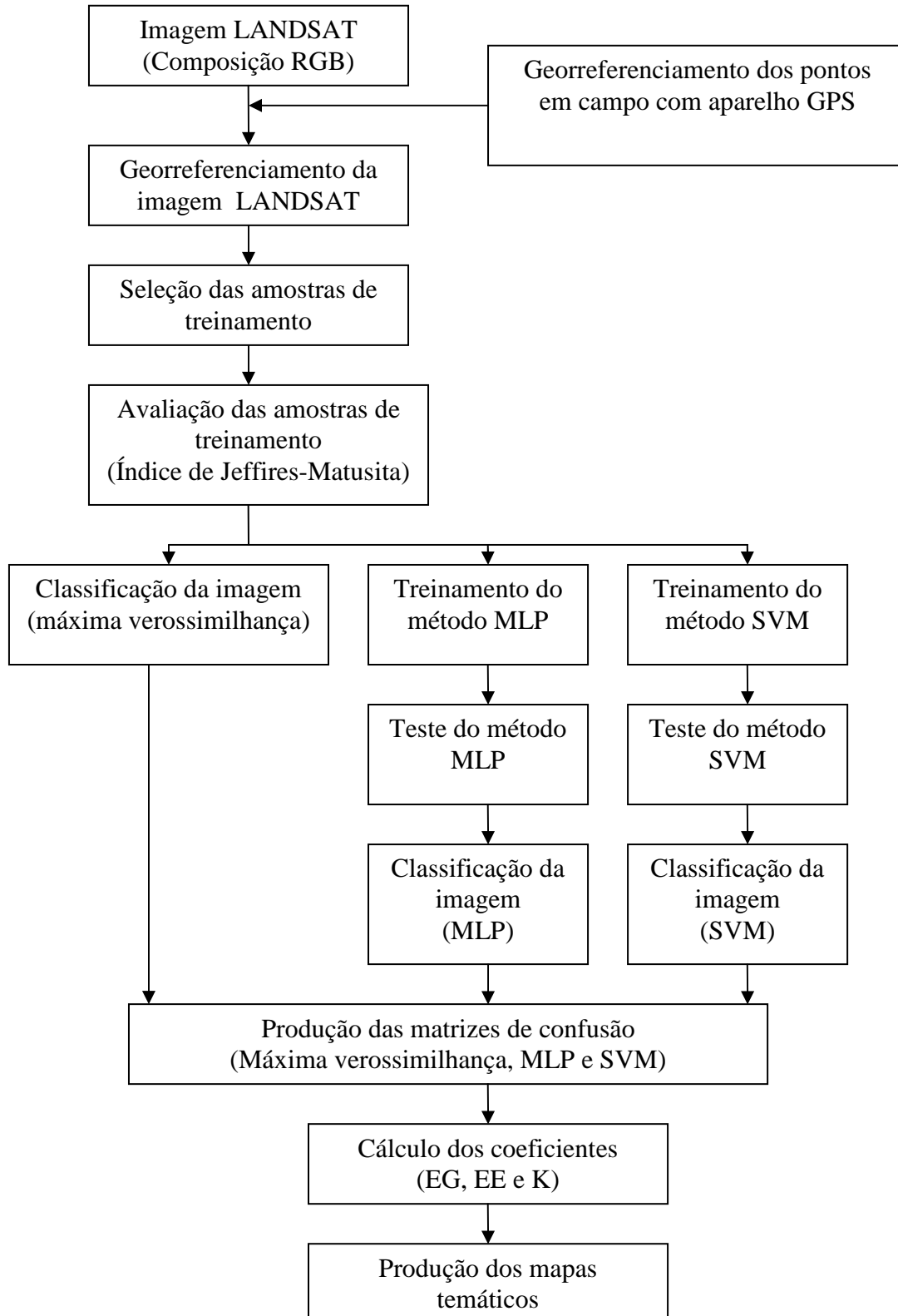


Figura 27 – Etapas e ordem de execução

3.10 Comentários Finais

Neste capítulo foram apresentados os métodos propostos para a classificação de imagens provenientes do satélite LANDSAT-5, assim como o método proposto para a validação da classificação. A validação da classificação é de fundamental importância, já que aponta para a qualidade do mapa resultante. No capítulo 4 serão expostos e discutidos os resultados obtidos.

4. RESULTADOS E DISCUSSÃO

Os resultados da metodologia proposta serão abordados neste capítulo, assim como a discussão destes diante dos objetivos propostos.

4.1 Análise da separabilidade das classes

A resposta espectral das amostras de treinamento foi analisada objetivando avaliar a separabilidade entre as classes. A Figura 28 expõe a resposta espectral média das amostras de treinamento das classes água, antropizada, agricultura, Caatinga Herbácea Arbustiva (CHA) e Caatinga Arbórea Densa (CAD) nas bandas 2, 3 e 4 do satélite LANSAT-5.

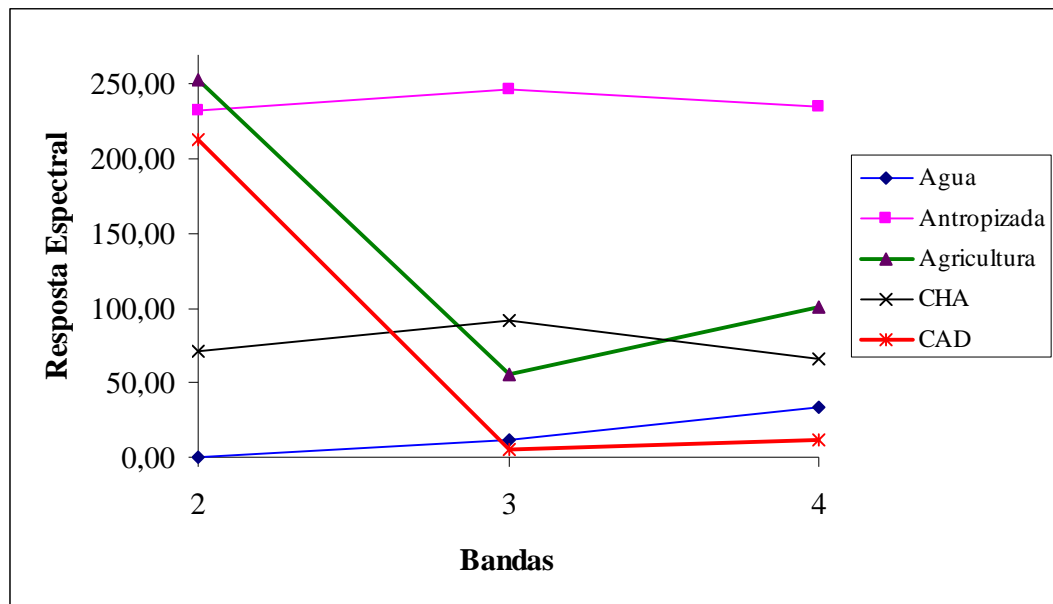


Figura 28 – Resposta espectral média das diferentes classes de cobertura do solo das amostras de treinamento

Por meio do gráfico apresentado na Figura 28 é possível notar que, em geral, as classes possuem respostas espectrais distintas, o que pode proporcionar melhores resultados na classificação dos dados. Nota-se ainda que as classes agricultura e CAD formam o par de classes que apresenta valores de resposta espectral média mais próximos, o que poderá interferir na separabilidade entre estas. Essa similaridade pode ter ocorrido devido aos altos índices de precipitação ocorridos neste ano, como apresentado na sessão 3.1.

A partir da identificação da assinatura espectral foi calculado a separabilidade entre os alvos terrestres. A Tabela 4 mostra as comparações entre os pares de classes e seus respectivos índices de Jeffries-Matushita. O valor de limiar que indica separabilidade ou não das classes é de 1,8.

Tabela 4 – Índice de separabilidade entre as classes

	Água	Antropizada	Agricultura	CHA	CAD
Água	-	-	-	-	-
Antropizada	2,0	-	-	-	-
Agricultura	2,0	1,99	-	-	-
CHA	1,97	1,98	1,99	-	-
CAD	1,99	2,0	1,89	1,99	-

Assim como no trabalho de Kraisch et al. (2007), todos pares de classes apresentaram valores do índice de Jeffries-Matusita acima do limiar, o que aponta para uma boa separabilidade das classes pelo ponto de vista da análise espectral (Tabela 4). Os pares de classes água e agricultura, água e antropizada e antropizada e CAD, apresentaram o máximo valor deste índice sendo, desta forma, os mais separáveis. Nota-se ainda que o menor valor de JM obtido, de 1,89, ocorreu para o par de classes agricultura e CAD.

4.2 Classificação pelo método estatístico da máxima verossimilhança

A classificação supervisionada pelo método da máxima verossimilhança foi utilizada como referência aos métodos MLP e SVM. O mapa resultante com as classes CHA, CAD, água, antropizada e agricultura está exposto na Figura 29.

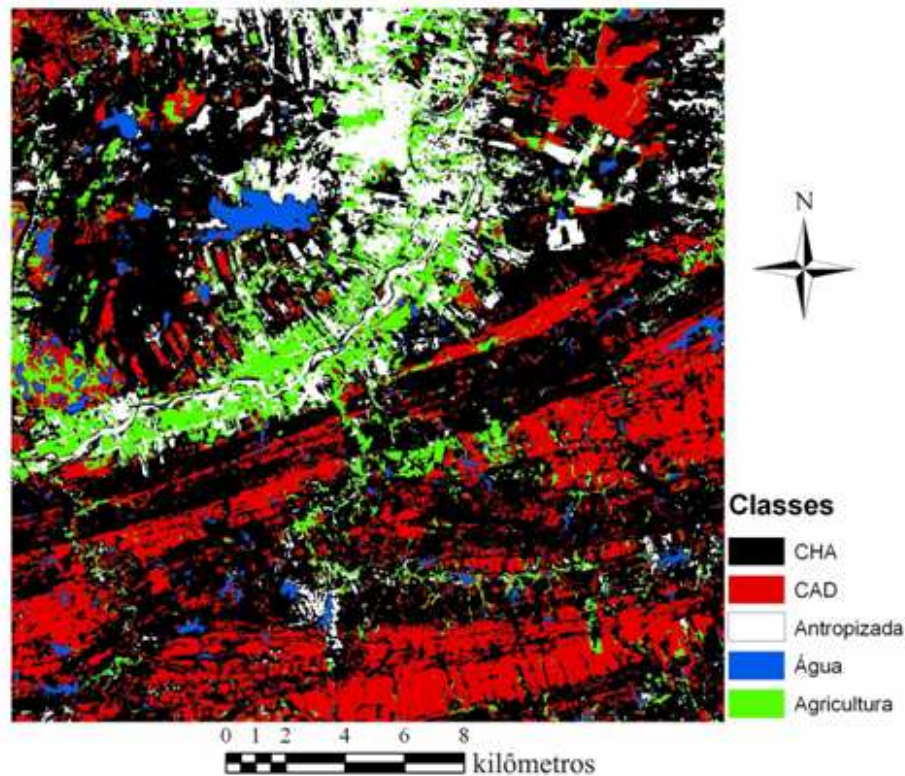


Figura 29 – Classificação pelo método da máxima verossimilhança

Os 112 pontos de GPS correspondentes à verdade terrestre foram comparados à classificação e a matriz de confusão foi montada. A Tabela 5 expõe a matriz com as classes: agricultura, água, antropizada, CHA e CAD. O total de linhas corresponde à soma dos pontos nas linhas e o total de colunas à soma dos pontos nas colunas. Os números da diagonal principal quantificam os pixels classificados corretamente. Por exemplo, para a classe antropizada 16 pixels foram classificados corretamente. Já os números dispostos fora da diagonal principal quantificam os pixels classificados incorretamente. Ainda para a classe antropizada, por exemplo, 3 pixels pertenciam a classe agricultura e outros 3 pixels à classe CHA.

Tabela 5 – Matriz de confusão para o método da máxima verossimilhança

		Verdade de Campo					Total Linhas
		Agricultura	Água	Antropizada	CHA	CAD	
Classificação	Agricultura	34		4		1	39
	Água		8		2		10
	Antropizada	3		16	3		22
	CHA	1		3	23		27
	CAD	2			2	10	14
Total colunas		40	8	23	30	11	112

A Exatidão Global (EG) foi obtida aplicando os dados obtidos na matriz de confusão convenientemente na Equação 56. Desta forma, para este método, a quantidade de pixels classificados corretamente para foi de 81,2%. Já o coeficiente de Kappa (K), obtido por meio da Equação 58, foi de 0,75. A classificação proposta por Landis e Koch (1977), exposta na Tabela 1, permite afirmar que a classificação resultante por meio do método da máxima verossimilhança é muito boa, já que o valor de Kappa encontrado está presente na faixa entre 0,6-0,8.

4.3 Classificação por meio do método Perceptron de Múltiplas Camadas (MLP)

Pode-se resolver a questão da definição do número de camadas escondidas por tentativa e erro ou pela experiência prévia do operador no domínio de determinado problema de classificação de padrões. Segundo Egmont et al. (2001), uma rede MLP com uma camada escondida é um mapeador universal, isto é, pode aproximar regiões de decisão contínua. Desta forma, optou-se por utilizar uma rede MLP de única camada e realizar testes apenas variando-se basicamente o número de neurônios na camada escondida (12,15 e18). No treinamento da rede, em nenhum dos testes o erro final desejado foi alcançado. Desta forma, a parada ocorreu devido a quantidade de iterações pré-estabelecidas (3000) ter chegado ao fim. A Tabela 6 mostra os parâmetros fornecidos à rede neural MLP em cada teste realizado, com os respectivos valores de erro final, EG, K e o tempo computacional.

Tabela 6 – Resultados dos testes pelo método MLP

	Teste	Número de neurônios	Erro final	EG (%)	K	Tempo (s)
MLP	1	12	0,008	82,14	0,76	657,5
	2	15	0,007	81,25	0,75	875,9
	3	18	0,008	70,53	0,61	1103,6

Observando os valores de EG e K na Tabela 6, nota-se que a rede realizou um bom mapeamento dos dados com 12 neurônios na camada escondida (teste 1). A performance, no entanto, caiu com o aumento do número de neurônios, podendo ter ocorrido um *overfitting*. Gelete e Volatão (2007), assim como no presente estudo, usaram a Toolbox Neural Network contida no software MATLAB (The MathWorks Inc., Natick, MA, USA) e obtiveram resultados satisfatórios com uma rede MLP com 11 neurônios na camada oculta. A Figura 30 expõe a classificação da área de estudo correspondente ao teste que apresentou melhor desempenho para este método (teste 1).

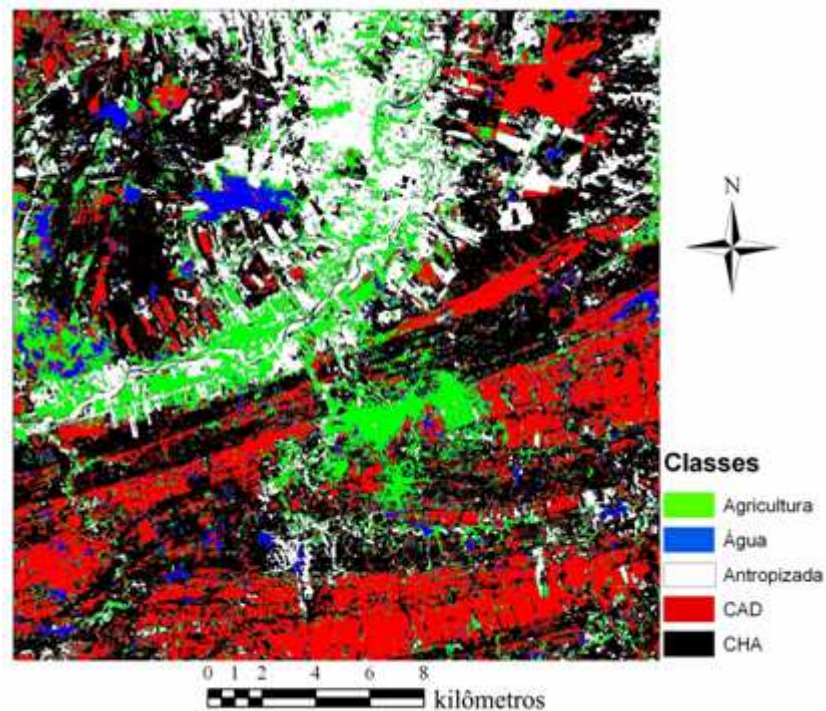


Figura 30 – Classificação por meio de uma rede MLP

Os valores de EG e K foram obtidos por meio das Equações 54 e 56 e os dados utilizados no cálculo destes coeficientes estão expostos na matriz de confusão ilustrada na Tabela 7.

Tabela 7 – Matriz de confusão para o método MLP

		Verdade de Campo					Total Linhas
		Agricultura	Água	Antropizada	CHA	CAD	
Classificação	Agricultura	33	1	4		1	39
	Água	1	7			2	10
	Antropizada	1		19	2		22
	CHA			6	21		27
	CAD				2	12	14
Total colunas		35	8	29	25	15	112

O valor de EG obtido para a rede MLP (82,14%) é próximo ao apresentado por Bischof et al. (1992), o que obteve um valor de EG de 85,9%. De acordo com os autores, este percentual pode ser melhorado se informações de textura forem apresentadas à rede.

O coeficiente de Kappa (K) calculado foi de 0,76. Assim como no método da máxima verossimilhança, o mapa temático gerado por este método apresenta-se, segundo Landis e Koch (1977), classificado como muito bom.

4.4 Classificação por meio do método Máquina de Vetor de Suporte (SVM)

O método SVM foi testado por meio da variação do fator de penalização (C) e da largura da função gaussiana (σ).

Primeiramente, objetivando analisar o comportamento de σ , foram escolhidos para C os valores de 100, 1000 e 1500 e, para cada um destes, foram realizados testes variando σ entre 1,5 e 10 (1,5; 1,8; 2,0; 2,2; 2,5; 10). A Tabela 8 expõe os resultados obtidos de EG, K e o tempo computacional de cada teste.

Tabela 8 – Desempenho do classificador SVM com a variação de σ

Teste	Fator C	σ	EG (%)	K	Tempo (s)
Teste 44	1000	1,8	82,14	0,766	350,98
Teste 6	1000	2	86,03	0,778	210,61
Teste 46	1000	2,2	81,03	0,752	266,95
Teste 36	1500	1,5	80,35	0,741	260,81
Teste 8	1500	2	82,14	0,766	207,12
Teste 38	1500	2,5	81,25	0,755	268,96
Teste 40	1500	5	80,35	0,742	276,60
Teste 42	1500	10	79,46	0,731	301,26
Teste 15	100	1,5	80,35	0,741	243,26
Teste 2	100	2	86,03	0,778	276,60
Teste 17	100	2,5	81,25	0,756	269,21
Teste 19	100	5	78,57	0,719	283,75
Teste 21	100	10	78,57	0,719	325,23

Pela análise da Tabela 8 nota-se que, para as amostras de treinamento utilizadas, os melhores valores de EG e K ocorrem com $\sigma=2$. Ainda é possível constatar que, para os três valores de C testados, e para valores de σ superiores a 2 os coeficientes de EG e K tendem a decrescer, já que com o aumento da largura da gaussiana pode ocorrer maior confusão na classificação dos dados. A Figura 31 exemplifica este comportamento para o teste em que C está fixado em 100.

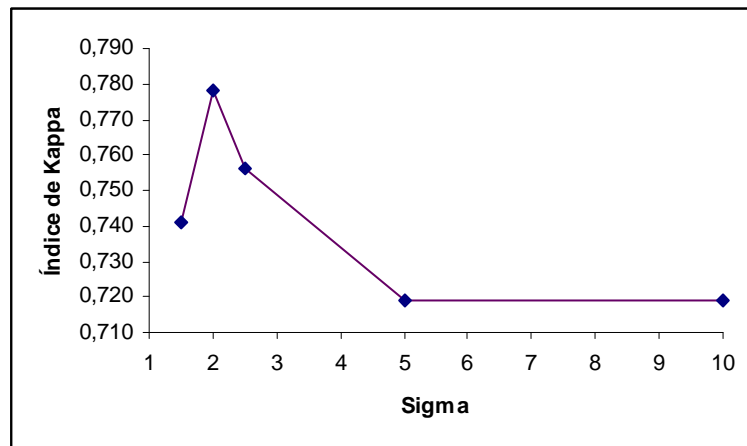


Figura 31 – Comportamento do coeficiente de Kappa com C mantido constante e variando o fator σ

Os demais testes foram realizados fixando σ no valor 2 e alterando o fator C entre 10 e 5000. A Tabela 9 expõe os resultados obtidos de EG, K e o tempo computacional de cada teste com a variação de C.

Tabela 9 – Desempenho do classificador SVM com a variação do fator C

Teste	Fator C	σ	EG (%)	K	Tempo (s)
Teste 13	10	2	81,25	0,755	215.25
Teste 11	50	2	82,14	0,767	202.68
Teste 2	100	2	86,03	0,778	276.60
Teste 4	500	2	86,03	0,778	206.90
Teste 6	1000	2	86,03	0,778	210.61
Teste 26	1300	2	83,03	0,778	337.34
Teste 28	1400	2	82,14	0,766	294.40
Teste 8	1500	2	82,14	0,766	207.12
Teste 30	2000	2	82,14	0,766	274.57
Teste 32	3000	2	80,35	0,741	303.20
Teste 34	4000	2	80,35	0,741	283.79
Teste 23	5000	2	80,35	0,741	327.68

Analisando a Tabela 9 nota-se que K apresentou tendência de crescimento com C entre 10 e 1300, porém, após este valor o coeficiente K passou a decrescer. Com valores de C maiores do que 1000 EG e K tendem a decrescer, já que a penalização é menos rígida e a classificação apresenta maior quantidade de erros. A Figura 32 expõe o comportamento de K com a variação do fator C (entre 10 e 5000) e com σ fixado em 2.

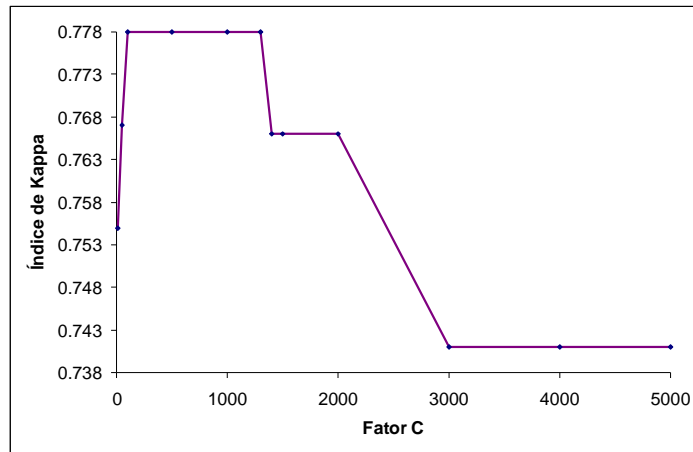


Figura 32 – Comportamento do coeficiente de Kappa com σ mantido constante e variando o fator C

Pela análise da Tabela 9 e da Figura 32 nota-se que as melhores classificações, ou seja, aquelas que apresentaram maiores valores de Exatidão Global (EG) e Kappa (K) para o método SVM, foram obtidas para valores de C entre 100 e 1000. Com valores de C maiores do que 1000 os valores dos coeficientes de EG e K tendem a decrescer e a classificação apresenta maior quantidade de erros.

A matriz de confusão correspondente ao teste 6 está exposta na Tabela 10. O total de linhas corresponde à soma dos pontos nas linhas e o total de colunas à soma dos pontos nas colunas.

Tabela 10 – Matriz de confusão para o método SVM

		Verdade de Campo					Total Linhas
		Agricultura	Água	Antropizada	CHA	CAD	
Classificação	Água	30		7		2	39
	Antropizada		10				10
	Agricultura			21	1		22
	CHA	3		4	20		27
	CAD	1			1	12	14
Total colunas		34	10	32	22	14	112

O mapa temático resultante com as classes CHA, CAD, água, antropizada e agricultura está exposto na Figura 33.

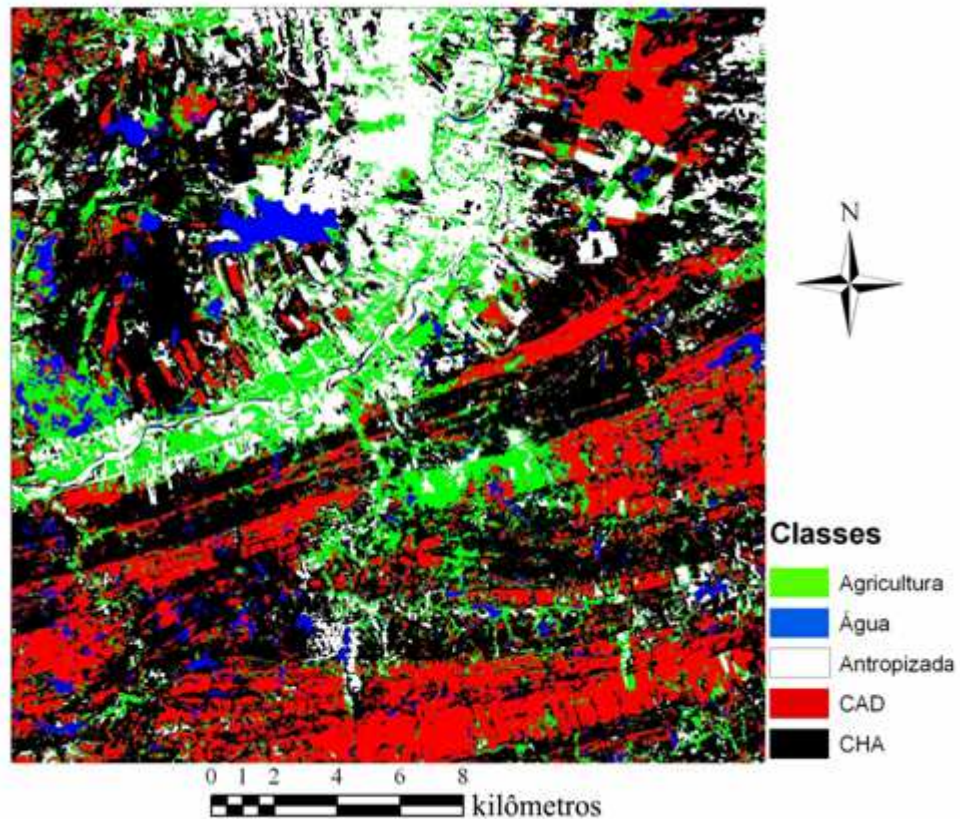


Figura 33 – Classificação utilizando o método SVM

A classificação realizada pelo método SVM, correspondente ao teste 6, obteve coeficiente de Kappa igual a 0,77. Segundo Landis e Koch (1977), o mapa temático resultante é rotulado como muito bom. Assim como nos estudos de Foody e Mathur (2004) e Pal e Mather (2004), a SVM obteve desempenho satisfatório quando comparado aos métodos MLP e da máxima verossimilhança, os quais apresentaram K de 0,76 e 0,73, respectivamente.

Um outro ponto importante a ser analisado é o custo computacional envolvido no processo de classificação. Para o método SVM, como pode ser observado na Tabela 9, o teste 6 realizou o processo de classificação em 210,51 segundos. Já para no método MLP, o teste 1 realizou o mesmo processo em 657,5 segundos. Desta forma, pode-se afirmar que o SVM apresentou menor custo computacional do que o método MLP ratificando os resultados encontrados por Foody e Mathur (2004), os quais observaram que o método SVM foi mais rápido quando comparado à rede MLP e a árvore de decisão.

4.5 Avaliação dos coeficientes de exatidão específica (EE)

Para avaliar o desempenho dos classificadores em cada classe, o coeficiente de EE foi calculado através da Equação 56. A Tabela 11 expõe os valores destes coeficientes para os métodos da máxima verossimilhança, MLP e SVM.

Tabela 11 – Exatidão específica das classes para cada método

Classe	Exatidão específica (%)		
	Máxima Verossimilhança	MLP	SVM
Agricultura	87,2	84,6	76,9
Água	80,0	70,0	100,0
Antropizada	72,7	86,4	95,5
CHA	85,2	77,8	74,1
CAD	71,4	85,7	85,7

A classe agricultura apresentou nível de acurácia de 87,2% quando classificada pelo método da máxima verossimilhança, sendo este o maior nível obtido para esta classe. Da mesma forma, os maiores níveis de acurácia encontrados para as classes água e antropizada foram de 100 e 95,5%, respectivamente, quando classificadas pelo método SVM sendo, portanto, este o método mais indicado para a classificação destas classes. O valor de EE para a classe Caatinga Herbácea Arbustiva (CHA) foi de 85,2, 77,8 e 74,1% para o método da máxima verossimilhança, MLP e SVM, respectivamente. Já para a classe Caatinga Arbórea Densa (CAD) o valor de EE foi de 85,7, para os métodos SVM e MLP e 74,1% para o método da máxima verossimilhança. Os resultados expostos apontam para o bom desempenho dos métodos baseados em aprendizado de máquina. Assim como no trabalho de Su et al. (2007), os resultados obtidos permitem afirmar que estes métodos estão aptos à classificação do bioma em Caatinga, já que apresentaram resultados equivalentes e até superiores ao método da máxima verossimilhança.

4.6 Quantificação das áreas

As áreas das classes agricultura, água, antropizada, CHA e CAD foram quantificadas para os métodos SVM, MLP e máxima verossimilhança. Os resultados, em km², estão expostos na Figura 34.

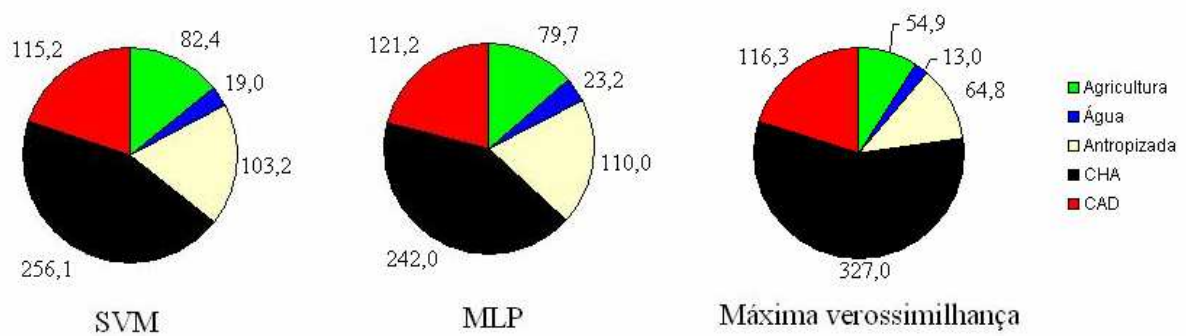


Figura 34 – Área total por cada classe (km²)

As classificações realizadas permitiram concluir que, na área de estudo, existe uma parcela significativa desta ocupada pelo bioma Caatinga, variando de 363 a 443 km² dependendo do método. A quantificação ainda permite afirmar que, entre os dois tipos de Caatinga, a predominante é a do tipo herbácea arbustiva (CHA) com área variando entre 242 e 327 km². Já a do tipo arbórea densa (CAD) apresenta área variando entre 115 e 121 km², dependendo do método. A área total de estudo é 576 km². Sendo o método SVM o que apresentou melhor desempenho dentre os testados, pode-se utilizar os valores de área resultantes deste como referência aos demais. Pela Figura 34, nota-se que método MLP apresentou valores de resultados de área semelhantes ao SVM. Já o da máxima verossimilhança superestimou o valor de área da classe CHA e subestimou o da classe antropizada quando comparados ao SVM e MLP.

5. Conclusões

As amostras de treinamento apresentaram boa separabilidade, o que contribuiu para o desempenho satisfatório dos classificadores utilizados.

Os classificadores baseados em aprendizado de máquina obtiveram desempenho superior ao classificador tradicional da máxima verossimilhança.

Os valores dos coeficientes de exatidão global (EG), de exatidão específica (EE) e de Kappa (K), obtidos a partir da verdade terrestre, revelam o alto potencial dos métodos baseados em aprendizado de máquinas na classificação do uso e ocupação do solo no bioma Caatinga.

O método SVM apresentou maior índice de Kappa. A escolha adequada do parâmetro da função kernel (σ) e do parâmetro de penalização (C) proporciona aumento na acurácia da classificação por este método.

Além do desempenho satisfatório na classificação do bioma Caatinga, o método SVM apresentou excelente desempenho para as classes água e antropizada pela análise dos coeficientes de exatidão específica (EE). Podendo, desta forma, ser usado para esta função em estudos onde se deseja realizar monitoramento dos recursos hídricos e da degradação ambiental em áreas com características semi-áridas.

Os métodos baseados em aprendizado de máquina apresentaram valores de área semelhantes para as cinco classes consideradas. Foi ainda possível constatar que, dentre os dois tipos de Caatinga considerados, o de herbácea arbustiva é predominante na área de estudo.

Apesar dos melhores resultados de EG e K obtidos em comparação ao método da máxima verossimilhança, classificadores baseados em aprendizado de máquina podem consumir maior tempo computacional além da necessidade do usuário despende boa parte do tempo na escolha dos parâmetros do método a ser utilizado, que pode ser extremamente desgastante. Assim, caso o tempo seja um fator limitante, pode-se utilizar o método da máxima verossimilhança.

6. Perspectivas

O trabalho apresenta parte do desenvolvimento de uma metodologia de pesquisa em busca por melhores dados de validação, principalmente no que concerne aos novos paradigmas de classificação. No entanto, o trabalho também objetiva consolidar referências de medições, às quais permitam montar uma estação piloto de aferição de sensores de solo, orbitais e aerotransportados, utilizando-se sensores certificados e georreferenciados, numa forma de auxiliar também a calibração da resposta espectral dos atuais e futuros satélites, principalmente, quando aplicados ao sensoriamento remoto do semi-árido.

Espera-se também que estas ferramentas permitam automação do monitoramento e quantificação da presença deste bioma e, por consequência, maior agilidade na produção de informações que subsidiem a adoção de práticas conservacionistas.

A proposta deverá evoluir para o desenvolvimento de sistemas automáticos de classificação e validação, através da coleta de dados de sensores certificados, de forma que, a partir de referências de dados cada vez mais confiáveis, seja possível indicar a potencialidade de novas ferramentas computacionais no processo de classificação de imagens de satélite que contemplem às condições semi-áridas. Pretende-se também desenvolver novos classificadores específicos e adequados à realização da classificação de diferentes fitofisionomias de Caatinga.

REFERÊNCIAS

- ASSAD, E. D.; SANO, E. E. **Sistema de informações geográficas: aplicações na agricultura**. 2. ed. Brasília: EMBRAPA-SPI/EMBRAPA-CPAC, 1998. 434p.
- BELOUSOV, A. I.; VERZAKOV, S. A.; FRESE, J. V. A flexible classification approach with optimal generalization performance: support vector machines. **Chemometrics and Intelligent Laboratory Systems**, v. 64, n. 1, p. 15-25, out. 2002.
- BISCHOF, H.; SCHNEIDER, W.; PINZ, A. J. Multispectral classification of LANDSAT-images using neural networks. **IEEE Transactions Geoscience and Remote Sensing**, v. 30, n. 3, p. 482-490, maio. 1992.
Disponível em:
< <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=00142926>>. Acesso em: 13 jul. 2008.
- BOSER, B.; GUYON, I; VAPNIK, V. N. A training algorithm for optimal margin classifiers. Annual Workshop on Computational Learning Theory, 5., 1992, Pittsburgh. **Proceedings...** SAnMateo, CA: Morgan Kayfmann, 2007. p. 144-152. Disponível em: <<http://portal.acm.org/citation.cfm?id=130401>>. Acesso em: 15 dez. 2008.
- BUCENE, L. C.; RODRIGUES, L. H. A. Utilização de redes neurais artificiais para avaliação de produtividade do solo, visando classificação de terras para irrigação. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v.8, n.2/3, p. 326-329, maio/dez. 2004. Disponível em:< <http://www.agriambi.com.br/>>. Acesso em: 8 de set. 2008.
- CANU, S.; GRANDVALET, Y.; GUIGUE, V.; RAKOTOMAMONJY, A. **SVM and Kernel Methods MATLAB Toolbox**. Perception Systèmes et Information, INSA de Rouen, Rouen, France, 2005.
- CANU, S.; MARY, X.; RAKOTOMAMONJY, A. Functional learning through kernels. **Advances in learning theory: methods, models and applications**. IOS Press: 2003. v. 190, cap. 5, p. 89-110.
- CARVALHO, L. M. T; CLEVERES, J. G. P. W. ; SKIDMORE, A. K.; JONG, S. M. J. Selection of imagery data and classifiers for mapping Brazilian semideciduous Atlantic forests. **International Journal of Applied Earth Observation and Geoinformation**, v. 5, n. 5, p. 173-186, set. 2004.
- CHAGAS, C. S.; FILHO, E. I. F.; VIEIRA, C. A. O.; JÚNIOR, W. de C. Utilização de redes neurais artificiais para predição de classes de solo em uma bacia hidrográfica no Domínio de Mar de Morros. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 13., 2007, Florianópolis. **Anais...** Florianópolis: INPE, 2007. p. 2421-2428. CD-ROM, On-line. ISBN 978-85-17-00031-7. Disponível em:
<<http://urlib.net/dpi.inpe.br/sbsr@80/2006/10.30.21.31>>. Acesso em: 15 set. 2008.
- CHAPELLE, O.; HAFFNER, P.; VAPNIK, V. N. Support Vector Machines for Histogram-Based Image Classification. **IEEE Transactions On Neural Networks**, v. 10, n. 5, p. 1055-1064, set. 1999. Disponível em:< <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=00788646>>. Acesso em 16 de dez. 2008.
- COGERH. 2005. **Anuário do monitoramento quantitativo dos principais açudes do Estado do Ceará**. Fortaleza: Companhia de Gestão dos Recursos Hídricos (COGERH). 260 p.
- CORTES, C.; VAPNIK, V. Support vector networks. **Machine Learning**. v. 20, n. 3, p. 273-297, set. 1995.

COVER, T. M. Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition. **IEEE Transactions on Electronic Computers**, v. 25, n. 3, p. 326-334, jun. 1965. Disponível em: <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4038448>. Acesso em: 10 jan. 2009.

CRISTIANINI, N.; TAYLOR, J. S. **Introduction to Support Vector Machines**. UK: Cambridge University Press, 2000.

EGMONT, M. P.; RIDDER, D. de.; HANDELS, H. Image processing with neural networks—a review. **Pattern Recognition**, v. 35, n. 10, p. 2279–2301, out.2002.

ERBERT, M. **O Uso da Análise Discriminante Regularizada (RDA) no Reconhecimento de Padrões em Imagens Digitais Hiperespectral de Sensoriamento Remoto**. 2001. 118 f. Dissertação (Mestrado em Sensoriamento Remoto) - Universidade Federal do Rio Grande do Sul, Porto Alegre, 2001. Disponível em: <<http://www.inf.ufrgs.br/~silvia/ipg/IntroSensRemoto.pdf>>. Acesso em: 15 ago. 2008.

FIGUEIREDO, G. C.; VIEIRA, C. A. O. Estudo do comportamento dos índices de Exatidão Global, Kappa e Tau, comumente usados para avaliar a classificação de imagens do sensoriamento remoto. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 13., 2007a, Florianópolis. **Anais...** Florianópolis: INPE, 2007. p. 5755-5762 CD-ROM, On-line. ISBN 978-85-17-00031-7. Disponível em: <<http://urlib.net/dpi.inpe.br/sbsr@80/2006/11.13.17.35>>. Acesso em: 04 jan. 2009.

FIGUEIREDO, S. M. M.; CARVALHO, L. M. T. Análise comparativa entre técnicas de classificação digital em imagens LANDSAT no mapeamento do uso e ocupação do solo em Capixaba, Acre. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 13., 2007, Florianópolis. **Anais...** Florianópolis: INPE, 2007. p. 6729-6736. CD-ROM, On-line. ISBN 978-85-17-00031-7. Disponível em: <<http://urlib.net/dpi.inpe.br/sbsr@80/2006/11.13.17.02>>. Acesso em: 15 jan. 2009.

FOODY, G. M.; MATHUR, A. A Relative Evaluation of Multiclass Image Classification by Support Vector Machines. **IEEE Transactions on Geoscience And Remote Sensing**, v. 42, n.6, p. 1335-1343, jun.2004. Disponível em: <<http://ieeexplore.ieee.org/search/wrapper.jsp?arnumber=1304900>>. Acesso em: 10 dez. 2008.

GELETE, C.; VOLATÃO, C. F. de S. Análise do potencial de classificação do uso e cobertura do solo por meio de rede neural. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 13., 2007, Florianópolis. **Anais...** Florianópolis: INPE, 2007. p. 5779-5785. CD-ROM, On-line. ISBN 978-85-17-00031-7. Disponível em: <<http://urlib.net/dpi.inpe.br/sbsr@80/2006/10.20.13.05>>. Acesso em: 13 out. 2008.

GIDUDU, A.; HULLEY G.; MARWALA, T. Image Classification Using SVMs: One-against-One Vs One-against-All. In: ASIAN CONFERENCE ON REMOTE SENSING, 28., 2007, Kuala Lumpur. **Proceedings...** Singapore: ACSR, 2007. Disponível em: <<http://arxiv.org/abs/0711.2914v1>>. Acesso em: 05 nov. 2008.

GIGANDET, X.; CUADRA, M. B.; POINTET, A.; CAMMOUN, L.; CALOZ, R.; THIRAN, J. Region-based satellite image classification: method and validation. **Image Processing**, v. 3, n. 4, p. 832-835, Set.2005. Disponível em: <<http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=01530521>>. Acesso em: 14 out. 2008.

GLEZAKOS, T.; TSILIGIRIDIS, T. *Neural Networks for Landscape Applications*. In: CONFERENCE OF HAICTA (HELLENIC ASSOCIATION OF INFORMATION AND COMMUNICATION TECHNOLOGY IN AGRICULTURE, FOOD AND ENVIRONMENT), 1., 2002, Athens. **Proceedings...** Athens: HAICTA, 2002. p. 220-233. Disponível em: <<http://www.aua.gr/tmhmeta/geniko/tsili/geostat.html>>. Acesso em: 12 set. 2008.

GONÇALVES, P.; CARRÃO, H.; PINHEIRO, A.; CAETANO, M. Land cover classification with Support Vector Machine Applied to MODIS imagery. In: GLOBAL DEVELOPMENTS IN ENVIRONMENTAL EARTH OBSERVATION FROM SPACE, 25., 2006, Porto. **Proceedings...**Porto: Millpress Science Publishers, 2006. p. 517-526. Disponível em: < <http://www.igeo.pt/gdr/pdf/Goncalves2006.pdf>>. Acesso em: 25 set. 2008.

HAYKIN, S. **Neural Networks: A comprehensive foundation**. 2. ed. New Jersey: Prentice Hall, 1999. 842 p.

HSU, C. W.; LIN, C. J. A Comparison of Methods for Multiclass Support Vector Machines. **IEEE Transactions On Neural Networks**, v. 13, n. 2, p. 415-425, mar. 2002. Disponível em: <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=00991427>. Acesso em: 2 dez. 2008.

IWAI, O. K. **Mapeamento do uso do solo urbano no município de São Bernardo do Campo, através de imagens de satélites**. 2003. 127 f. Dissertação (Mestrado em Engenharia de Transportes) – Escola Politécnica da Universidade de São Paulo. São Paulo, 2003. Disponível em: <<http://www.teses.usp.br/teses/disponiveis/3/3138/tde-03022004-152010/>>. Acesso em: 15 dez. 2007.

JACOMINE, P. K. T.; ALMEIDA, J. C.; MEDEIROS, L. A. R. **Levantamento exploratório-reconhecimento de solos do estado do Ceará**. Recife, Ministério da Agricultura/ Superintendência do Desenvolvimento do Nordeste, 1973, v.1. 301p. (Boletim técnico, 28; Série Pedologia, 16).

JI, C. Y. Land-Use Classification of Remotely Sensed Data Using Kohonen Self-organizing Feature Map Neural Networks. **Photogrammetric Engineering & Remote Sensing**, v. 66, n. 12, p. 1451-1460, dez. 2000.

JIANWEN, M.; BAGAN, H. Land-use classification using ASTER data and self-organized neural networks. **International Journal of Applied Earth Observation and Geoinformation**, v.3, n.3, p. 183–188, nov. 2005.

KÖPPEN, W.1948. **Climatologia: con un estudio de los climas de la tierra**. México: Fondo de Cultura Económica. 479 p.

KRAISCH, H.; REFOSCO, J.C; VIBRANS, A. C. Estudo de comportamento espectral de alvos e análise multitemporal da região da foz do rio Itajaí – Santa Catarina. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 2007, 13., Florianópolis. **Anais...** Florianópolis: INPE, 2007. p. 5329-5336. CD-ROM, On-line. ISBN 978-85-17-00031-7. Disponível em: <<http://urlib.net/dpi.inpe.br/sbsr@80/2006/11.15.20.25>>. Acesso em: 15 jan. 2009.

LANDIS, J. R.; KOCH, G. C. The measurement of observer agreement for categorical data. **Biometrics**, v. 33, n. 1, p. 159-174, mar. 1977.

LEAL, I. R. ; SILVA, J. M. C.; TABARELLI, M.; JÚNIOR, T. E. L. Mudando o curso da conservação da biodiversidade na Caatinga do Nordeste do Brasil. **Megadiversidade**, Belo Horizonte, v. 1, n. 1, p. 139-146, jul. 2005.

LEDUC, R.; GRAÇA, P.M.L. de A.; PEIXOTO, J.M.A. Classificação da cobertura do solo da Reserva Extrativista do Baixo – Juruá/AM. . In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 2007, 13., Florianópolis. **Anais...** Florianópolis: INPE, 2007. p. 5329-5336. CD-ROM, On-line. ISBN 978-85-17-00031-7. Disponível em: <<http://martedpi.inpe.br/col/dpi.inpe.br/sbsr@80/2006/11.16.02.53.01/doc/6787-6789.pdf>>. Acesso em: 13 mar. 2009.

LIN, W. C. A hierarchical multiple view approach to three-dimensional object recognition. **IEEE Transactions On Neural Networks**, v. 2, n. 1, p. 84-92, jan. 1991.

MAIA, G. N. **Caatinga: Árvores e arbustos e suas utilidades**. São Paulo: Ed. Computação Gráfica e Editora, 2004. 413p.

MANGABEIRA, J. A. de C., AZEVEDO, E. C. de, LAMPARELLI, R. A. C. **Avaliação do levantamento do uso das terras por imagens de satélite de alta e média resolução espacial**. Campinas: Embrapa Monitoramento por Satélite, 2003. 16 p. (Comunicado Técnico, 11). Disponível em: <<http://www.cnpm.embrapa.br/publica/serie.html>>. Acesso em: 15 jan. 2008.

MEDEIROS, J. S. **Bancos de Dados Geográficos e Redes Neurais Artificiais: Tecnologias de Apoio à Gestão de Território**. 1999. 207f. Tese (Doutorado em geografia) – Universidade de São Paulo, São Paulo, 1999. Disponível em: <<http://www.dpi.inpe.br/teses/simeao/>>. Acesso em: 25 jul. 2008.

MELGANI, F. Classification of Hyperspectral Remote Sensing Images With Support Vector Machines. **IEEE Transactions on geoscience and remote sensing**, v. 42, n. 8, p. 1778-1790, ago.2004. Disponível em: <<http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=01323134>>. Acesso em: 7 out. 2008.

MERCIER, G; LENNON, M. Support vector machines for hyperspectral image classification with spectral based kernels. In: IEEE INTERNATIONAL GEOSCIENCE & REMOTE SENSING SYMPOSIUM, 23., 2003, Toulouse, France. **Proceedings...** Toulouse: GRSS, 2003. p.288-290. Disponível em: <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1293752>. Acesso em: 13 set. 2008

MINSKY, M. L., PAPERT, S. A. **Perceptrons**. Cambridge, MA: MIT Press, 1969.

MITCHELL, T. **Machine Learning**. New York: McGraw Hill. 1997. 414p.

MMA – **Ministério do Meio Ambiente, dos Recursos Hídricos e da Amazônia Legal**. 2002. Avaliação e ações prioritárias para a conservação da biodiversidade da Caatinga. Universidade Federal de Pernambuco, Conservation International do Brasil e Fundação Biodiversitas, Brasília. Disponível em: <http://www.cepan.org.br/docs/publicacoes/livro_caatinga/25_caatinga_cap20_areas_prioritarias.pdf>. Acesso em: 12 ago. 2008.

MOREIRA, M. A. **Fundamentos do sensoriamento remoto e metodologias de aplicação**. 3. ed. Viçosa: UFV, 2005. 320p.

NISHIDA, W. **Uma rede neural artificial para classificação de imagens multiespectrais de sensoriamento remoto**. 1998. Dissertação (Mestrado em Engenharia de Produção) - Universidade Federal de Santa Catarina, Florianópolis, 1998. Disponível em: <http://biblioteca.universia.net/html_bura/ficha/params/id/600620.html>. Acesso em: 11 jan. 2008.

NOVO, E. M. L. M. **Sensoriamento Remoto, Princípios e Aplicações**; 2. ed. São Paulo: Edgrad Blucher Ltda., 1989. 269 p.

OLIVEIRA, S. B. P.; SOUZA, M. J. N.; LEITE, F. R. B.; COSTA, R. N. S. Contribuição ao estudo da degradação ambiental no município de Caridade-CE. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 11., 2003, Belo Horizonte. **Anais...** Belo Horizonte: INPE, 2003. p. 1391 - 1398. CD-ROM, Online. ISBN 85-17-00017-X. Disponível em: <<http://urlib.net/ltid.inpe.br/sbsr/2002/11.11.09.39>>. Acesso em: 24 nov. 2009.

ORUC, M.; MARANGOZ, A. M.; BÜYÜKSALIH, G.; Comparison of Pixel-based and Object-oriented Classification Approaches Using LANDSAT-7 ETM Spectral Bands. In: PROCEEDINGS OF THE ISRPS 2004 ANNUAL CONFERENCE, 2004, Istanbul, Turkey. **Proceedings...**Istanbul: ISPRS, 2004. p. 19 - 23. CD-ROM, Online. Disponível em: <http://www.isprs.org/new/congresses/istanbul2004/comm4/papers/510.pdf>. Acesso em: 12 dez. 2008.

OSUNA, E.; FREUND, R.; GIROS, F. **Support vector machines: Training and applications**. Massachusetts: M.I.T Artificial Intelligence Laboratory, 1997. 42 p.

PAL, M.; MATHER, P. M. Assessment of the effectiveness of support vector machines for hyperspectral data. **Future Generation Computer Systems**, v. 20, n. 7, p. 1215–1225, out. 2004.

QUEIROZ, R. B.; SEVERINO, P. A. R.; RODRIGUES, A. G.; GÓMEZ, A. T. U. Redes Neurais: Um comparativo com Máxima Verossimilhança Gaussiana na Classificação de Imagens CBERS. In: WORKSHOP DE TECNOLOGIA DA INFORMAÇÃO APLICADA AO MEIO AMBIENTE, 2., 2004, Itajaí. **Anais...** Universidade Vale do Itajaí: UVI, 2004. p. 746-749. CD-ROM, On-line. ISBN 1677-2822. Disponível em: <http://www.niee.ufrgs.br/cbcomp/cbcomp2004/html/pdf/Workshop_Ambiente/Intelig%eancia_Artificial/t170100151_3.pdf>. Acesso em: 05 nov. 2008.

RAKOTOMAMONJY, A.; CANU, S. Frames, Reproducing Kernels, Regularization and Learning. **Journal of Machine Learning Research**, v. 6, p. 1485–1515, set. 2005. Disponível em: <<http://portal.acm.org/citation.cfm?id=1194895>>. Acesso em: 19 jan. 2009.

RIBEIRO, R. M. P.; SOARES, V. P.; VIEIRA, C. A. O. Avaliação de métodos de classificação de imagens IKONOS para o mapeamento da cobertura terrestre. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 2005, 12., Goiânia. **Anais...** Goiânia: INPE, 2005. p. 4277-4283. CD-ROM, On-line. ISBN 85-17-00018-8. Disponível em: <<http://urlib.net/ltid.inpe.br/sbsr/2004/11.21.17.32>>. Acesso em: 24 ago. 2008.

RIBEIRO, S. R. A.; CENTENO, J. S. Classificação do Uso do Solo Utilizando Redes Neurais e o Algoritmo MAXVER. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 2001, 10., Foz do Iguaçu. **Anais...** Foz do Iguaçu: INPE, 2001. p. 1341-1348. CD-ROM, On-line. ISBN 85-17-00016-1. Disponível em: <<http://urlib.net/dpi.inpe.br/lise/2001/09.20.17.56>>. Acesso em: 07 abr. 2008.

RICHARDS, J.A. **Remote sensing digital image analysis: An introduction**. 4 ed. New York: Springer, 1993. 439 p.

ROSENBLATT, F. The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain. **Psychological Review**, v. 65, n. 6, p. 386-408, nov. 1958.

RUMELHART, D. E.; McCLELLAND, J. L. **Parallel Distributed Processing: Explorations in the Mricrostructure of Cognition**, v. 1, Cambridge: MIT Press, 1986.

RUMELHART, D.E.; HINTON, G.E.; WILLIAMS, R. J. Learning Representations of Back-Propagation Errors. **Nature (London)**, v. 323, n.9, p. 533-536, out. 1986.

SCHOLKOPF, B.; SMOLA, A. **Learning with Kernels**. Cambridge: Mit Press, 2001.

SEMOLINI, R. **Support vector machines, inferência transdutiva e o problema de classificação**. 2003. 128 f. Tese (Mestrado em Engenharia Elétrica) – Universidade Estadual de Campinas, Campinas, São Paulo, 2003. Disponível em: <<http://libdigi.unicamp.br/document/?code=vtls000286513>>. Acesso em: 4 jan. 2009.

SILVA, M. L. M.; PEREIRA, G. P. M. Avaliação de algoritmos de classificação supervisionada para imagens do CBERS-2 da Região do Parque Estadual do Rio Doce-MG. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 2007, 13., Florianópolis. **Anais...** Florianópolis: INPE, 2007. p. 6223-6228. CD-ROM, On-line. ISBN 978-85-17-00031-7. Disponível em: <<http://urlib.net/dpi.inpe.br/sbsr@80/2006/11.15.20.25>>. Acesso em: 15 jan. 2009.

SMOLA, A. J.; BARLETT, P.; SCHÖLKOPF, B.; SCHUURMANS, D. **Advances in Large Margin Classifiers**. Cambridge: MIT Press, 2000. 423 p.

SOUSA, B. F. S.; TEIXEIRA, A. dos S.; LEÃO, R. A. de O.; FILHO, A. B. C. Uso do solo da bacia hidrográfica do Alto Piauí através de imagens do satélite CBERS. **Revista Ciência Agronômica**, Fortaleza, v. 38, n. 4, p. 327-334, out. 2007.

SU, L.; MARK, J. C.; RANGO, A.; MARTONCHIK, J. V.; PETERS, D. P. C. Support vector machines for recognition of semi-arid vegetation types using MISR multi-angle imagery. **Remote Sensing of Environment**, v. 107, n.1 p. 299-311, jan. 2007.

SULSOFT (2009). **Guia do Envi em português**. Disponível em: <www.sulsoft.com.br>. Acesso em: 27 jan. 2009.

Vapnik, V.N. and Chervonenkis, A. (1971), On the uniforme convergence of relative frequencies of events to their probabilities. *Theoretical Probability and Its Applications*, vol. 17, pp. 264-280.

VAPNIK, V.N. **Statistical Learning Theory**, New York: Wiley. 1998.

VAPNIK, V.N. **The Nature of Statistical Learning Theory**. New York: Springer-Verlag. 1995. 332 p.

VAPNIK, V.N., CHERVONENKIS, A.Y. **Theory of pattern recognition**. Moscow: Nauka. 1974.

VONK, E.; JAIN, L. C.; JOHNSON, R. P. **Automatic Generation of Neural Network Architecture Using Evolutionary Computation**. New Jersey: World Scientific Publishing Co. Pte. Ltda., 1997. 200 p.

WANDSCHEER, L. Caatinga está sendo destruída mais rápido do que a Amazônia. **Agência Brasil**, Brasília, 26 out 2008. Disponível em: <<http://www.agenciabrasil.gov.br/noticias/2008/10/29/materia.2008-10-29.7369241684/view>>. Acesso em: 29 out.2008.

WANG, Q., YANG, J. Eye Detection in Facial Images with Unconstrained Background. **Journal of Pattern Recognition Research**, China, v. 1, n. 5, p. 285-289, dec. 2006.

WASSERMAN, P.D. **Neural computing: theory and practice**. New York: Van Nostrand Reinhold. 1989. 230p.

WATZLAWICK, L. F.; KIRCHNER, F. F.; MADRUGA, P. R de A. Avaliação de classificação digital em imagens de vídeo multiespectral utilizando matriz confusão. **Revista de Ciências Exatas e Naturais**, Guarapuava-PR, v. 5, n. 1, p. 47-57, jul./dez. 2003.

WESTON, J.; WATKINS, C. Multi-Class Support Vector Machines. Maio, 1998. Comunicado técnico.

APÊNDICE 1

```

% -----
% Método MLP
% Universidade Federal do Ceará
% Departamento de Engenharia Agrícola
% Curso de mestrado em Engenharia Agrícola
% Beatriz F. S. Sousa
% Orientador: Prof. Adunias dos Santos Teixeira (UFC)
% Coorientador: Prof. Francisco de Assis Tavares (INPE)
% -----

PASTA='E:\_Beatriz \';

% Imagem à ser classificada
IGUATU=imread(cat(2,PASTA,'imagem_teste_1_2008.tif'));

% Amostras de treinamento
agricultura=load(cat(2,PASTA,'agricultura_RGB.txt'));
agua=load(cat(2,PASTA,'agua_RGB.txt'));
antropizada=load(cat(2,PASTA,'antropizada_RGB.txt'));
caatinga=load(cat(2,PASTA,'CHA_RGB.txt')); % Classe Caatinga Herbácea Arbustiva
vegetacaonativa=load(cat(2,PASTA,'CAD_RGB.txt')); % Classe Caatinga Arbórea Densa

agricultura=agricultura(1:2000,:);
agua=agua(1:2000,:);
antropizada=antropizada(1:2000,:);
caatinga=caatinga(1:2000,:);
vegetacaonativa=vegetacaonativa(1:2000,:);

metros_agricultura=0;
metros_agua=0;
metros_antropizada=0;
metros_caatinga=0;
metros_vegetacaonativa=0;

% Armazenando de dimensões da imagem e número de pixels
W=size(IGUATU);
W_pixel=W(1)*W(2);

% Armazenando quantidade de exemplos da classe
SIZE_AGRICULTURA=size(agricultura);
SIZE_AGUA=size(agua);
SIZE_ANTROPIZADA=size(antropizada);
SIZE_CAATINGA=size(caatinga);
SIZE_VEGETACAONATIVA=size(vegetacaonativa);

```

```

%Transformando a imagem para vetor
for I=1:W(1)
    for J=1:W(2)
        IGUATU_VETOR((I-1)*W(2)+J,:)= IGUATU(I,J,:);
    end;
end;

%-----
% Definição das variáveis de entrada
%-----

tic
%Matriz de treino
treino_mlp = [agricultura;agua;antropizada;caatinga;vegetacaonativa];
treino_mlp = treino_mlp';

%matriz de saída
saida_treino = [ones(1,SIZE_AGRICULTURA(1)) 2.*ones(1,SIZE_AGUA(1))
3.*ones(1,SIZE_ANTROPIZADA(1)) 4.*ones(1,SIZE_CAATINGA(1))
5.*ones(1,SIZE_VEGETACAONATIVA(1))];

%-----
% Criando a rede
%-----

net = newff ([min(treino_mlp')max(treino_mlp')], [15,1], {'tansig','tansig'}, 'trainlm');

net.trainParam.show = 50;
net.trainParam.epochs = 3000;
net.trainParam.goal = 0.001;
net.trainParam.lr = 0.05;

%-----
% Treinando a rede (AMOSTRAS DE TREINAMENTO)
%-----

[treino_mlp, min_treino_mlp, max_treino_mlp, saida_treino, min_saida_treino,
max_saida_treino] = premmx(treino_mlp, saida_treino);

[net,erro] = train(net, treino_mlp, saida_treino);
toc
Treinando_MLP=toc

```

```

%-----
% Aplicando a rede treinada na imagem
%-----
tic
teste_mlp = double(IGUATU_VETOR);
teste_mlp = teste_mlp';
teste_mlp = trammx(teste_mlp, min_treino_mlp, max_treino_mlp);
teste_mlp = teste_mlp';

    for i=1:W_pixel
        aux=sim(net,teste_mlp(i,:));
        teste_mlp_out(:,i)=aux(:,1);
    end;

toc
testando_MLP=toc

%-----
% Pós-processamento os dados de saída da rede
%-----

resposta2 = postmnmx(teste_mlp_out,min_saida_treino,max_saida_treino);
resposta_final=uint8(resposta2);

% -----
% Tratamento dos resultados
% -----

for I=1:W(1)
    for J=1:W(2)
        switch resposta_final((I-1)*W(2)+J)
            case 1
                IGUATU_FINAL(I,J,:)= [0 255 0];                % Agricultura
                metros_agricultura=metros_agricultura+1;
            case 2
                IGUATU_FINAL(I,J,:)= [0 0 255];                % Água
                metros_agua=metros_agua+1;
            case 3
                IGUATU_FINAL(I,J,:)= [255 255 255];            % Antropizada
                metros_antropizada=metros_antropizada+1;
            case 4
                IGUATU_FINAL(I,J,:)= [0 0 0];                %CHA
                metros_caatinga=metros_caatinga+1;
            case 5
                IGUATU_FINAL(I,J,:)= [255 0 0];                %CAD
                metros_vegetacaonativa=metros_vegetacaonativa+1;
            end
        end
    end
end
end

```

```
% -----  
% Cálculo da área  
% -----  
  
metros_agricultura=metros_agricultura*900;  
metros_agua=metros_agua*900;  
metros_antropizada=metros_antropizada*900;  
metros_caatinga=metros_caatinga*900;  
metros_vegetacaonativa=metros_vegetacaonativa*900;  
  
% -----  
% Salvar a imagem em um diretório  
% -----  
  
figure();  
image(IGUATU_FINAL);  
imwrite(IGUATU_FINAL,('E:\_Beatriz\IGUATU_FINAL_MLP_TESTE_5.TIF'));  
save();
```

APÊNDICE 2

```

% -----
% Método SVM
% Universidade Federal do Ceará
% Departamento de Engenharia Agrícola
% Curso de mestrado em Engenharia Agrícola
% Beatriz F. S. Sousa
% Orientador: Prof. Adunias dos Santos Teixeira (UFC)
% Coorientador: Prof. Francisco de Assis Tavares (INPE)
% -----

PASTA='E:\_Beatriz \';

% Imagem a ser classificada
IGUATU=imread(cat(2,PASTA,'imagem_teste_1_2008.tif'));

% Amostras de treinamento
agricultura=load(cat(2,PASTA,'agricultura_RGB.txt'));
agua=load(cat(2,PASTA,'agua_RGB.txt'));
antropizada=load(cat(2,PASTA,'antropizada_RGB.txt'));
caatinga=load(cat(2,PASTA,'CHA_RGB.txt'));
vegetacaonativa=load(cat(2,PASTA,'CAD_RGB.txt'));

agricultura=agricultura(1:2000,:);
agua=agua(1:2000,:);
antropizada=antropizada(1:2000,:);
caatinga=caatinga(1:2000,:);
vegetacaonativa=vegetacaonativa(1:2000,:);

metros_agricultura=0;
metros_agua=0;
metros_antropizada=0;
metros_caatinga=0;
metros_vegetacaonativa=0;

% Armazenando de dimensões da imagem e número de pixels
W=size(IGUATU);
W_pixel=W(1)*W(2);

% Armazenando quantidade de exemplos da classe
SIZE_AGRICULTURA=size(agricultura);
SIZE_AGUA=size(agua);
SIZE_ANTROPIZADA=size(antropizada);
SIZE_CAATINGA=size(caatinga);
SIZE_VEGETACAONATIVA=size(vegetacaonativa);

```

```

% Transformando a imagem para vetor
for I=1:W(1)
    for J=1:W(2)
        IGUATU_VETOR((I-1)*W(2)+J,:)= IGUATU(I,J,:);
    end;
end;

%-----
% Definição das variáveis de entrada
%-----
% Matriz de treino
treino_mlp = [agricultura;agua;antropizada;caatinga;vegetacaonativa];
treino_mlp = treino_mlp';

% matriz de saída
saida_treino = [ones(1,SIZE_AGRICULTURA(1)) 2.*ones(1,SIZE_AGUA(1))
3.*ones(1,SIZE_ANTROPIZADA(1)) 4.*ones(1,SIZE_CAATINGA(1))
5.*ones(1,SIZE_VEGETACAONATIVA(1))];

%-----
% Criando SVM
%-----
n=2000; % Número de exemplos por classe
nbtest=2000; % Número de exemplos por classe no Teste
nbiter=1; % Número de iterações
nbclass=5; % Número de classes
xapp=treino_mlp';

for i=1:nbiter
    fprintf('%d... ',i);
    yapp=[];
    yindice=[];

    for j=1:nbclass
        yapp=[yapp;j*ones(n,1)];
    end;

%-----Parâmetros da rede-----%
c = 1000;
kerneloption= 2;
kernel='gaussian';

```



```

%-----Treinamento da rede-----%

tic
[xsup,w,b,nbsv,classifier]=svmmultclassoneagainstone(xapp,yapp,nbclass,c,lambda,kernel,ke
rneloption,verbose);
end;
toc
treinando_SVM=toc

%-----
Teste da rede
%-----
tic
teste_mlp = double(IGUATU_VETOR);
teste_mlp = teste_mlp';
teste_mlp = trannmx(teste_mlp, min_treino_mlp, max_treino_mlp);
teste_mlp = teste_mlp';

for i=1:W_pixel
    [aux,maxi]=svmmultivalneagainstone(teste_mlp(i,:),xsup,w,b,nbsv,kernel,kerneloptio
n);
    teste_mlp_out(:,i)=aux(:,1);
end;
toc
teste_svm=toc

% -----
% Tratamento dos resultados
% -----
for I=1:W(1)
    for J=1:W(2)
        switch resposta_final((I-1)*W(2)+J)
            case 1
                IGUATU_FINAL(I,J,:)= [0 255 0];                % Agricultura
                metros_agricultura=metros_agricultura+1;
            case 2
                IGUATU_FINAL(I,J,:)= [0 0 255];                % Água
                metros_agua=metros_agua+1;
            case 3
                IGUATU_FINAL(I,J,:)= [255 255 255];            % Antropizada
                metros_antropizada=metros_antropizada+1;
            case 4
                IGUATU_FINAL(I,J,:)= [0 0 0];                % CHA
                metros_caatinga=metros_caatinga+1;
            case 5
                IGUATU_FINAL(I,J,:)= [255 0 0];                % CAD
                metros_vegetacaonativa=metros_vegetacaonativa+1;
            end
        end
    end
end
end

```

```
% -----  
% Cálculo da área  
% -----  
metros_agricultura=metros_agricultura*900;  
metros_agua=metros_agua*900;  
metros_antropizada=metros_antropizada*900;  
metros_caatinga=metros_caatinga*900;  
metros_vegetacaonativa=metros_vegetacaonativa*900;  
  
% -----  
% Salvar a imagem em um diretório  
% -----  
IGUATU_FINAL=uint8(IGUATU_FINAL);  
figure();  
image(IGUATU_FINAL);  
imwrite(IGUATU_FINAL,('E:\_Beatriz\IGUATU_FINAL_TESTE_49.TIF'));  
save();
```