

UNIVERSIDADE FEDERAL DO AMAZONAS
FACULDADE DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

JOEL PARENTE DE OLIVEIRA

DETECÇÃO DE ÁREAS DESMATADAS NA PORÇÃO SUL DO ESTADO DO
AMAZONAS, UTILIZANDO TÉCNICAS DE EXTRAÇÃO DE CARACTERÍSTICAS E
REDES NEURAIIS
ARTIFICIAIS

MANAUS
2017

UNIVERSIDADE FEDERAL DO AMAZONAS
FACULDADE DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

JOEL PARENTE DE OLIVEIRA

DETECÇÃO DE ÁREAS DESMATADAS NA PORÇÃO SUL DO ESTADO DO
AMAZONAS, UTILIZANDO TÉCNICAS DE EXTRAÇÃO DE CARACTERÍSTICAS E
REDES NEURAIIS
ARTIFICIAIS

Dissertação apresentada ao Curso de
Mestrado em Engenharia Elétrica, área de
concentração Controle e Automação de
Sistemas do Programa de Pós-Graduação em
Engenharia Elétrica da Universidade Federal
do Amazonas.

Orientador: Prof. Dr. Cícero Ferreira Fernandes Costa Filho
Co-Orientadora: Prof^a. Dra. Solange dos Santos Costa

MANAUS
2017

Ficha Catalográfica

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

Oliveira, Joel Parente de
O48d Detecção de áreas desmatadas na porção sul do estado do amazonas, utilizando técnicas de extração de características e redes neurais artificiais / Joel Parente de Oliveira. 2017
148 f.: il. color; 31 cm.

Orientador: Cícero Ferreira Fernandes Costa Filho
Coorientadora: Solange dos Santos Costa
Dissertação (Mestrado em Engenharia Elétrica) - Universidade Federal do Amazonas.

1. Sensoriamento Remoto. 2. Redes Neuras Artificias. 3. Desmatamento. 4. Seleção Escalar de Características. 5. Análise de Componentes Principais. I. Costa Filho, Cícero Ferreira Fernandes II. Universidade Federal do Amazonas III. Título

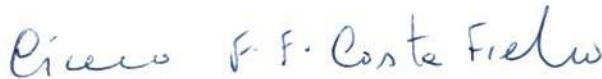
JOEL PARENTE DE OLIVEIRA

DETECÇÃO DE ÁREAS DESMATADAS NA PORÇÃO SUL DO ESTADO DO
AMAZONAS, UTILIZANDO TÉCNICAS DE EXTRAÇÃO DE
CARACTERÍSTICAS E REDES NEURAIS ARTIFICIAIS

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Amazonas, como requisito parcial para obtenção do título de Mestre em Engenharia Elétrica na área de concentração Controle e Automação de Sistemas.

Aprovado em 02 de junho de 2017.

BANCA EXAMINADORA



Prof. Dr. Cícero Ferreira Fernandes Costa Filho, Presidente

Universidade Federal do Amazonas- UFAM



Prof. Dr. Carlos Benedito Santana da Silva Soares, Membro

Centro Gestor e Operacional do Sistema de Proteção da Amazônia-

CENSIPAM



Prof. Dr. José Raimundo Gomes Pereira, Membro, Membro

Universidade Federal do Amazonas- UFAM

AGRADECIMENTOS

Primeiro a Jeová Deus (Salmos 83:18). Segundo Apo 4:11, Ele merece toda honra pois criou todas as coisas e tudo veio a existência segundo a sua vontade.

A minha querida esposa Evelyn Oliveira pelo apoio e compreensão durante todo o período de realização do meu mestrado.

Ao meu orientador, Prof. Dr. Cícero Ferreira Fernandes Costa Filho e a Prof. Marly Guimarães Fernandes Costa pelo apoio e grande ajuda prestada durante todo o desenvolvimento do trabalho. Meus sinceros agradecimentos.

À minha co-orientadora Dr. Solange dos Santos Costa que com seus conhecimentos sempre esteve pronta a colaborar, tirar dúvidas e opinar. Obrigado pela dedicação e apoio prestados.

Ao meu irmão Prof. Dr. Jozias Parente de Oliveira que sempre esteve disponível para tirar dúvidas e sugerir correções no trabalho.

Ao colega de trabalho Dr. Carlos Benedito Santana da Silva Soares pelo auxílio oferecido que certamente contribuíram para melhoria deste trabalho.

RESUMO

A realização do monitoramento do desmatamento na Amazônia é um grande desafio devido, entre outros fatores, a enorme extensão territorial. Para facilitar esse trabalho, imagens geradas a partir de sistemas de sensoriamento remoto têm sido utilizadas para realizar a detecção de regiões desmatadas. Nesse sentido, desde 1988, Instituto Nacional de Pesquisas Espaciais realiza o monitoramento da Floresta Amazônica por meio de imagens ópticas LANDSAT adquiridas por meio de sistemas de sensoriamento remoto. Atualmente, a metodologia adotada pelo INPE tem por base realizar a demarcação das áreas desmatadas de forma manual baseado no conhecimento de especialistas. Porém, realizar esse trabalho de maneira manual, demanda tempo e grande esforço. Na literatura, podemos encontrar diversos trabalhos que propõem realizar uma classificação automática de imagens de sensoriamento remoto. Porém, com base na revisão realizada, os trabalhos propostos não utilizam métodos de generalização ou critérios objetivos para extração de características. Este trabalho propõe uma metodologia para o estudo de áreas de desmatamento utilizando imagens ópticas LANDSAT-8/OLI obtidas por sensoriamento remoto. É proposta a associação de métodos para melhorar a generalização das redes neurais, como a parada antecipada e regularização, com técnicas de extração de características, como seleção escalar, análise de componentes principais e análise de discriminante linear. O desempenho da metodologia é avaliado utilizando medidas como precisão, sensibilidade, especificidade, área sob curva ROC. Por fim, foram calculadas métricas de similaridade, Dice e Jaccard. Os melhores resultados foram obtidos com o método de generalização parada antecipada associado à técnica de análise discriminante linear. O estudo foi realizado em três regiões do Estado do Amazonas, localizadas nos municípios de Apuí, Humaitá e Lábrea. Além disso, a metodologia proposta foi comparada com três trabalhos da literatura que também utilizaram imagens ópticas da região amazônica.

Palavras-chave: Sensoriamento Remoto, Redes Neuras Artificias, Parada Antecipada, Regularização, Seleção Escalar de Características, Análise de Componentes Principais, Análise de Discriminante Linear.

ABSTRACT

The monitoring of deforestation in the Amazon is a great challenge due to, among other factors, the enormous territorial extension. In this sense, images generated from remote sensing systems have been used to perform the detection of deforested regions. To facilitate this work, images generated from remote sensing systems have been used to perform the detection of deforested regions. In this sense, since 1988, the National Institute of Space Research has been monitoring the Amazon Forest through LANDSAT optical images acquired through remote sensing systems. Currently, the methodology adopted by INPE is based on the manual demarcation of deforested areas based on the knowledge of specialists. However, doing this work manually requires time and effort. In the literature, we can find several works that propose to perform an automatic classification of remote sensing images. This work proposes a methodology for the study of deforestation areas using LANDSAT-8/OLI optical images obtained by remote sensing. It is proposed the association of methods to improve the generalization of neural networks, such as early stopping and regularization, with characteristic extraction techniques, such as scalar selection, principal component analysis and linear discriminant analysis. The performance of the methodology is evaluated using measures such as precision, sensitivity, specificity, area under ROC curve. Finally, similarity metrics, Dice and Jaccard, were calculated. The best results were obtained with the early stop generalization method associated to the linear discriminant analysis technique. The study was carried out in three regions of the State of Amazonas, located in the municipalities of Apuí, Humaitá and Lábrea. In addition, the proposed methodology was compared with three papers in the literature that also used optical images from the Amazon region.

Keywords: Remote Sensing, Artificial Neural Networks, Early Stopping, Regularization, Characteristic Scalar Selection, Principal Component Analysis, Linear Discriminant Analysis.

LISTA DE ILUSTRAÇÕES

Figura 1 – Obtenção de imagem por sensoriamento remoto (FLOREZANO, 2011).....	34
Figura 2 – Espectro eletromagnético (MundoGeo)	35
Figura 3 – Exemplo do efeito da resolução espectral no contraste do rio com a vegetação. Imagens da região amazônica nas bandas: (a) visível e (b) infravermelho próximo. (MENESES e ALMEIDA, 2012)	38
Figura 4 – Exemplos de imagens com diferentes níveis de quantização ou de resolução radiométrica. (MENESES e ALMEIDA, 2012).....	39
Figura 5 – Modelo não linear de um neurônio. Fonte: Adaptado de Haykin (2004)	43
Figura 6 – Rede BP Fonte: PATROCINIO (2004).....	47
Figura 7 – Estado do Amazonas com as regiões analisadas.....	57
Figura 8 – Identificação das imagens LANDSAT-8/OLI para a Região de Apuí.....	58
Figura 9 – Identificação das imagens LANDSAT-8/OLI para a Região de Humaitá.....	58
Figura 10 – Identificação das imagens LANDSAT-8/OLI para a Região de Lábrea.....	59
Figura 11 – Diagrama em blocos com as etapas de reconhecimento de padrões.....	60
Figura 12 – Região de Apuí.....	61
Figura 13 – Região de Humaitá.....	62
Figura 14 – Região de Lábrea.....	62
Figura 15 – Regiões utilizadas por Reis (2014) e Zanotta (2014) no estado do Pará.	63
Figura 16 – Ampliação da região utilizada por Reis (2014)	64
Figura 17 – Ampliação da região utilizada por Zanotta (2014)	65
Figura 18 – Estado de Mato Grosso (Região utilizada por Pimentel, 2014).....	65
Figura 19 – Região utilizada por Pimentel (2014)	66
Figura 20 – Estrutura utilizada	70
Figura 21 – Exemplo de (a) sobreposição de duas funções de densidade de probabilidade de uma característica em duas classes (b) Curva ROC.	78
Figura 22 – Mapa de classificação. Apuí	115
Figura 23 – Mapa classificação. Apuí	116
Figura 24 – Mapa classificação Humaitá	117
Figura 25 – Mapa classificação Humaitá	118
Figura 26 – Mapa classificação Lábrea	119
Figura 27 – Mapa classificação Lábrea	120
Figura 28 – Imagem 229/065 LANDSAT-8/OLI.....	129
Figura 29 – Imagem 229/066 LANDSAT-8/OLI.....	130
Figura 30 – Imagem 230/064 LANDSAT-8/OLI.....	130
Figura 31 – Imagem 230/065 LANDSAT-8/OLI.....	131
Figura 32 – Imagem 230/066 LANDSAT-8/OLI.....	131
Figura 33 – Imagem 231/064 LANDSAT-8/OLI.....	132
Figura 34 – Imagem 231/065 LANDSAT-8/OLI.....	132
Figura 35 – Imagem 231/066 LANDSAT-8/OLI.....	133
Figura 36 – Imagem 232/064 LANDSAT-8/OLI.....	133
Figura 37 – Imagem 232/065 LANDSAT-8/OLI.....	134
Figura 38 – Imagem 232/066 LANDSAT-8/OLI.....	134
Figura 39 – Imagem 001/065 LANDSAT-8/OLI.....	135
Figura 40 – Imagem 001/066 LANDSAT-8/OLI.....	135

Figura 41 – Imagem 001/067 LANDSAT-8/OLI.....	136
Figura 42 – Imagem 002/065 LANDSAT-8/OLI.....	136
Figura 43 – Imagem 233/065 LANDSAT-8/OLI.....	137
Figura 44 – Imagem 233/066 LANDSAT-8/OLI.....	137
Figura 45 – Imagem 233/066 LANDSAT-8/OLI.....	138

LISTA DE TABELAS

Tabela 1 – Taxa de acerto para o trabalho de Pimentel (2014).....	21
Tabela 2 – Resumo dos conjuntos de dados utilizados no trabalho de Reis (2014).....	22
Tabela 3 – Taxa de acerto (REIS, 2014)	23
Tabela 4 – Valores médios de Acurácia, Sensibilidade e Especificidade, com os respectivos	23
Tabela 5 – Valores médios de acurácia com os respectivos desvios padrão para a região	
227/68 (Barbosa <i>et al.</i> , 2015)	24
Tabela 6 – Bandas LANDSAT-8/OLI USGS (2016).....	40
Tabela 7 – Características do computador utilizado.....	56
Tabela 8 – Lista de Imagens utilizadas.....	57
Tabela 9 – Lista de Regiões.....	61
Tabela 10 – Imagens LANDSAT-5/TM utilizadas	63
Tabela 11 – Lista de Características Utilizadas.....	68
Tabela 12 – Sumário dos passos metodológicos realizados.....	71
Tabela 13 – Detalhamento da divisão dos grupos nos cinco experimentos segundo a técnica de	
validação cruzada com 5 pastas.....	75
Tabela 14 – Valores de FDR para o conjunto de variáveis - Apuí.....	80
Tabela 15 – Conjunto de Características selecionados - Apuí	81
Tabela 16 – Valores de FDR para o conjunto de variáveis - Humaitá	82
Tabela 17 – Conjunto de Características selecionados - Humaitá	82
Tabela 18 – Valores de FDR para o conjunto de variáveis - Lábrea.....	83
Tabela 19 – Conjunto de Características selecionados - Lábrea	84
Tabela 20 – Valores de FDR para o conjunto de variáveis	84
Tabela 21 – Conjunto de Características selecionados.....	85
Tabela 22 – Resultados médios para acurácia, AUC, sensibilidade e especificidade, para os	
métodos de parada de treinamento MSE, parada antecipada e regularização, combinados as	
técnicas de extração de características SEC, PCA e LDA, para a região de Apuí.....	88
Tabela 23 – Valores para os testes de significância estatística relativos as diferenças entre as	
áreas sob a curva de ROC, para as técnicas SEC, PCA e LDA, para a região de Apuí.....	89
Tabela 24 – Valores médios para os testes de significância estatística relativos as diferenças	
entre as áreas sob a curva de ROC, para as técnicas SEC, PCA e LDA, para a região de Apuí.	
.....	89
Tabela 25 – Valores para os testes de significância estatística relativos as diferenças entre as	
áreas sob a curva de ROC, para os métodos de parada MSE, Parada Antecipada e	
Regularização, para a região de Apuí.....	89
Tabela 26 – Valores médios para os testes de significância estatística relativos as diferenças	
entre as áreas sob a curva de ROC, para os métodos de parada MSE, Parada Antecipada e	
Regularização, para a região de Apuí.....	90
Tabela 27 – Resultados médios para acurácia, AUC, sensibilidade e especificidade, para os	
métodos de treinamento MSE, parada antecipada e regularização, combinadas as técnicas de	
extração de características SEC, PCA e LDA, para a região de Humaitá.....	91
Tabela 28 – Valores para os testes de significância estatística relativos as diferenças entre as	
áreas sob a curva de ROC, para as técnicas SEC, PCA e LDA, para a região de Humaitá.	92

Tabela 29 – Valores médios para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para as técnicas SEC, PCA e LDA, para a região de Humaitá.	92
Tabela 30 – Valores para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para os métodos de parada MSE, Parada Antecipada e Regularização, para a região de Humaitá.	92
Tabela 31 – Valores médios para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para os métodos de parada MSE, Parada Antecipada e Regularização, para a região de Humaitá.	93
Tabela 32 – Resultados médios para acurácia, AUC, sensibilidade e especificidade, para os métodos de treinamento MSE, parada antecipada e regularização, combinadas as técnicas de extração de características SEC, PCA e LDA, para a região de Lábrea.....	94
Tabela 33 – Valores para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para as técnicas SEC, PCA e LDA, para a região de Lábrea.	95
Tabela 34 – Valores médios para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para as técnicas SEC, PCA e LDA, para a região de Lábrea.	95
Tabela 35 – Valores para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para os métodos MSE, Parada Antecipada e Regularização, para a região de Lábrea.	95
Tabela 36 – Valores médios para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para os métodos MSE, Parada Antecipada e Regularização, para a região de Lábrea.....	96
Tabela 37 – Resultados médios para acurácia, AUC, sensibilidade e especificidade, para os métodos de treinamento MSE, parada antecipada e regularização, combinadas as técnicas de extração de características SEC, PCA e LDA, para as três regiões (Apuí, Humaitá e Lábrea)	99
Tabela 38 – Valores para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para as técnicas SEC, PCA e LDA, para as três regiões (Apuí, Humaitá e Lábrea).	100
Tabela 39 – Valores médios para os testes de significância estatística relativos as diferenças entre as áreas sob a curva ROC, para as técnicas SEC, PCA e LDA, para as três regiões (Apuí, Humaitá e Lábrea).	100
Tabela 40 – Valores para os testes de significância estatística relativos as diferenças entre as áreas sob a curva ROC, para os métodos MSE, Parada Antecipada e Regularização, para as três regiões (Apuí, Humaitá e Lábrea).....	100
Tabela 41 – Valores médios para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para os métodos MSE, Parada Antecipada e Regularização, para as três regiões (Apuí, Humaitá e Lábrea).	101
Tabela 42 – Resultados médios para acurácia, AUC, sensibilidade e especificidade, para os métodos de treinamento MSE, parada antecipada e regularização, combinadas as técnicas de extração de características SEC, PCA e LDA, para a região analisada por Reis (2014)	104
Tabela 43 – Valores para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para as técnicas SEC, PCA e LDA, para a região analisada por Reis (2014).	105
Tabela 44 – Valores médios para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para as técnicas SEC, PCA e LDA, para a região analisada por Reis (2014).	105

Tabela 45 – Valores para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para os métodos MSE, Parada Antecipada e Regularização, para a região analisada por Reis (2014).	105
Tabela 46 – Valores médios para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para os métodos MSE, Parada Antecipada e Regularização, para a região analisada por Reis (2014).	106
Tabela 47 – Resultados médios para acurácia, AUC, sensibilidade e especificidade, para os métodos de treinamento MSE, parada antecipada e regularização, combinadas as técnicas de extração de características SEC, PCA e LDA, para a região analisada por Pimentel (2014)	108
Tabela 48 – Valores para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para as técnicas SEC, PCA e LDA, para a região analisada por Pimentel (2014).	109
Tabela 49 – Valores médios para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para as técnicas SEC, PCA e LDA, para a região analisada por Pimentel (2014).	109
Tabela 50 – Valores para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para os métodos MSE, Parada Antecipada e Regularização, para a região analisada por Pimentel (2014).	109
Tabela 51 – Valores médios para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para os métodos MSE, Parada Antecipada e Regularização, para a região analisada por Pimentel (2014).	110
Tabela 52 – Resultados médios para acurácia, AUC, sensibilidade e especificidade, para os métodos de treinamento MSE, parada antecipada e regularização, combinadas as técnicas de extração de características SEC, PCA e LDA, para a região utilizada por Zanotta (2014) ...	112
Tabela 53 – Valores para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para as técnicas SEC, PCA e LDA, para a região analisada por Zanotta (2014).	113
Tabela 54 – Valores médios para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para as técnicas SEC, PCA e LDA, para a região analisada por Zanotta (2014).	113
Tabela 55 – Valores para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para os métodos MSE, Parada Antecipada e Regularização, para a região analisada por Zanotta (2014).	113
Tabela 56 – Valores médios para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para os métodos MSE, Parada Antecipada e Regularização, para a região analisada por Zanotta (2014).	114
Tabela 57 – Resultados de sensibilidade, especificidade e coeficientes de Dice e Jacard para as três regiões apresentadas na Tabela 8	120

LISTA DE ABREVIATURAS E SIGLAS

ACUR	Acurácia
AUC	Área sob a curva
CORINE	<i>Coordination of information on the environment</i>
EVI	Índice de Vegetação Melhorado
ESP	Especificidade
FDR	<i>Fisher's Discriminant Ratio</i>
LANDSAT	<i>Land Remote Sensing Satellite</i>
LANDSAT-5/TM	<i>Land Remote Sensing Satellite / Thematic Mapper</i>
LANDSAT-7/ETM	<i>Land Remote Sensing Satellite / Enhanced Thematic Mapper Plus</i>
LANDSAT-8/OLI	<i>Land Remote Sensing Satellite / Operational Land Sensor</i>
LDA	Análise de Discriminante Linear
MLP	Perceptron Multicamadas
MLP-G	MLP utilizando o grau de ambiguidade
MLP-D	MLP usando a saída difusa desejada
MLP-GD	MLP utilizando grau de ambiguidade e saída difusa desejada
MSE	Erro Médio Quadrático
NDVI	Índice de Vegetação de Diferença Normalizada
OBT	Observação da Terra
PAN	Parada Antecipada
PCA	Análise de Componentes Principais
PRODES	Programa de Monitoramento do Desflorestamento na Amazônia Legal
REG	Regularização
ROC	<i>The Receiver Operating Characteristics</i>
SAR	Synthetic Aperture Radar – Radar de Abertura Sintética
SENS	Sensibilidade
SPRING	Sistema de Processamento de Informações Georeferenciadas
SVM	<i>Support Vector Machine</i>
TIRS	<i>Thermal InfraRed Sensor</i>

LISTA DE QUADROS

Quadro 1 – Sumário da Revisão Bibliográfica.....	27
--	----

SUMÁRIO

1	INTRODUÇÃO	16
1.1	OBJETIVOS	18
1.1.1	GERAL	18
1.1.2	ESPECÍFICOS	18
1.2	ORGANIZAÇÃO	19
2	REVISÃO BIBLIOGRÁFICA	20
3	FUNDAMENTAÇÃO TEÓRICA	33
3.1	Introdução	33
3.2	Sensoriamento Remoto	33
3.3	LANDSAT	39
3.4	PRODES	42
3.5	REDES NEURAIS ARTIFICIAS	42
3.6	SOBRETREINAMENTO (<i>OVERFITTING</i>)	49
3.7	SELEÇÃO DE CARACTERÍSTICAS	50
4	MATERIAIS E MÉTODOS	56
4.1	Materiais	56
4.2	Métodos	59
4.2.1	Aquisição de dados	60
4.2.1	Extração de Características	66
4.2.2	Classificação	69
5	RESULTADOS	80
5.1	Seleção escalar: conjunto de características selecionado	80
5.1.1	Região de Apuí	80
5.1.2	Região de Humaitá	81
5.1.3	Região de Lábrea	83
5.1.4	Três Regiões (Apuí, Humaitá, Lábrea)	84
5.2	Resultados dos classificadores	86
5.2.1	Região de Apuí	86
5.2.2	Região de Humaitá	90
5.2.3	Região de Lábrea	93
5.2.4	Avaliação dos resultados relativos a utilização das três regiões separadamente 96	
5.2.5	Dados das três Regiões (Apuí, Humaitá e Lábrea) juntas	97
5.2.6	Avaliação dos resultados relativos a utilização das três regiões conjuntamente 101	

5.2.7	Região analisada por Reis (2014).....	102
5.2.8	Região analisada por Pimentel (2014).....	106
5.2.9	Região analisada por Zanotta (2014).....	110
5.3	Resultados: Mapa de classificação	114
6	CONCLUSÕES E TRABALHOS FUTUROS	121
7	REFERÊNCIAS	123
	APÊNDICE 1	129
	APÊNDICE 2	139

1 INTRODUÇÃO

Sensoriamento remoto é obtenção de dados sobre um objeto sem estar em contato físico direto com o mesmo (JENSEN, 2009). De forma mais específica, segundo Novo (2008), pode-se definir sensoriamento remoto como sendo a utilização de diversas tecnologias, como sensores, equipamentos instalados a bordo de aeronaves, espaçonaves, ou outras plataformas, com o objetivo de estudar os fenômenos que ocorrem na superfície da Terra. Isso é possível por meio do registro e análise das interações entre a radiação eletromagnética e as substâncias que o compõem em suas diversas manifestações.

Atualmente, existem diversos sistemas de sensoriamento remoto. Os dados gerados a partir desses sistemas são de grande utilidade para várias aplicações, dentre as quais podemos citar: urbanas (interferência demográfica, cadastro, planejamento urbano, suporte ao setor imobiliário); agrícolas (condição de culturas, previsão de safras, erosão de solos; geológicas (minerais, petróleo, gás natural); ecológicas (regiões alagadas, solos, florestas, oceanos, águas continentais); florestais (produção de madeira, estimativa de biomassa, monitoramento de desmatamento) (NOVO, 2008).

Para realizar o monitoramento do desmatamento na Amazônia, as imagens geradas a partir de sistemas de sensoriamento remoto possuem grande importância, pois segundo Noma *et al.* (2013), imagens de satélite permitem realizar o monitoramento do desmatamento em larga escala e compreender a dinâmica da paisagem. Neste sentido, desde 1988, o Instituto Nacional de Pesquisas Espaciais (INPE) é o principal responsável pelo monitoramento da Floresta Amazônica. Neste aspecto, Noma *et al.* (2013) destacam dois projetos: Programa de Monitoramento do Desflorestamento na Amazônia Legal (PRODES) e TerraClass. O PRODES fornece dados por meio de mapas anuais de desmatamento na região Amazônica. O projeto TerraClass utiliza os dados gerados pelo PRODES para realizar uma classificação de uso e cobertura de solo. Nesse caso, o objetivo é identificar os vários tipos de uso nas regiões desmatadas, como: pastagem, agricultura e áreas urbanas. Com isso, é possível realizar uma fiscalização e desenvolver políticas públicas de prevenção para conter o avanço do desmatamento. De acordo com o Inpe (2016), nos anos de 2015 e 2016, o desmatamento no estado do Amazonas foi de 712 km² e 1099 km², respectivamente, segundo dados gerados pelo PRODES. Isso representa um aumento de 54% de desmatamento de 2015 para 2016.

Para agilizar e tornar mais eficiente o trabalho que envolva a análise de imagens de sensoriamento remoto, por exemplo, para detectar áreas desmatadas, diversos trabalhos foram propostos na literatura tendo como objetivo realizar uma classificação automática das áreas em estudo. Reis (2014) utilizou imagens SAR (*Synthetic Aperture Radar* – Radar de Abertura Sintética), ópticas e a combinação das duas para realizar detecções de mudanças em uma área da Floresta Nacional do Tapajós, no estado do Pará. As imagens ópticas utilizadas foram do satélite LANDSAT-5/TM e as imagens SAR foram do sensor PALSAR do satélite ALOS. Utilizando o classificador de mínima distância estocástica Bhattacharyya, Reis (2014) obteve uma exatidão global de 84,02%.

Zanotta (2014) propôs uma nova metodologia para detecção semi-supervisionada de desflorestamento na Amazônia brasileira utilizando imagens LANDSAT-7/ETM. As classes definidas por Zanotta (2014) foram desmatamento e floresta. A técnica utilizada foi um algoritmo de crescimento de regiões. A acurácia obtida por Zanotta (2014) foi de 82%.

Pimentel (2014) desenvolveu um sistema computacional para classificar padrões de uso do solo e cobertura da terra em séries temporais através de índice de vegetação. A área de estudo foi o estado do Mato Grosso e as classes definidas para classificação foram: pastagem, agricultura, floresta e desflorestamento. A classificação foi realizada utilizando quatro classificadores de redes neurais, MLP (Perceptron Multicamadas), MLP-G (MLP utilizando o grau de ambiguidade), MLP-D (MLP usando a saída difusa desejada), MLP-GD (MLP utilizando grau de ambiguidade e saída difusa desejada). Pimentel (2014) obteve como melhor resultado um acerto de 96,06% para o classificador MLP-GD.

Symeonakis (2016) desenvolveu uma metodologia para identificar mudanças ocorridas na ilha mediterrânica de Lesbos (Grécia) por meio de imagens LANDSAT-5/TM. As classes definidas foram: solo exposto, área urbana, floresta, outras culturas (irrigadas e não irrigadas) e corpos de água. Utilizando redes neurais, o autor conseguiu uma precisão de 93%.

Os trabalhos apresentados não utilizaram estratégias contra o sobre ajustamento (*overfitting*), fenômeno que ocorre quando uma rede neural é excessivamente treinada, trazendo como consequência o sobre-ajustamento dos parâmetros na tarefa de reconhecimento dos padrões apresentados durante a fase de treinamento. Uma rede excessivamente treinada possui um baixo poder de generalização, o que implica em um desempenho fraco na classificação de novos padrões apresentadas a mesma. Para evitar o *overfitting*, existem métodos de aprendizagem como Parada Antecipada (*Early Stopping*) e Regularização Bayesiana.

Outra crítica que fazemos aos trabalhos previamente publicados é que os mesmos não se preocupam em selecionar as melhores variáveis para a classificação. Existem diversas

técnicas que podem ser utilizadas para extração de características, como: Seleção Escalar de Características (SEC), Análise de Componentes Principais (*Principal Component Analysis* - PCA) ou Análise Discriminante Linear (*Linear Discriminant Analysis* - LDA).

Considerando o aumento do desmatamento na porção sul do estado do Amazonas e ausência de pesquisas envolvendo as estratégias e critérios aplicadas em análises nessa região, este trabalho ora apresentado, pretende preencher as lacunas existentes na literatura quanto a utilização de métodos para melhoria da generalização de redes neurais, na tarefa de classificação de áreas desmatadas, bem como quanto a utilização de técnicas de seleção de características que permitam uma melhoria do desempenho de um classificador, as redes neurais.

1.1 OBJETIVOS

1.1.1 GERAL

Avaliar e comparar diferentes metodologias utilizando redes neurais para classificação de área desmatada nas regiões de Apuí, Humaitá e Lábrea do estado do Amazonas a partir de imagens ópticas LANDSAT-8/OLI.

1.1.2 ESPECÍFICOS

- Comparar o desempenho de métodos para melhoria da generalização das redes neurais, como regularização (REG) e parada antecipada (PAN), com o método de parada utilizando o erro médio quadrático (MSE), na quantificação de áreas desmatadas do Amazonas;
- Avaliar a influência da redução de características, utilizando as técnicas SEC, PCA e LDA, sobre o desempenho das redes neurais;
- Avaliar o desempenho da metodologia proposta nas regiões de Apuí, Humaitá e Lábrea do estado do Amazonas;
- Avaliar a metodologia proposta em áreas utilizadas nos estudos de outros autores, com o objetivo de comparação de desempenho.

1.2 ORGANIZAÇÃO

Este trabalho está organizado em 6 capítulos descritos a seguir:

- Introdução (Capítulo 1);
- Revisão Bibliográfica (Capítulo 2);
- Fundamentação Teórica (Capítulo 3);
- Materiais e Métodos (Capítulo 4);
- Resultados (Capítulo 5);
- Discussões e Conclusões (Capítulo 6);
- Referências Bibliográficas;
- Apêndice.

No Capítulo 1, é apresentada a introdução deste trabalho.

No Capítulo 2, identificam-se os principais sistemas de classificação de imagens de satélite até então publicados, por meio de uma pesquisa bibliográfica nas principais bases de dados literárias.

No Capítulo 3, apresenta-se a fundamentação teórica necessária ao desenvolvimento do presente trabalho por meio da descrição das ferramentas utilizadas.

No Capítulo 4, apresenta-se o conjunto de dados e os passos metodológicos utilizados.

No Capítulo 5, exibem-se os resultados obtidos com o método proposto, utilizando medidas tais como: acurácia (ACUR), sensibilidade (SENS), especificidade (ESP), área sob a curva (AUC) e desvio padrão da AUC.

No Capítulo 6, apresentam-se as conclusões do trabalho e as sugestões de trabalhos futuros.

No Apêndice 1 encontra-se as imagens LANDSAT-8/OLI das regiões de Apuí, Humaitá e Lábrea analisadas neste trabalho de dissertação.

Por fim, no Apêndice 2 encontra-se uma cópia do artigo publicado intitulado: “Detecção de áreas desmatadas utilizando redes neurais artificiais associados a métodos para melhoria da generalização”, apresentado no XIII Encontro Anual de Computação, realizado na cidade de Catalão, Goiás, de 24 a 26 de maio de 2017.

2 REVISÃO BIBLIOGRÁFICA

Neste Capítulo é apresentada uma revisão bibliográfica dos principais trabalhos na área de classificação do solo utilizando imagens de sensoriamento remoto, a partir de uma busca feita nas bases de dados do *Institute of Electrical and Electronic Engineers* (IEEE), do Instituto Nacional de Pesquisas Espaciais (INPE) e do *ScienceDirect*.

Yin *et al.* (2014) desenvolveram uma abordagem automatizada para o mapeamento da cobertura da terra em intervalo de tempo anual, no período de 2000 a 2011, utilizando imagens MODIS. A área de estudo escolhida foi o interior da Mongólia, área em que, na última década, ocorreram grandes mudanças na cobertura da terra. As classes definidas para reconhecimento foram: cultivo, floresta, pastagem e áreas não vegetais (água, áreas construídas e solo exposto). Para reduzir os níveis de ruído nas imagens utilizou-se o filtro Savitzky-Golay. Os dados de amostra para treinamento foram extraídos do CAS-LUCC *database*, imagens LANDSAT-5/TM e imagens do *Google Earth*. As características extraídas foram média dos pixels, valor mínimo do pixel, faixa dos pixels e desvio padrão. Os pixels das imagens foram classificados segundo a probabilidade de pertencer a determinada classe. Segundo os autores do trabalho, as probabilidades estão mais ligadas às mudanças de cobertura da terra do que a valores espectrais ou índices de vegetação. Utilizando um classificador *random forest* disponível no software *EnMAP-BOX* foi obtida uma taxa de acerto de 92%.

Wang *et al.* (2009) classificaram imagens da região da bacia do rio Dongjiang (China) utilizando índices de vegetação EVI. Para a filtragem das imagens foi utilizado o filtro *Hant*. As amostras para treinamento foram geradas por meio de inspeção visual, a partir de imagens LANDSAT-5/TM. Foram utilizados também dados de campo e dados de um modelo de elevação. As características extraídas foram: média dos pixels, valor mínimo de pixel e valor máximo de pixel. Através de um classificador de árvore de decisão, sete classes (agricultura, pasto, represa, rio, matagal, floresta e área urbana) foram identificadas com uma precisão de 87,1%.

Kuck *et al.* (2011) apresentaram uma metodologia de classificação do uso da terra baseada na análise de séries multitemporais de produtos MODIS. A área de estudo foi o estado do Mato Grosso, restringindo-se apenas a área compreendida pelo bioma amazônico. O período escolhido foi de 2001 a 2009. Os dados utilizados como amostras foram gerados a partir de dados de campo e inspeção visual de imagens LANDSAT-7/ETM. As características extraídas foram média dos pixels, variância dos pixels, médias mínima e máxima dos pixels. Utilizando

um classificador de árvore de decisão, as seguintes classes foram classificadas: pasto, agricultura, floresta e floresta alterada. A precisão do método foi de 94,9%.

Pimentel (2014) desenvolveu um sistema computacional que classificou padrões de uso do solo e cobertura da terra em séries temporais através de índice de vegetação. A área de estudo selecionada foi o estado do Mato Grosso e as classes definidas para classificação foram: pastagem, agricultura, floresta e desflorestamento. As características utilizadas foram média dos pixels e desvio padrão. O conjunto de dados de amostra foram gerados a partir de inspeção visual. Dois conjuntos de dados foram gerados, um com 205 amostras e outro com 1363 amostras. A classificação foi realizada utilizando quatro classificadores, MLP (Perceptron Multicamadas), MLP-G (MLP utilizando o grau de ambiguidade), MLP-D (MLP usando a saída difusa desejada), MLP-GD (MLP utilizando grau de ambiguidade e saída difusa desejada). Os dois conjuntos de amostra foram aplicados em cada classificador. A taxa de acerto de classificação é apresentada na Tabela 1.

Tabela 1 – Taxa de acerto para o trabalho de Pimentel (2014)

Classificador	Medida de Acerto (%)	
	Conjunto de dados I	Conjunto de dados II
MLP	89,22	93,16
MLP-G	93,45	93,50
MLP-D	92,50	92,05
MLP-GD	96,05	93,55

Em relação aos dados do Terraclass2010, a acurácia do classificador desenvolvido por Pimentel (2014) foi de 81%.

Costa *et al.* (2015) desenvolveram um trabalho com o objetivo de classificar e criar mapas temáticos de áreas de formações campestres nativas e regiões de pastagens cultivadas no Cerrado brasileiro, a partir da análise de séries temporais EVI disponíveis no produto MODIS. O estudo realizado produziu um mapeamento das imagens nas seguintes classes: áreas de formações campestres nativas e regiões de pastagens cultivadas no Cerrado brasileiro. O conjunto de dados para treinamento foi gerado a partir de dois *shapefiles* que fornecem polígonos que delimitam áreas de pastagem nativa e cultivada. O primeiro é um mapa de cobertura vegetal do cerrado fornecido pelo Inventário Florestal de Minas Gerais do ano de 2009. Esse mapa foi utilizado para a obtenção das formações campestres nativas. O segundo *shapefile* é um mapa disponibilizado pelo Ministério do Meio Ambiente, que fornece dados relativos ao ano de 2006 sobre pastagens cultivadas de cerrado. A característica utilizada foi a média dos valores de todos os pixels, no caso referente a imagem do índice de vegetação EVI, associado a região delimitada por cada polígono. Utilizando os classificadores SVM (*Support*

Vector Machine), árvore de decisão e *random forest*, foram obtidas taxas de precisões 85,96%, 76,06% e 81,87%, respectivamente.

Reis (2014) avaliou o potencial de imagens SAR, ópticas e a combinação das duas para detectar mudanças em uma área da Floresta Nacional do Tapajós. As imagens ópticas utilizadas no trabalho são oriundas do satélite LANDSAT-5/TM. Já as imagens de micro-ondas são do sensor PALSAR do satélite ALOS. As imagens LANDSAT-5/TM datam de 23 de junho de 2008 e de 29 de junho de 2010 e as imagens ALOS/PALSAR datam de 15 de junho de 2008 e 21 de junho de 2010. Segundo o autor, essas imagens foram escolhidas por serem as datas mais próximas em que havia imagens dos dois sensores disponíveis e em épocas nas quais espera-se que as condições climáticas e do calendário de atividade agrícola sejam semelhantes. A partir dos dois conjuntos de imagens citados, foram gerados quatro conjuntos de dados, totalizando seis conjuntos de dados. A Tabela 2 apresenta as características desses seis conjuntos de dados.

Tabela 2 – Resumo dos conjuntos de dados utilizados no trabalho de Reis (2014)

Conjunto de Dados	Composição
TM	Bandas selecionadas das imagens LANDSAT-5/TM ortorretificadas
Atributos TM	Atributos selecionados do conjunto formado por medidas de textura calculados sobre o conjunto TM, PCs das bandas 1 a 5 e 7 das imagens LANDSAT5/TM ortorretificadas e fração vegetação, solo e sombra calculadas usando as bandas 1 a 5 e 7 das imagens LANDSAT-5/TM ortorretificadas
PALSAR	Polarizações HH e HV das imagens ALOS/PALSAR ortorretificadas
PALSAR_SDNLM	Conjunto PALSAR filtrado utilizando o filtro SDNLM
Atributos PALSAR_SDNLM	Atributos selecionados de textura calculados sobre PALSAR_SDNLM
Fusão	Imagens LANDSAT-5/TM ortorretificadas e PALSAR_SDNLM fusionadas por SPC-SAR

Reis (2014) baseou-se em trabalhos de campo para definir dez classes de cobertura da terra, utilizadas para a classificação. As classes escolhidas pelo autor foram: Floresta Primária, Floresta Degradada, Vegetação Secundária Inicial, Vegetação Secundária Intermediária, Vegetação Secundária Avançada, Pasto Limpo, Pasto Sujo, Área Cultivada, Área em Repouso e Solo Exposto. Reis (2014) também definiu duas classes que foram mapeadas de forma manual, pois não possuíam amostras suficientes para a classificação. Essas classes foram Água e Sem Informação. O conjunto de dados de treinamento foi criado a partir de interpretação

visual de imagens LANDSAT-5/TM, dados do projeto TerraClass e algumas fotos georeferenciadas de 2008 presentes na biblioteca digital de fotos da Observação da Terra (OBT) (INPE, 2013). Utilizando o classificador de mínima distância estocástica Bhattacharyya, foram obtidas taxas de acerto conforme mostradas na Tabela 3.

Tabela 3 – Taxa de acerto (REIS, 2014)

Conjunto de dados	Medida de acerto (%)
TM	98
Atributos TM	97,4
Fusão	97,6
PALSAR	60,5
PALSAR_SDNLM	67
Atributos PALSAR_SDNLM	83,9

Noma *et al.* (2013) desenvolveram uma metodologia para classificação do uso da terra por meio da utilização de índices de vegetação EVI disponíveis nas séries temporais MODIS e imagens LANDSAT-5/TM. A área de estudo foi o estado do Mato Grosso. Segundo o autor, nesta região existem extensas áreas com concentração de agricultura, pastagem e floresta (HUETE *et al.*, 2002). Noma *et al.* (2013) definiram as seguintes classes para classificação: agricultura, floresta e pasto limpo. O conjunto de dados de treinamento e testes foi construído a partir de informações do projeto TerraClass2008. Noma *et al.* (2013) utilizaram cinco classificadores: Árvore de Decisão (DT), Bayes Ingênuo (NB), Vizinho Mais Próximo (NN), Máquina de Vetores de Suporte (SVM), Floresta de Caminhos Ótimos (OPF). O classificador SVM obteve uma maior taxa de acerto em relação aos demais. Os melhores resultados foram obtidos para a classificação da imagem LANDSAT-5/TM correspondente a órbita/ponto 227/68, conforme demonstrado na Tabela 4.

Tabela 4 – Valores médios de Acurácia, Sensibilidade e Especificidade, com os respectivos desvios padrão, obtidos a partir da validação cruzada k-fold para a região órbita/ponto 227/68 (Noma *et al.*, 2013)

Método	Acurácia	Sensibilidade	Especificidade
NB	0,8496 ± 0,0223	0,8705 ± 0,0211	0,8286 ± 0,0318
NN	0,8254 ± 0,0256	0,8169 ± 0,0294	0,8339 ± 0,0337
DT	0,8259 ± 0,0276	0,8243 ± 0,0303	0,8274 ± 0,0398
OPF	0,8164 ± 0,0240	0,8015 ± 0,0276	0,8312 ± 0,0342
SVM	0,8896 ± 0,0120	0,9166 ± 0,0176	0,8625 ± 0,0161

Barbosa *et al.* (2015) utilizaram a metodologia de Noma *et al.* (2013) e a mesma área de estudo para classificar a região em quatro classes distintas, quais sejam: agricultura, floresta, não-floresta, pasto limpo. Além disso, Barbosa *et al.* (2015) utilizaram um novo classificador,

método de votação (VOT), para classificar o conjunto de dados. Em relação a região correspondente a imagem 227/68 (Imagem LANDSAT-5/TM), os resultados mostraram que o classificador SVM obteve uma maior taxa de acerto, conforme mostrado na Tabela 5.

Tabela 5 – Valores médios de acurácia com os respectivos desvios padrão para a região 227/68 (Barbosa *et al.*, 2015)

Ano	Método	Classe 1	Classe 2	Classe 3	Classe 4
2008	NB	80,46 ± 5,05	87,36 ± 1,95	47,59 ± 4,31	86,10 ± 4,64
	NN	77,70 ± 3,64	79,77 ± 4,51	62,18 ± 4,32	84,14 ± 2,64
	DT	75,06 ± 2,77	79,43 ± 3,36	60,58 ± 5,79	83,79 ± 2,26
	OPF	74,83 ± 3,73	78,85 ± 4,76	62,41 ± 4,27	83,79 ± 3,13
	SVM	87,47 ± 3,57	93,45 ± 1,88	72,07 ± 6,11	91,15 ± 3,47
	VOT	84,88 ± 3,26	89,43 ± 1,98	64,94 ± 3,99	89,66 ± 2,57
2010	NB	82,22 ± 3,33	86,91 ± 3,48	49,60 ± 7,17	83,73 ± 2,93
	NN	75,32 ± 3,60	79,13 ± 3,28	61,19 ± 3,81	79,52 ± 5,03
	DT	74,76 ± 3,36	79,60 ± 4,60	59,52 ± 5,96	78,02 ± 3,74
	OPF	73,25 ± 3,51	78,10 ± 3,76	61,35 ± 3,78	78,57 ± 4,76
	SVM	84,68 ± 3,43	90,64 ± 1,97	65,56 ± 4,10	91,67 ± 3,29
	VOT	82,64 ± 3,77	87,14 ± 3,31	62,70 ± 2,88	87,30 ± 3,67

Zanotta (2014) propôs uma nova metodologia para detecção semi-supervisionada de desflorestamento na Amazônia brasileira utilizando imagens LANDSAT-5/TM. Segundo o autor, a metodologia atual do PRODES, utiliza detecção visual de incrementos a partir de imagens da série de satélites LANDSAT-5/TM. Ainda segundo o autor, o procedimento atualmente adotado produz resultados com precisão adequada, porém baseia-se em regras subjetivas de cada intérprete e possui heterogeneidade espaço-temporal, que prejudica a uniformidade do mapeamento. O autor utilizou um algoritmo baseado em crescimento de regiões para segmentar a imagem em regiões homogêneas. O critério de similaridade utilizado no algoritmo de segmentação foi a média e as informações estatísticas. A classificação proposta por Zanotta (2014) é realizada por meio de uma análise por vetor de mudança em coordenadas esféricas. Foi utilizado um método de seleção de características através do algoritmo C4.5 que emprega árvores de decisão na sua metodologia. O algoritmo desenvolvido pelo autor apresenta uma regra de decisão hierárquica em forma de árvore nos quais os nós superiores possuem um maior grau de generalização. As características selecionadas foram: valor espectral na banda 5, NDVI e entropia. O melhor resultado alçado foi uma acurácia de 82%.

Arango *et al.* (2016) desenvolveram uma metodologia para a delimitação automática de terra capaz de identificar áreas de agricultura utilizando algoritmos de agrupamento. Foram utilizadas imagens de satélite LANDSAT-8/OLI da região de Galiza, Espanha, um local onde

existe cultivo de uva. O algoritmo de agrupamento utilizado foi o *Partitioning Around Medoids* (PAM). Segundo os autores, o algoritmo PAM é mais robusto que o *k-means*, pois minimiza a soma de diferenças ao invés da soma de distâncias euclidianas ao quadrado. O PAM utiliza medóides como centro dos agrupamentos e estes medóides são selecionados entre os objetos a serem agrupados. Os resultados dessa abordagem se mostraram promissores. Segundo os autores, estes resultados podem ser utilizados para monitorar terras cultiváveis a partir de imagens multiespectrais de sensoriamento remoto, servindo de base por exemplo, para um censo de terras agrícolas ou auxiliando no processo de verificação do uso de terra agrícolas fornecidas pelo governo.

Li *et al.* (2016) utilizando imagens de satélite, realizaram uma comparação sistemática de diferentes técnicas de classificação baseadas em objetos. A região de estudo foi a Província de Sichuan, China. As classes escolhidas para a classificação foram: construção, floresta, área de agricultura, área desmatada e estrada. Os padrões de amostras utilizados nos classificadores foram gerados a partir da utilização do algoritmo de segmentação Baatz. Esse algoritmo de segmentação é baseado na abordagem de crescimento de regiões. Os resultados indicam que os classificadores *random forest* e máquinas de vetores de suporte são os mais adequados para classificações de áreas agrícolas.

Symeonakis (2016) desenvolveu uma metodologia para identificar mudanças ocorridas na ilha mediterrânea de Lesbos (Grécia) por meio da aplicação de redes neurais. Foram utilizadas três cenas LANDSAT-5/TM (julho de 1995, junho de 2001 e agosto de 2007), que abrangem 12 anos. Os dados de entrada para o classificador foram gerados por meio de um modelo de elevação digital, fotos aéreas de 1995, mapa de cobertura terrestre da *Coordination of information on the environment* (CORINE) e dados do *Google Earth*. As classes definidas foram: solo exposto, área urbana, floresta, outras culturas (irrigadas e não irrigadas) e corpos de água. Utilizando redes neurais, o autor conseguiu uma precisão de 93%.

CONSIDERAÇÕES FINAIS

A análise dos trabalhos mostrou que as metodologias elaboradas abordam várias técnicas de visão computacional, abrangendo classificação supervisionada e não-supervisionada.

Em relação aos métodos de classificação supervisionada, por exemplo, Yin *et al.* (2014) e Costa *et al.* (2015) utilizaram árvores de decisão e *random forests* em seus respectivos

trabalhos. Redes neurais foram empregadas no trabalho de Pimentel (2014). Noma *et al.* (2013), Barbosa *et al.* (2015) e Li *et al.* (2016) empregaram SVM como método de classificação.

Sobre os métodos de classificação não supervisionados, Arango *et al.* (2016) utilizaram o algoritmo de agrupamento *Partitioning Around Medoids* (PAM) para realizar uma delimitação automática de terra capaz de identificar áreas de agricultura. O classificador k-vizinhos mais próximos foi uma das técnicas utilizadas por Li *et al.* (2016).

O recurso de seleção de características foi utilizado apenas em dois trabalhos. Zanotta (2014) elencou 17 características e empregou em seu trabalho somente as três melhores características indicadas pelo algoritmo C4.5. Este algoritmo emprega árvores de decisão na sua metodologia. Li *et al.* (2016) utilizaram seleção de características baseada em correlação e, juntamente com o classificador SVM, obtiveram melhores resultados que aqueles obtidos com a utilização de todo o conjunto de características.

Os trabalhos apresentados nesta revisão que utilizaram redes neurais, não se preocuparam em utilizar estratégias para evitar o fenômeno do *overfitting*, fenômeno no qual, devido a um sobre-treinamento da rede, a mesma termina por memorizar os padrões apresentados nos treinamentos da rede, perdendo assim a capacidade de generalização, ou seja, de classificar corretamente novos padrões não utilizados para o treinamento da rede.

No trabalho ora apresentado nesta dissertação, métodos de aprendizagem como Parada Antecipada (*Early Stopping*) e Regularização Bayesiana, que previnem a ocorrência de *overfitting*, são utilizados para treinamento das redes neurais. Além disso, avalia-se o potencial de diversas técnicas de seleção de características, como SEC, PCA e LDA.

No Quadro 1, ao final desta seção, é apresentado um resumo dos artigos citados, esclarecendo suas características e os resultados alcançados em cada um.

Dessa forma, o presente trabalho explora as lacunas elencadas durante o processo de revisão dos trabalhos, quais sejam, a utilização de métodos de classificação com um melhor poder de generalização e uma avaliação de um rol de técnicas de seleção de características. Como metodologia de treinamento e teste será utilizado o método de validação cruzada. As imagens utilizadas foram imagens ópticas e a região de estudo escolhida foram três municípios do estado do Amazonas, Apuí, Humaitá e Lábrea. Segundo o Inpe (2016), dados levantados pelos PRODES mostraram que essas três regiões tiveram uma área desmatada de 358 km² nos anos de 2014 à 2015. A descrição completa da metodologia adotada nessa dissertação é descrita no Capítulo 4, denominado por materiais e métodos.

Quadro 1 – Sumário da Revisão Bibliográfica

Ano/autor	Título	Objetivos	Materiais	Filtragem	Características	Classificador	Medida de Desempenho e Resultado
2014 / He Yin, Dirk Pflugmacher, Robert E. Kennedy, Damien Sulla- Menashe, Patrick Hostert	<i>Mapping Annual Land Use and Land Cover Changes Using MODIS Time Series</i>	Mapeamento da cobertura/uso da terra Classes: cultivo, floresta, pastagem e áreas não vegetais (áreas construídas e terra nua).	Banco de Dados: Série de Dados MODIS Padrão ouro: <i>The Chinese Academy of Sciences' (CAS) Land Use/Land Cover Change (CAS- LUCC) Database Imagens Landsat Google Earth</i> Tamanho amostra: 3x3	Filtro Savitzky- Golay (Aproximação polinomial de mínimos quadrados locais)	média, mínimo, máximo, range e desvio padrão para primeiro semestre (Março-junho) e segundo semestre (Julho- Outubro)	Classes de probabilidade, Random Forests (RF) / árvore de decisão.	Medida de acerto: 92%
2009 / Wen Wang Hai-Xu Hu Juan Hu	<i>Land cover change detection based on MODIS 250m vegetation index time series data</i>	Mapeamento da cobertura/uso da terra Classes: agricultura, pasto, represa, rio, matagal, floresta, área urbana.	Banco de Dados: Série de Dados MODIS Padrão ouro: Inspeção visual imagens LANDSAT-5/TM, dados de campo, dados de modelo de elevação	Filtro HANT	Média, mínimo, máximo	Árvore de decisão, Classificação não- supervisionada ISODATA	Medida de acerto: 87,1%
2011/ Tahisa Neitzel Kuck Edwin Willem Hermanus Keizer Pablo Pacheco Roni Von Cascais de Lira Sâmia Amorim de Vasconcelos André Nóbrega	<i>Mapeamento multitemporal (2001-2009) do uso da terra no bioma Amazônia do estado do Mato Grosso através de imagens MODIS</i>	Mapear uso do solo nas classes: pasto, agricultura, floresta e floresta alterada Área de estudo: Mato Grosso	Banco de Dados: Série de Dados MODIS(EVI) Imagens LANDSAT- 7/ETM Padrão ouro: inspeção visual LANDSAT-7/ETM Imagens georeferenciadas de sobrevo	MNF-Inversa	Validação: Inspeção visual em imagens Landsat (1206 pontos) e dados de campo Características: média, variância, médias mínimas e médias máximas	Árvore de decisão	Medida de acerto: 94,9%

Ano/autor	Título	Objetivos	Materiais	Filtragem	Características	Classificador	Medida de Desempenho e Resultado
2014 / Toni Roberto Gomes Pimentel	<i>Classificação de padrões temporais de uso do solo e cobertura da terra em séries temporais de índice de vegetação utilizando um sistema neuro-difuso</i>	Mapear uso do solo nas classes: pastagem, agricultura, floresta e desflorestamento Área de estudo: Mato Grosso	Banco de Dados: LANDSAT-7/ETM TerraClass 2010 Padrão ouro: pixels classificados por inspeção visual (Dois conjuntos de dados: Conjunto I: 205 amostras Conjunto II: 1363 Tamanho amostra: 2x2	Não informado	Validação: Dados do próprio conjunto de dados / Comparação com TerraClass2010 Características: média e desvio padrão	MLP MLP-G MLP-D MLP-GD	Medida de acerto: I II MLP: 89,22% 93,16% MLP-G: 93,45% 93,50% MLP-D: 92,50% 92,05% MLP-GD: 96,05% 93,55% Medida de acerto com o TerraClass2010: 81%
2015 / Wanderson Santos Costa Leila Maria Garcia Fonseca Thales Sehn Körting	<i>Classificação de pastagens cultivadas e formações campestres nativa</i>	Classificar cobertura da terra classes: pastagens nativas e cultivadas no Cerrado brasileiro	Banco de Dados: Série de Dados Modis (EVI) Padrão ouro: polígonos de um mapa de cobertura vegetal do Cerrado 2009 e mapa de pastagens cultivadas 2006 (508 amostras, 337:treinamento, 171: testes)	Não informado	Validação: Dados do próprio conjunto de dados Características: mínimo, máximo, amplitude, soma, moda, média e desvio padrão	SVM, random forests, árvore de decisão	Medida de acerto: SVM: 85,96% RF: 81,87% DT: 76,02%

Ano/autor	Título	Objetivos	Materiais	Filtragem	Características	Classificador	Medida de Desempenho e Resultado																								
2014 / Mariane Souza Reis	<i>Detecção de mudanças de uso e cobertura da terra utilizando dados óticos e de micro-ondas em uma região da Amazônia brasileira</i>	Classificar cobertura da terra nas seguintes classes: Florestas, Vegetação Secundária, Pasto Limpo, Pasto Sujo, Área Cultivada, Área em Pousio, Solo Exposto, Água, Sem informação Local: município de Belterra, no estado do Pará	Banco de Dados: Imagens LANDSAT-5/TM (óptica), imagens PALSAR (SAR) Padrão ouro: Dados de campo, interpretação visual, série histórica de imagens LANDSAT-5/TM, dados do projeto TerraClass, fotos de campo de 2008 (221279 amostras, 106897:treinamento, 114382:testes)	Não informado	Validação: Dados do próprio conjunto de dados Características: atributos de textura de segunda ordem: Média, Variância, Homogeneidade, Contraste, Dissimilaridade, Entropia, Segundo Momento e Correlação	Classificador de Mínima Distância Estocástica - Bhattacharyya	Medida de acerto: TM: 79,6% Atributos TM: 84,2% Fusão: 79,8,6% Palsar: 39,8% Palsar_SDNLM: 40,09% Atributos Palsar_SDNLM: 48,2%																								
2013 / Alexandre Noma Thales Sehn Korting Leila Maria Garcia Fonseca	<i>Uma Comparação entre Classificadores usando Regiões e Perfis EVI para Agricultura</i>	Classificar cobertura/uso da terra nas seguintes classes: Agricultura, Floresta, Pasto	Banco de Dados: Imagens LANDSAT-5/TM, Imagens MODIS, TerraClass (imagem segmentada pelo algoritmo Baaz) Padrão ouro: TerraClass2008, TerraClass2010	HANT	Validação: Dados do TerraClass Características: média, variância, valor mínimo, valor máximo	DT NB NN SVM OPF VOT	<table border="1"> <thead> <tr> <th>Método</th> <th>Acurácia</th> <th>Sens.</th> <th>Espec.</th> </tr> </thead> <tbody> <tr> <td>NB</td> <td>0,8496</td> <td>0,8705</td> <td>0,8286</td> </tr> <tr> <td>NN</td> <td>0,8254</td> <td>0,8169</td> <td>0,8339</td> </tr> <tr> <td>DT</td> <td>0,8259</td> <td>0,8243</td> <td>0,8274</td> </tr> <tr> <td>OPF</td> <td>0,8164</td> <td>0,8015</td> <td>0,8312</td> </tr> <tr> <td>SVM</td> <td>0,8896</td> <td>0,9166</td> <td>0,8625</td> </tr> </tbody> </table>	Método	Acurácia	Sens.	Espec.	NB	0,8496	0,8705	0,8286	NN	0,8254	0,8169	0,8339	DT	0,8259	0,8243	0,8274	OPF	0,8164	0,8015	0,8312	SVM	0,8896	0,9166	0,8625
Método	Acurácia	Sens.	Espec.																												
NB	0,8496	0,8705	0,8286																												
NN	0,8254	0,8169	0,8339																												
DT	0,8259	0,8243	0,8274																												
OPF	0,8164	0,8015	0,8312																												
SVM	0,8896	0,9166	0,8625																												

Ano/autor	Título	Objetivos	Materiais	Filtragem	Características	Classificador	Medida de Desempenho e Resultado				
							Ano	Método	Classe 1	Classe 2	Classe 3
2015 / David Pereira Barbosa Alexandre Noma Thales Sehn Korting Leila Maria Garcia Fonseca	<i>Um Estudo Experimental com Classificadores baseados em Regiões e Perfis EVI</i>	Classificar cobertura/uso da terra nas seguintes classes: Agricultura, Floresta, Não-Floresta, Pasto	Banco de Dados: Imagens LANDSAT-5/TM, Imagens MODIS, TerraClass Padrão ouro: TerraClass2008, TerraClass2010 (imagem segmentada pelo algoritmo Baaz)	HANT	Validação: Dados do TerraClass Características: média, variância, valor mínimo, valor máximo	DT NB NN SVM OPF	2008	NB	77,22	89,26	84,26
								NN	73,70	91,20	78,61
								DT	75,09	91,48	75,09
								OPF	73,43	91,20	77,32
								SVM	78,80	93,43	90,56
								VOT	78,51	92,59	86,30
							2010	NB	77,04	92,50	83,61
								NN	75,19	91,48	76,76
								DT	77,32	90,46	75,19
								OPF	74,54	91,30	75,83
								SVM	80,19	92,87	90,83
								VOT	79,35	92,87	85,37
2014 / Daniel Capella Zanotta	<i>Uma metodologia adaptativa baseada em Objetos para detecção semi-automática de Desmatamento na amazônia por imagens de Sensoriamento remoto</i>	Classificar áreas desmatadas na região Amazônica	Banco de Dados: Imagens LANDSAT-7/TM Padrão ouro: PRODES Imagem segmentada pelo algoritmo de crescimento de regiões	Não informado	Validação: Dados do PRODES Características: valor espectral na banda 5, NDVI, entropia	Análise de vetor de mudança	Acurácia: 82%				

Ano/autor	Título	Objetivos	Materiais	Filtragem	Características	Classificador	Medida de Desempenho e Resultado
2016 / R.B. Arangoa A.M. Campos E.F. Combarroa E.R. Canas I. Diaz	<i>Mapping cultivable land from satellite imagery with clustering algorithms</i>	Identificar áreas de agricultura utilizando algoritmos de agrupamento	Banco de Dados: Imagens LANDSAT-8/OLI Padrão ouro: Utilizado o índice de qualidade <i>Calinsky-Harabasz</i> para a geração dos dados de amostra.	Não informado	Validação: Dados do próprio conjunto de dados Características: índice Calinski-Harabasz	<i>Partitioning Around Medoids</i>	Precisão: 90%
2016 / Manchun Li Lei Maa, Thomas Blaschke Liang Chenga Dirk Tiede	<i>A systematic comparison of different object-based classification techniques using high spatial resolution imagery in agricultural environments</i>	Comparar diferentes técnicas de classificação baseadas em objetos	Banco de Dados: Imagens da província de Sichuan, China, adquiridas através de um veículo aéreo não tripulado Padrão ouro: polígonos de referência. Amostras geradas pelo algoritmo Baatz.	Não informado	Validação: Dados do próprio conjunto de dados Características: Média, desvio padrão, homogeneidade, contraste, dissimilaridade, entropia, correlação, área, densidade.	<i>Random Forest</i> Árvore de Decisão K-vizinhos mais próximos SVM	SVM: 80% RF: 89% KNN: 70% DT: 80%

Ano/autor	Título	Objetivos	Materiais	Filtragem	Características	Classificador	Medida de Desempenho e Resultado
2016 / Elias Symeonakis	<i>Modelling land cover change in a mediterranean environment using Random forests and a multi-layer neural network model</i>	Identificar mudanças ocorridas na ilha mediterrânica de Lesvos (Grécia) por meio de redes neurais	Banco de Dados: Três cenas de LANDSAT-5/TM (julho de 1995, junho de 2001 e agosto de 2007) Padrão ouro: modelo de elevação digital, mapa de cobertura terrestre CORINE, dados do <i>Google Earth</i> .	Não informado	Validação: Dados do próprio conjunto de dados Características: Média, desvio padrão, área.	Rede neural	Acurácia: 93%

3 FUNDAMENTAÇÃO TEÓRICA

3.1 Introdução

Neste Capítulo são apresentados os conceitos essenciais para uma melhor compreensão desta dissertação. Os assuntos apresentados versam sobre sensoriamento remoto, imagens LANDSAT-8/OLI, índices de vegetação, PRODES, redes neurais artificiais, estratégias contra o sobre-treinamento (*overfitting*): regularização e parada antecipada, uma breve introdução sobre as medidas de separabilidade de classes e FDR. Por fim, será feita uma revisão conceitual das técnicas de extração das características: SEC, PCA e LDA.

3.2 Sensoriamento Remoto

Sensoriamento remoto é uma técnica que consiste na aquisição de informações de um objeto ou fenômeno sem o contato físico com o mesmo. Esses dados podem ser utilizados para a extração de informações, monitoramento dinâmico, mapeamento temático, classificação de imagens, etc. Essa técnica desempenha um papel vital para interpretar e identificar objetos no solo por meio de classificação de imagens, atribuindo à cada pixel ou região da imagem uma determinada categoria. (BHARATKAR *et al.*, 2013).

Segundo Meneses e Almeida (2012), sensoriamento remoto visa a obtenção de imagens da superfície terrestre por meio da detecção e medição quantitativa das respostas das interações da radiação eletromagnética com os materiais terrestres.

A radiação eletromagnética possui uma dualidade de comportamento: onda e energia. Segundo Meneses e Almeida (2012), a radiação eletromagnética que se propaga pelo espaço, como luz solar, é ao mesmo tempo uma forma de onda e uma forma de energia.

Na Figura 1, podemos observar um exemplo de obtenção de dados por meio de sensoriamento remoto. Nela, vemos que o Sol ilumina a superfície terrestre. A energia proveniente do Sol, refletida pela superfície em direção ao sensor, é captada e registrada em forma de imagem, com o posterior processamento das mesmas (FLOREZANO, 2011).

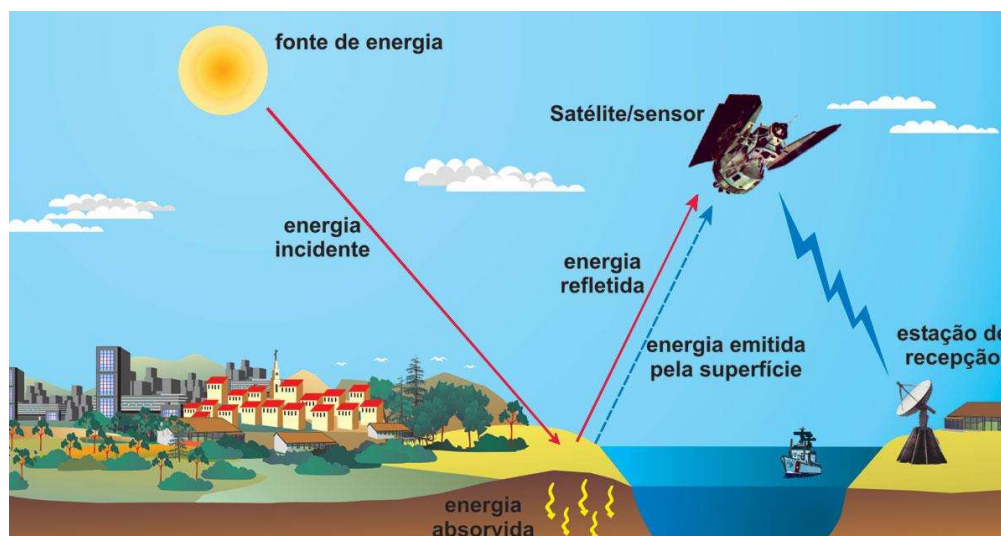


Figura 1 – Obtenção de imagem por sensoriamento remoto (FLOREZANO, 2011)

3.1.1 Fontes de Energia usadas em sensoriamento remoto

A obtenção de dados por sensoriamento remoto, como em qualquer outra atividade envolvendo equipamentos elétricos, requer o uso de energia. A energia com a qual operam os sensores remotos pode ser proveniente de uma fonte natural, como a luz do sol ou o calor emitido pela superfície da Terra, ou pode ser de uma fonte artificial como, por exemplo, a energia do *flash* utilizado em uma máquina fotográfica ou o sinal produzido por um radar (FLOREZANO, 2011).

A energia utilizada em sensoriamento remoto para produção das imagens são ondas eletromagnéticas, que se propagam à velocidade da luz. A frequência das mesmas é expressa em unidades de *hertz* (Hz) e seus múltiplos, como quilohertz ($1 \text{ kHz} = 10^3 \text{ Hz}$) e mega-hertz ($1 \text{ MHz} = 10^6 \text{ Hz}$). O comprimento de onda (λ) é medido em unidades de metro. A frequência de onda é o número de vezes que uma onda se repete por unidade de tempo, sendo diretamente proporcional à velocidade de propagação e inversamente proporcional ao comprimento da onda. Dessa maneira, como mostra a Figura 2, quanto menor for o comprimento da onda, maior será a frequência da mesma. O comprimento de onda é a distância entre dois picos de ondas sucessivas: quanto mais distantes esses picos, maior é o comprimento e, quanto menos distantes, menor será o comprimento de onda.

O espectro eletromagnético representa a distribuição da radiação eletromagnética, por regiões, segundo o comprimento de onda e a frequência. Na Figura 2, observa-se que o espectro eletromagnético abrange desde curtos comprimentos de onda, como os raios cósmicos e os raios gama (γ), de alta frequência, até longos comprimentos de onda, como as ondas de rádio e TV,

de baixa frequência. Na região do espectro visível, o olho humano enxerga a energia (luz) eletromagnética, sendo capaz de distinguir as cores do violeta ao vermelho (FLOREZANO, 2011).

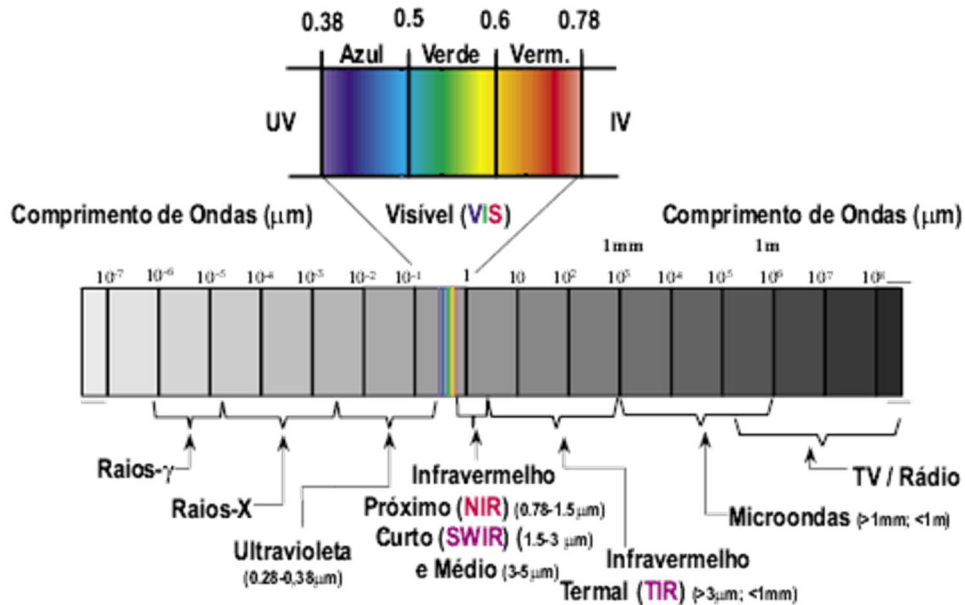


Figura 2 – Espectro eletromagnético (MundoGeo)

Segundo Meneses e Almeida (2012), o espectro eletromagnético foi arbitrariamente dividido pelo homem em intervalos de comprimentos de onda com base nos mecanismos físicos geradores da energia eletromagnética e nos mecanismos físicos de sua detecção. Segundo os autores, a nomenclatura de cada um dos intervalos foi realizada com base no uso que o homem encontrou para as suas aplicações. Meneses e Almeida (2012) acrescentam que é essencial saber os valores dos comprimentos de onda desses intervalos, as denominações que recebem, e quais comprimentos de ondas são possíveis de serem detectados por cada tipo de sensor.

A seguir é apresentado um resumo das características dos intervalos espectrais mostrados na Figura 2.

Visível (0,45-0,76 μm): Essa região do espectro apresenta a intensidade de fluxo radiante mais alta e é onde se verifica a melhor janela atmosférica, com a passagem de grande quantidade de radiação. É também responsável pela interação com os minerais, o que dá origem às suas cores, e com os pigmentos da vegetação. Porém, o problema dessa faixa espectral é o alto espalhamento da radiação solar incidente pelos gases atmosféricos. Isso pode provocar uma redução do contraste da refletância dos alvos terrestres (MENESES e ALMEIDA, 2012).

Infravermelho próximo (0,76 - 1,2 μm): Segundo Meneses e Almeida (2012), nessa região do espectro solar a atmosfera é bastante transparente, mas ocorrem algumas bandas de

absorções, impedindo que todo o intervalo possa ser continuamente utilizado por sensoriamento remoto. Segundo os autores, esse intervalo espectral gera feições espectrais que são diagnósticas para identificar a natureza de vários tipos de rochas, principalmente as de composição mineral com metais de transição.

Infravermelho de ondas curtas (1,2 – 3,0 μm): É a região espectral geológica, pois é nesta faixa espectral que os vários minerais de alteração hidrotermal têm as suas diagnósticas feições de absorção. Nessa região os comprimentos de onda em 1,4 μm e em 1,9 μm são totalmente absorvidos pelas moléculas de vapor d'água da atmosfera, inviabilizando a utilização do sensoriamento remoto e, por consequência, de se determinar nos materiais terrestres a presença de água molecular nas suas estruturas (MENESES e ALMEIDA, 2012).

Infravermelho médio (3,0 - 5,0 μm) – Nessa região o Sol e a Terra não emitem quantidades suficientes de energia que possam ser detectadas pelos sensores. Apenas alvos com elevadas temperaturas, como vulcões e incêndios, podem ser detectados, pois agem como fontes próprias de emissão de radiação. É uma região espectral pouco usada no sensoriamento remoto, à exceção de sensores meteorológicos ou atmosféricos (MENESES e ALMEIDA, 2012).

Infravermelho termal (5,0 – 1,0 mm) – Segundo Meneses e Almeida (2012), essa região é conhecida como a região termal devido à radiação emitida pelos objetos terrestres em função das suas temperaturas de superfícies. Segundo os autores, a melhor janela atmosférica nessa região espectral para imageamento orbital é o intervalo de 8,0 μm a 14,0 μm , porque acima de 30 km, a camada de ozônio absorve toda a radiação além de 14 μm emitida pela Terra.

Micro-ondas (3,0 - 100 cm) – Essa região, segundo Meneses e Almeida (2012) é de uso de sensores ativos (radar), que utilizam fontes artificiais para a geração da radiação eletromagnética. Segundo os autores, por causa do tamanho do comprimento de onda, o radar tem a habilidade de operar em condições atmosféricas adversas, com coberturas de nuvens ou chuvas, e pode operar tanto de dia como à noite. É utilizado para realizar o mapeamento geológico, pois segundo os autores a interação das micro-ondas com as rochas é controlada pelas texturas de relevo.

3.2.1 Resolução

Atualmente, o sensoriamento é constituído por uma razoável constelação de satélites que oferecem imagens para atender as necessidades de uma ampla demanda de usuários. Uma forma de se abordar as potencialidades de um sensor é pelo dimensionamento de suas resoluções. A detecção ou identificação de um objeto nas imagens de sensoriamento remoto é

determinada por quatro diferentes formas de medições: resolução espacial, espectral, radiométrica e temporal (MENESES e ALMEIDA, 2012).

3.2.1.1 Resolução espacial

Segundo Meneses e Almeida (2012), resolução espacial determina o tamanho do menor objeto que pode ser identificado em uma imagem. Segundo os autores, um objeto somente pode ser detectado, quando o tamanho deste é, no mínimo, igual ou maior do que o tamanho do elemento de resolução no terreno. Ainda segundo Meneses e Almeida (2012), por exemplo, se uma casa tem 20 m x 20 m de tamanho, a resolução espacial da imagem deveria ser, no mínimo, de 20 metros para que essa casa possa ser identificada na imagem.

3.2.1.2 Resolução espectral

Uma outra propriedade de sensores imageadores é a resolução espectral. Segundo Meneses e Almeida (2012), o termo resolução espectral envolve pelo menos três parâmetros de medida:

- i) o número de bandas que o sensor possui;
- ii) a largura em comprimento de onda das bandas;
- iii) as posições que as bandas estão situadas no espectro eletromagnético.

De forma comparativa, Meneses e Almeida (2012) afirmam que, um sensor tem melhor resolução espectral se ele possui maior número de bandas situadas em diferentes regiões espectrais e com larguras estreitas de comprimentos de onda. Os autores acrescentam que essa necessidade é devida às diferenças relativas de refletância entre os materiais da superfície da terra, permitindo assim distinguir um material do outro, em determinados comprimentos de onda. Os autores também afirmam que a largura da banda deve ser dimensionada em concordância com as larguras das feições de absorção exibidas pelos espectros de refletância de cada material.

As feições de absorções são identificadores de composições dos tipos de rochas, solos, vegetação e água, e normalmente, são da ordem de 10 nm a 20 nm. Sensores que possuem bandas muito largas, além de 20 nm tendem a não ser capazes de diferenciar um objeto do outro em função de sua composição. Dessa forma, o objeto é detectado somente em razão da resolução espacial (MENESES e ALMEIDA, 2012).

Para ilustrar o efeito da resolução espectral na detecção ou identificação de um alvo, pode-se considerar o exemplo apresentado por Meneses e Almeida (2012). Para isso, considere

a Figura 3. Nela é apresentada uma área de floresta da região amazônica. As duas imagens estão na resolução espacial de 30 metros. Na Figura 3(a), que corresponde a uma banda do visível na faixa entre $0,63\mu\text{m}$ a $0,69\mu\text{m}$, os rios, apesar de possuírem dezenas de metros de largura, não são facilmente identificados, pois a refletância da água e da vegetação são, praticamente, iguais nesse comprimento de onda, não existindo contraste entre os dois tipos de alvos. Segundo Meneses e Almeida (2012), para a região amazônica, seria um erro selecionar esta banda espectral para se elaborar mapas de drenagem. Por outro lado, a Figura 3(b) mostra uma banda do infravermelho próximo. Segundo Meneses e Almeida (2012), nessa banda os rios são facilmente identificados, pois a água possui uma baixa refletância, enquanto a vegetação da floresta tem uma alta refletância. Nesse caso, ocorre uma razão de alto contraste entre os dois alvos.

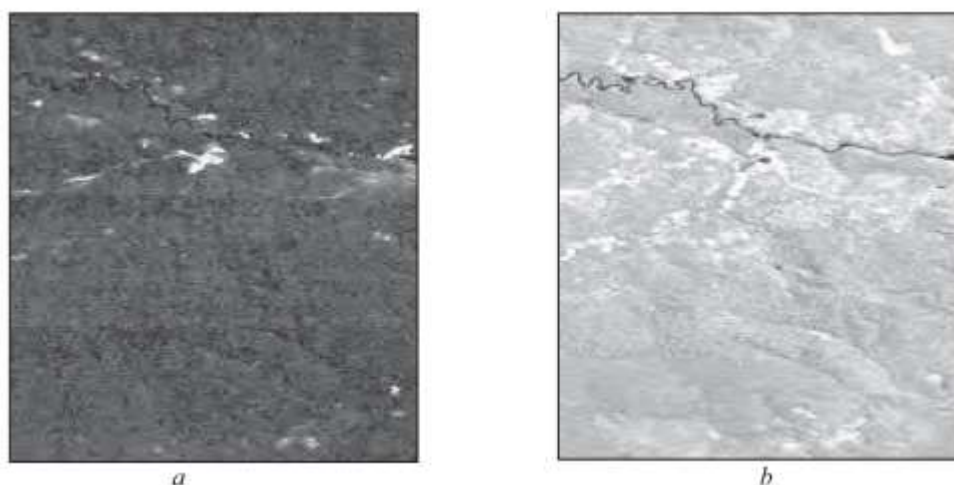


Figura 3 – Exemplo do efeito da resolução espectral no contraste do rio com a vegetação. Imagens da região amazônica nas bandas: (a) visível e (b) infravermelho próximo. (MENESES e ALMEIDA, 2012)

3.2.1.3 Resolução radiométrica

Segundo Novo (2010), a resolução radiométrica de um sensor descreve a sua habilidade de distinguir variações no nível de energia refletida, emitida ou retro-espalhada que deixa a superfície do alvo. Novo (2010) acrescenta que quanto maior for a capacidade do sensor de distinguir diferenças de intensidade do sinal, maior será sua resolução radiométrica.

A medida radiométrica também pode ser chamada de quantização. De maneira prática, a quantização é medida pelo intervalo de número de valores digitais usados para expressar os valores de radiância medidos pelo detector. Normalmente, é expressa em *bits*. Por exemplo,

uma imagem discretizada em 8 bits pode ser representada com 256 níveis de cinza. Conforme exemplificado na Figura 4, quanto maior a quantização, melhor será a qualidade visual da imagem. Na Figura 4 podemos observar que a imagem de 8 *bits* tem detalhes visuais melhores que os das imagens de menores resoluções radiométricas, de 6, 4 e 2 bits. A imagem de 2 *bits* possui apenas 4 níveis de cinza (MENESES e ALMEIDA, 2012). Em processamento digital de imagens, a medida radiométrica é também denominada de resolução de profundidade (GONZÁLEZ e WOODS, 2010).

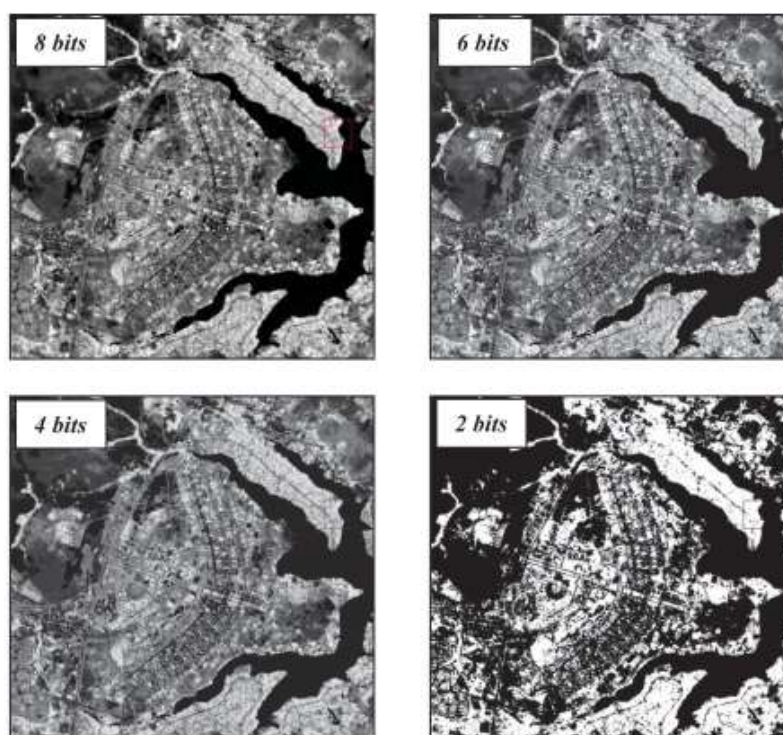


Figura 4 – Exemplos de imagens com diferentes níveis de quantização ou de resolução radiométrica. (MENESES e ALMEIDA, 2012)

3.2.1.4 Resolução temporal

O sensor possui ainda a resolução temporal. Essa resolução refere-se à frequência de imageamento sobre uma mesma área. Por exemplo, o sensor TM-Landsat possui uma resolução temporal de 16 dias, já o sensor instalado no satélite meteorológico Goes possui uma resolução temporal de meia hora (FLOREZANO, 2011).

3.3 LANDSAT

LANDSAT é um programa de satélite de observação da Terra de origem norte americana. As atividades iniciaram em 1972 com o lançamento do satélite ERTS-1. O programa teve sequência com o LANDSAT 2, 3, 4, 5, 6, 7 e atualmente encontra-se em funcionamento o

LANDSAT-8/OLI. O objetivo do programa é fornecer dados multiespectrais da superfície da Terra (USGS, 2016).

3.3.1 LANDSAT-8/OLI

As imagens geradas possuem um ciclo de repetição de 16 dias com resolução radiométrica de 16 *bits*. Os dados são coletados por meio de 11 bandas espectrais cuja resolução espacial é de 30m, com exceção da banda pancromática, que possui resolução de 15m e a banda infravermelho termal que possui resolução de 100 m. Na Tabela 6 são exibidas as bandas do LANDSAT-8/OLI com suas respectivas descrições, bem como, os dois tipos de sensores: Óptico, denominado *Operational Land Imager (OLI)* e o outro Termal, denominado de *Thermal InfraRed Sensor (TIRS)*.

Tabela 6 – Bandas LANDSAT-8/OLI USGS (2016)

Sensor		Banda	Resolução Espacial	Comprimento de Onda (μm)
OLI	1	Aerosol	30 m	0,435 - 0,451
	2	Azul	30 m	0,452 - 0,512
	3	Verde	30 m	0,533 - 0,590
	4	Vermelho	30 m	0,636 - 0,673
	5	Infravermelho próximo	30 m	0,851 - 0,879
	6	Infravermelho de ondas curtas 1	30 m	1,5166 - 1,651
	7	Infravermelho de ondas curtas 2	30 m	2,107 - 2,294
	8	Pancromática	15 m	0,503 - 0,676
	9	<i>Cirrus</i>	30 m	1,363 - 1,384
TIRS	10	Infravermelho termal 1	100 m	10,60 - 11,19
	11	Infravermelho termal 2	100 m	11,50 - 12,51

Cada imagem gerada pelo sistema LANDSAT tem um sistema de identificação das imagens composto de 2 números: o primeiro é o número da órbita e o segundo, o número da cena dentro daquela órbita, chamado de ponto. Assim, por exemplo, a imagem 231/62 (correspondente à região de Manaus-AM).

3.3.2 Índice de vegetação

Índices de vegetação são transformações espectrais de duas ou mais bandas. Foram desenvolvidos para explorar as propriedades espectrais de vegetação. Esse conceito está

baseado no fato de que a energia refletida na banda espectral vermelha e na banda infravermelha próxima encontra-se relacionada à atividade fotossintética da vegetação e à sua estrutura.

3.3.2.1 Índice de Vegetação de Diferença Normalizada (NDVI)

Segundo Huete *et al.*, (2002), o índice de Vegetação de Diferença Normalizada (IVDN ou NDVI), resulta da resposta espectral da vegetação nas faixas do vermelho e do infravermelho próximo, conforme mostrado na Eq. (1):

$$NDVI = \frac{(\rho_{nir} - \rho_{red})}{(\rho_{nir} + \rho_{red})} \quad (1)$$

Na Eq. (1), ρ_{nir} são valores de refletância para o infravermelho próximo e ρ_{red} são valores de refletância para o vermelho. O NDVI apresenta o resultado normalizado para a razão no intervalo de -1 a +1. Para alvos terrestres, os valores menores e iguais a zero referem-se a pixels sem vegetação e os valores próximos de 0,8 referem-se à vegetação mais densa. Embora extremamente útil na estimativa de parâmetros biofísicos da vegetação, o NDVI apresenta problemas de saturação em áreas com densa vegetação (FERREIRA *et al.*, 2008 apud GINCIENE e BITENCOURT, 2011).

3.3.2.2 Índice de Vegetação Melhorado (EVI)

O EVI foi proposto por foi proposto por Huete *et al.*, (2002) com o objetivo de contornar a limitação do NDVI. Esse índice de vegetação é uma variação do NDVI, que procura corrigir efeitos de influências atmosféricas e do solo, conforme descrito na Eq. (2).

$$EVI = G \frac{(\rho_{nir} - \rho_{red})}{(L + \rho_{nir} + C_1 \rho_{red} - C_2 \rho_{blue})} \quad (2)$$

Na Eq. (2), o parâmetro ρ_{nir} corresponde ao valor de reflectância para o infravermelho próximo. O parâmetro ρ_{red} corresponde ao valor de reflectância para o vermelho. O parâmetro ρ_{blue} corresponde ao valor de reflectância para o azul. Os valores para as constantes $G = 2,5$, $L = 1$, $C_1 = 6$ e $C_2 = 7,5$ foram obtidos empiricamente (HUETE *et al.*, 2002). Entretanto, segundo Huete *et al.*, (2002) a introdução do espectro do azul à equação original e a introdução das constantes G , L , C_1 e C_2 , tornaram o índice sensível à discriminação de variações estruturais na vegetação de áreas com densa vegetação.

3.4 PRODES

Segundo o Inpe (2017), desde 1988, o projeto PRODES realiza o monitoramento por satélite para detecção do desmatamento por corte raso na Amazônia, disponibilizando taxas anuais de desmatamento na região. As informações geradas são de grande importância, pois servem de base para o governo brasileiro estabelecer políticas públicas. Anualmente, disponibiliza-se as imagens utilizadas na identificação das áreas desmatadas, acompanhadas dos respectivos mapas que contêm os polígonos de desmatamento. O mapa disponibilizado delimita as regiões nas seguintes classes: desmatamento, floresta, não-floresta (cerrado), hidrografia e nuvem.

A primeira metodologia adotada pelo INPE, chamada de PRODES Analógico, consistia em realizar a detecção dos polígonos de desmatamento por meio de interpretação visual. Esse trabalho era realizado em 220 imagens de satélite LANDSAT 5/TM que eram impressas em papel fotográfico na escala 1:250.000. Após a interpretação, os polígonos eram digitalizados manualmente. Essa metodologia foi adotada de 1988 a 2002 (INPE, 2013).

Entre 2003 a 2005, o INPE utilizou o processo de interpretação através de classificação digital realizada por meio do Sistema de Processamento de Informações Georeferenciadas (SPRING). Essa fase é conhecida como PRODES Digital (INPE, 2013).

Após 2005, o monitoramento passou a ser realizado com a ajuda do software de informações geográficas TerraAmazon. Esse software permite que a interpretação de imagens seja realizada em um ambiente corporativo facilitando o fluxo de trabalho. Com isso a forma de classificar e interpretar as imagens passou a ser feita de maneira visual delimitando as regiões de desmatamento diretamente na tela do computador (INPE, 2013).

3.5 REDES NEURAIAS ARTIFICIAS

Segundo Mendes Filho (2000) apud Raia (2000), Redes Neurais Artificiais (RNA) são técnicas computacionais que apresentam um modelo matemático inspirado na estrutura neural de organismos inteligentes e que adquirem conhecimento por meio da experiência.

As principais vantagens das redes neurais artificiais são: tolerância à falhas; aplicações em tempo real; alta capacidade de auto-adaptação; capacidade de resolver problemas práticos sem a necessidade da definição de lista de regras ou de modelos precisos. As redes neurais artificiais caracterizam-se por ajustarem a sua estrutura através de aprendizado e encontram diversas aplicações: reconhecimentos de padrões; associação de padrões; filtragem de ruído;

aproximação de funções. Caracterizam-se por serem a principal ferramenta utilizada nas aplicações anteriormente citadas em áreas onde é difícil criar modelos precisos da realidade e onde são encontrados problemas com frequentes mudanças de ambiente de atuação (BRAGA; LUDEMIR; CARVALHO, 2000).

Segundo Tubb (1993) apud Almeida (2015), um neurônio artificial pode ser definido como uma unidade de processamento simples que recebe uma ou mais entradas que são multiplicadas por pesos que são aprendidos. A aplicação de uma função de ativação ao somatório desses produtos define a saída da rede. Cada entrada está conectada a unidade somadora através de um peso.

A Figura 5 ilustra um típico neurônio artificial com suas partes mais importantes: as sinapses, caracterizadas pelos valores de peso associados a cada entrada do neurônio, uma função somadora, e uma função de ativação. A função de ativação é utilizada para limitar a amplitude do sinal de saída do neurônio a algum valor finito (RUMELHART e MCCLELLAND, 1986).

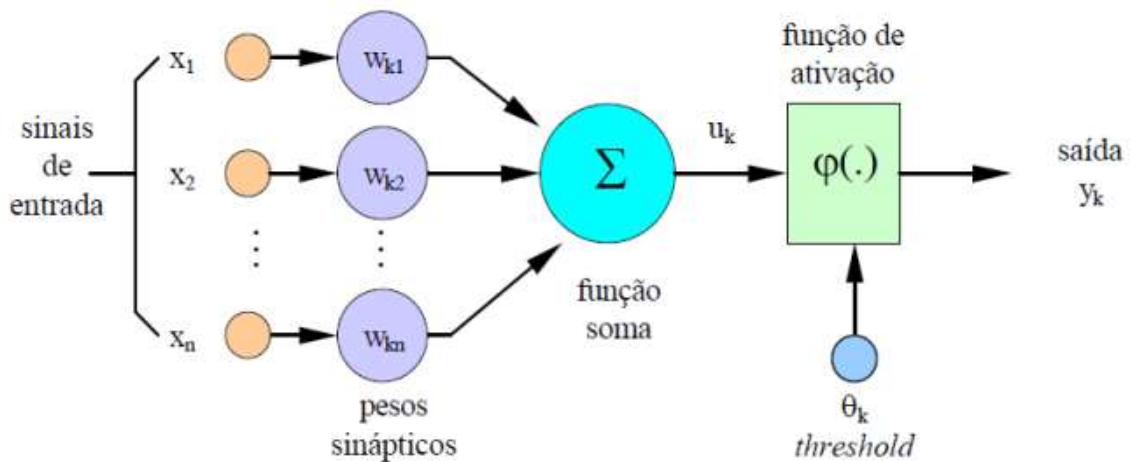


Figura 5 – Modelo não linear de um neurônio. Fonte: Adaptado de Haykin (2004)

Segundo Tubb (1993) apud Raia (2000), a forma como os neurônios são organizados é denominada de topologia da rede. A topologia da rede irá afetar o desempenho da mesma, bem como as aplicações para as quais ela é utilizada. Segundo o autor, sua estrutura está intimamente ligada ao algoritmo de aprendizado usado na fase de treinamento da rede. Algumas redes, denominadas redes recorrentes, permitem que a realimentação do sinal de saída para a entrada da rede. Por outro lado, outras redes, denominadas de redes competitivas, permitem que os neurônios da mesma camada estejam conectados. Existem ainda aquelas redes que permitem que os neurônios enviem sinais de volta para ele mesmo.

O algoritmo de aprendizado pode ser descrito como um conjunto de regras bem definidas para a solução de um problema de aprendizado. Existem vários tipos de algoritmos de aprendizado. Alguns são específicos para determinadas topologias de redes neurais e diferem entre si, principalmente, pela forma com que os pesos são atualizados (BRAGA *et al.*, 1998 apud RAIA JUNIOR, 2000). Os paradigmas de aprendizado, por sua vez, definem a maneira como a rede se relaciona com o ambiente e se dividem em dois grupos principais, segundo Brondino (1999) apud Raia Junior, 2000:

- **Supervisionado:** na fase de treinamento, é apresentado à rede um conjunto de entradas acompanhadas de suas respectivas saídas. O objetivo é minimizar o sinal de erro, que é uma função da diferença entre a saída desejada e aquela fornecida pela rede. Esta minimização é feita por meio do ajuste dos pesos da rede. Um exemplo deste paradigma de aprendizado é o algoritmo de propagação reversa (*BackPropagation – BP*), explicado mais a frente;
- **Não supervisionado:** a rede aprende automaticamente, sem uma mensagem de erro externa, ou seja, não há supervisão externa. Para tanto, é necessário que entradas similares sejam apresentadas à rede, com o objetivo que essa possa extrair as características estatisticamente relevantes e criar classes de maneira automática;
- **Híbrido:** este método consiste em uma combinação dos aprendizados não supervisionado e supervisionado. Nessa metodologia, a única informação que a rede recebe é a indicação de que a resposta fornecida está correta ou não.

3.1.2 Aprendizado Supervisionado: O Algoritmo de Propagação Reversa

Segundo Patrocínio (2004), o algoritmo de Propagação Reversa é baseado em um método estatístico proposto em 1951 por Hebert Robbins e Sutton Monro (ROBBINS, 1951), os quais o denominaram como “método de aproximação estocástica”. Várias aplicações dessa técnica têm obtido êxito em reconhecimento de padrões e sistemas de identificação. O Algoritmo BP tornou-se popular com os trabalhos de Rumelhart, Hilton e Williams em 1986, apresentados no livro PDP (*Parallel Distributed Processing*) (RUMELHART, 1986 apud PATROCINIO, 2004).

A regra de aprendizagem mais usada no algoritmo BP é a Regra Delta Generalizada ou *Backpropagation Rule*, ou seja, uma generalização da Regra Delta para redes com múltiplas camadas de neurônios (PATROCINIO, 2004).

Segundo Hertz (1991) apud Patrocínio (2004), o objetivo da regra de aprendizagem é aproximar as saídas desejada e obtida. Em seguida, é calculado o erro para camada de saída e o mesmo é propagado no sentido saída-entrada, dessa forma os pesos vão se ajustando como funções destes erros. Segundo o autor, os erros para uma unidade em particular devem ser ajustados na proporção direta ao erro nas unidades com as quais ela está conectada. Por isso, realizar a retropropagação desses erros através da rede permite que os pesos entre todas camadas sejam corretamente ajustados. Portanto, mudanças na entrada de uma camada qualquer da rede acarretam mudanças também nas saídas da mesma (HERTZ, 1991 apud PATROCINIO, 2004).

A seguir, para realizar a explicação do algoritmo BP tradicional, será utilizado as seguintes definições de variáveis (HERTZ, 1991 apud PATROCINIO, 2004):

w_{ij} = pesos das conexões da camada intermediária j para camada de saída i ;

w_{jk} = pesos das conexões da camada intermediária j para camada de entrada k ;

ξ_k^μ = entradas;

ζ_i^μ = saídas desejadas;

O_i = valores das saída da rede (camada i);

V_j = valores das saídas dos neurônios da camada intermediária j .

As entradas são introduzidas à rede com um valor particular, sendo referenciadas pelo símbolo μ . O índice i sempre se refere a uma unidade de saída, o índice j uma unidade intermediária (ou da camada escondida) e k a uma unidade de entrada. Usa-se N para o número de unidades de entrada e p para número de padrões de entradas ($\mu = 1, 2, \dots, p$) (PATROCINIO, 2004).

Segundo Patrocínio (2004), dado um padrão de entrada μ , as unidades j da camada escondida recebem uma entrada da rede:

$$h_j^\mu = \sum_k w_{jk} \xi_k^\mu \quad (3)$$

As saídas são processadas:

$$V_j^\mu = g(h_j^\mu) = g\left(\sum_k w_{jk} \xi_k^\mu\right) \quad (4)$$

onde V_j^μ é a saída dos neurônios da camada intermediária e g a função de ativação da rede.

As saídas da unidade i recebem:

$$h_i^\mu = \sum_j w_{ij} V_j^\mu = \sum_k w_{ij} g\left(\sum_k w_{jk} \xi_k^\mu\right) \quad (5)$$

onde W_{ij} são os pesos das conexões dos neurônios da camada intermediária para os neurônios da camada.

As saídas da rede são processadas (O_i), sendo então as chamadas saídas obtidas pela rede:

$$O_i^\mu = g(h_i^\mu) = g\left(\sum_j w_{ij} V_j^\mu\right) = g\left(\sum_j w_{ij} g\left(\sum_k w_{jk} \xi_k^\mu\right)\right) \quad (6)$$

A função custo ou erro dada por:

$$E[w] = \frac{1}{2} [\xi_i^\mu - O_i^\mu]^2 \quad (7)$$

torna-se:

$$E[w] = \frac{1}{2} [\xi_i^\mu - g(\sum_j W_{ij} g(\sum_k w_{jk} \xi_k^\mu))]^2 \quad (8)$$

A regra de atualização é local. Considerando-se:

η = taxa de aprendizagem;

δ = erro;

Para calcular as saídas das camadas escondidas tem-se:

$$\Delta W_{ij} = -\eta \left(\frac{\partial E}{\partial W_{ij}} \right) = \eta \sum_\mu [\xi_i^\mu - O_i^\mu] g'(h_i^\mu) V_j^\mu = \eta \sum_\mu \delta_i^\mu V_j^\mu \quad (9)$$

onde δ_i^μ é definido por:

$$\delta_i^\mu = g'(h_i^\mu) [\xi_i^\mu - O_i^\mu] \quad (10)$$

$$\Delta w_{jk} = -\eta \left(\frac{\partial E}{\partial W_{ij}} \right) = -\eta \sum_\mu \left(\frac{\partial E}{\partial V_j^\mu} \right) \left(\frac{\partial V_j^\mu}{\partial w_{jk}} \right)$$

$$\Delta w_{jk} = \eta \sum_\mu [\xi_i^\mu - O_i^\mu] g'(h_i^\mu) W_{ij} g'(h_j^\mu) \xi_k^\mu$$

$$\Delta w_{jk} = \eta \sum_\mu \delta_i^\mu W_{ij} g'(h_j^\mu) \xi_k^\mu$$

$$\Delta w_{jk} = \eta \sum_\mu \delta_j^\mu \xi_k^\mu$$

Dessa forma o erro de um padrão de entrada μ de uma camada j é dada pela Eq. (11).

$$\delta_j^\mu = g'(h_j^\mu) \sum_i W_{ij} \delta_i^\mu \quad (11)$$

Esse é o método de atualização dos pesos para redes de múltiplas camadas, conhecido como Gradiente Descendente. O método do gradiente descendente é uma técnica de otimização usada para minimizar a dimensão dos erros. Ela pode ser generalizada da seguinte forma (PATROCINIO, 2004):

$$\Delta W_{pq} = \eta \sum_{\text{padrões}} \delta_{\text{saídas}} x V_{\text{entradas}} \quad (12)$$

onde p e q são as conexões entres duas camadas e V entradas e/ou saídas processadas nas camadas intermediárias (ou escondidas). O método requer funções de ativação que variem suavemente, pois descontinuidades abruptas produzem resultados indesejáveis no espaço de estado da rede. Na Figura 6 está representada uma rede BP, onde as setas cheias indicam a fase de propagação das informações e as setas pontilhadas indicam a fase de retropropagação.

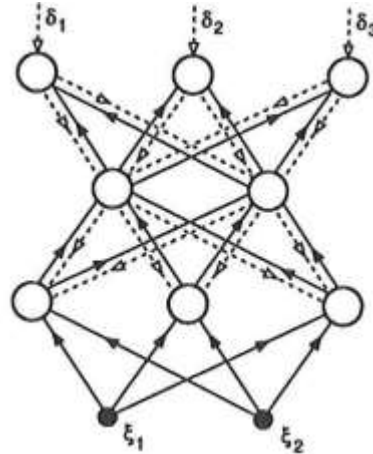


Figura 6 – Rede BP Fonte: PATROCINIO (2004)

Em geral, segundo Patrocínio (2004) o funcionamento de uma rede BP pode ser entendido através de duas fases distintas: fase de aprendizagem ou treinamento e fase de uso ou teste.

Os passos de mais alto nível do algoritmo BP são sintetizados da seguinte maneira (PATROCINIO, 2004):

- 1- Inicializar os pesos com pequenos valores aleatórios (entre 0 e 1);
- 2- Escolher um par de entrada ξ_k^μ e aplicar na camada de entrada ($m = 0$) tal que

$$V_k^0 = \xi_k^\mu \text{ para todos } k;$$

- 3- Propagar o sinal através da rede diretamente usando

$$V_i^m = g(h_i^m) = g\left(\sum_j w_{ij} V_j^{m-1}\right) \quad (13)$$

para cada i e m até que todas saídas V_i^m tenham sido calculadas;

- 4- Calcular os deltas para a camada de saída

$$\delta_j^m = g'(h_j^m) [\xi_j^\mu - V_j^m] \quad (14)$$

Para comparar a saída atual V_i^m com a saída desejada ζ_i^μ para o padrão μ inicialmente considerado.

5 - Calcular os deltas das camadas precedentes para retropropagar os erros

$$\delta_j^{m-1} = g'(h_i^{m-1}) \sum_j W_{ij}^m \delta_i^m \quad (15)$$

Para $m = M, M-1, \dots, 2$ até que o delta tenha sido calculado para todas as unidades;

6- Usar a Eq. (16) para atualizar as conexões de acordo com $w_{ij}^{novo} = w_{ij}^{antigo} + \Delta w_{ij}$

$$\Delta w_{ij} = \eta \delta_i^m V_j^{m-1} \quad (16)$$

7 – Voltar ao passo 2 e repetir para todos os padrões.

Juntamente, os passos 2 e 3 são conhecidos como passos de propagação para frente ou propagação direta (*feedforward*). Um vetor de entrada é aplicado e a saída resultante é computada. As computações são realizadas camada por camada. Os passos 4 e 5 são conhecidos conjuntamente como propagação reversa ou retropropagação. No passo 4, o vetor de saída produzido é subtraído do vetor de saídas desejadas do conjunto de treinamento, produzindo um erro. Esse erro é então utilizado pelo passo 5 para ajustar os pesos da rede (PATROCINIO, 2004)

Segundo Patrocínio (2004), a rede só é considerada treinada quando, após várias repetições desses 4 passos, o erro entre as saídas obtidas e as saídas desejadas estiver reduzido a um patamar aceitável, ou seja, um valor bastante baixo, que varia de acordo com cada aplicação. O autor acrescenta que o treinamento deve ser interrompido quando a rede apresentar uma boa capacidade de generalização e um erro global menor do que um limiar previamente estabelecido. Dessa forma, o ponto ótimo de parada deve ocorrer com erro mínimo e capacidade de generalização máxima. A fase de uso da rede BP corresponde aos passos 2 e 3 de seu processo de treinamento. Um vetor de entrada é aplicado a rede e vai progressivamente sendo propagado, através das ativações dos neurônios, desde a camada de entrada, passando pelas camadas escondidas, até atingir a camada de saída. Ao final do processo, a camada de saída manterá os valores correspondentes às entradas aplicadas.

Segundo Haykin (2004) apud Almeida (2015) o algoritmo BP apresenta alguns problemas em relação ao aprendizado. Por exemplo, em problemas muito complexos, o treinamento é lento e a rede pode cair num mínimo local. Outro problema existente é o *overfitting*, no qual o processo de classificação sofre uma piora devido a rede memorizar padrões que já foram apresentados. Nos tópicos seguintes são apresentadas soluções para este problema.

3.6 SOBRETREINAMENTO (*OVERFITTING*)

Segundo Mackay (1992) apud Almeida (2015), sobre-treinamento (*overfitting*) ocorre quando uma rede neural aprende “mais do que deveria” durante o período de treinamento. Como resultado, o treinamento excessivo da rede pode não resultar em um bom desempenho na fase de teste, devido ao fato da mesma, devido ao sobre-treinamento, não apresentar uma boa capacidade de generalização. Várias abordagens têm sido propostas na literatura para resolver esse problema. Dentre as mais utilizadas com redes neurais tem-se a parada antecipada (*early stopping*) e a regularização. A seguir são apresentadas essas duas abordagens.

3.6.1 Parada Antecipada

Segundo Sarle (1995), parada antecipada é uma técnica utilizada no processo de generalização de redes neurais. A ideia é separar o conjunto de dados em três subgrupos: treinamento, validação e teste. O conjunto de treinamento é utilizado na fase de aprendizado, para, através do algoritmo do gradiente descendente, ajustar os pesos da rede. O segundo conjunto, o de validação, é usado para decidir quando o treinamento será finalizado.

O terceiro conjunto, o de teste, é usado para avaliar o desempenho do modelo treinado. Este conjunto não é usado durante o treinamento da rede, mas apenas para avaliar o desempenho da mesma. Observa-se que, se o erro no conjunto de teste atinge um mínimo em um número de iteração significativamente diferente do que o do conjunto de validação, isto pode indicar uma divisão deficiente do conjunto de dados (DOAN e LIONG, 2004).

3.6.2 Regularização

Segundo Medeiros (2004) a regularização é um método que busca melhorar a capacidade de generalização dos algoritmos de aprendizado por meio de alguma restrição durante a etapa de treinamento. A teoria de regularização foi proposta por Tikhonov (TIKHONOV e ARSENIN, 1977 apud MEDEIROS, 2004), para resolver problemas mal-condicionados (mal postos). Segundo os autores, a ideia básica da regularização é estabilizar a solução por meio de algum funcional não-negativo que incorpore informação prévia sobre a solução.

A partir do trabalho de Tikhonov, vários tipos de regularização já foram propostos na literatura. Neste trabalho será abordado a Regularização Bayesiana, proposta por Mackay

(1992). Segundo Doan e Liong (2004), esse tipo de regularização faz com que o desempenho da rede tenha pesos menores, e isso força uma resposta mais suave da rede, tornando-se menos propensa ao sobre-treinamento.

A regularização acrescenta um termo à função objetivo representada pelo erro médio quadrático. Dessa forma, a função objetivo passa a ter o formato mostrado na Eq. (17).

$$F = \beta E_{QM} + \alpha E_W \quad (17)$$

Na Eq. (17) o termo E_{QM} refere-se ao erro médio quadrático, enquanto que o segundo termo E_W é o somatório dos pesos da rede elevados ao quadrado. Os parâmetros α e β regulam a importância dos dois termos anteriormente citados. O treinamento com esse tipo de regularização dispensa o uso de um conjunto de validação para determinação do valor ótimo dos parâmetros de regularização.

A magnitude relativa dos parâmetros da função objetivo traduz a ênfase que se quer dar ao treinamento. Se α for muito menor do que β , então o treinamento terá maior ênfase na minimização do erro médio quadrático (E_D). Se, por outro lado, α for muito maior que β , o treinamento irá enfatizar a redução dos valores dos pesos da rede, produzindo, então uma resposta da rede mais suave.

3.7 SELEÇÃO DE CARACTERÍSTICAS

As características utilizadas como entrada de um classificador têm grande influência sobre o desempenho do mesmo. Muitas vezes, um pequeno número destas características adiciona pouco ou nada à interpretação dos resultados. A decisão sobre quais características são importantes é baseada, geralmente, em critérios subjetivos (NETO e MOITA, 1998 apud ALMEIDA, 2015).

A redução do número de características de entrada de um classificador deve ser realizada através de técnicas objetivas. Algumas dessas técnicas utilizadas são: seleção escalar de características, análise de componentes principais (PCA) e análise discriminante linear (LDA). Na sequência abordaremos cada uma dessas técnicas, pois as mesmas foram utilizadas nessa dissertação para seleção de características.

3.7.1 Medida de separabilidade de classes

A correlação que existe entre as várias características influencia na capacidade de classificação de um sistema de reconhecimento de padrões. A medida da eficiência da

discriminação de um vetor de características pode ser usada como parâmetro para combinar características apropriadas e obter o melhor conjunto que separa as classes de objetos (THEODORIDIS e KOUTROUMBAS, 2009 apud XAVIER, 2012).

3.7.1.1 Fisher's Discriminant Ratio (FDR)

A medida FDR é utilizada para quantificar a capacidade de separabilidade de características individuais, considerando as características como estatisticamente independentes (THEODORIDIS e KOUTROUMBAS, 2009). Na Eq. (18) mostra-se a equação que expressa a medida FDR

$$FDR = \frac{(\mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2} \quad (18)$$

Segundo Theodoridis e Koutroumbas (2009), para um problema multiclasse, a medida FDR pode ser estendido conforme a Eq.

(19).

$$FDR = \sum_i^M \sum_{j \neq i}^M \frac{(\mu_i - \mu_j)^2}{\sigma_i^2 + \sigma_j^2} \quad (19)$$

Em que:

- μ_i é a média das amostras da classe i
- μ_j a média das amostras da classe j
- σ_i^2 é a variância das amostras da classe i
- σ_j^2 é a variância das amostras da classe j
- M é o número de classes

3.7.2 Seleção escalar de características

Seleção escalar de características trata do problema de selecionar um subconjunto de características a partir de um vetor de características, tratando cada característica como um

escalar (THEODORIDIS e KOUTROUMBAS, 2009). Para isso, é necessário escolher uma medida de separabilidade de classe, por exemplo a medida FDR. Após feita essa escolha, um critério $C(k)$, baseado nessa medida, é calculado para cada característica $k = 1, 2, \dots$

Segundo Theodoridis e Koutroumbas (2009) o procedimento de seleção de características possui os seguintes passos:

1. Selecione um critério $C(k)$ de separabilidade de classe e compute todos os seus valores para todas as características disponíveis x_k para $k = 1, 2, \dots, l$. Ordene de forma decrescente e escolha o melhor valor de $C(k)$. Vamos chamar de x_{i1} ;
2. Selecione uma segunda característica, calculando o coeficiente de correlação cruzada entre a característica escolhida x_{i1} e cada um das características restantes. O coeficiente de correlação cruzada é dado por:

$$p_{ij} = \frac{\sum_{n=1}^N x_{ni} x_{nj}}{\sqrt{\sum_{n=1}^N x_{ni}^2 \sum_{n=1}^N x_{nj}^2}} \quad (20)$$

3. Escolha a característica x_{i2} através da Eq. (21). Onde α_1 e α_2 são fatores de peso para determinar a importância relativa para cada termo.

$$i_2 = \arg \max_{x_j} \{ \alpha_1 C(j) - \alpha_2 \rho_{i_1 j} \} \quad (21)$$

4. Escolha a característica x_{ik} , $k = 3, \dots, l$, de modo que:

$$i_2 = \arg \max_j \left\{ \alpha_1 C(j) - \frac{\alpha_2}{k-1} \sum_{r=1}^{k-1} |\rho_{i_r j}| \right\} \quad (22)$$

3.7.3 ANÁLISE DE COMPONENTES PRINCIPAIS (PCA)

A análise de componentes principais é uma técnica de transformação de variáveis. Cada variável medida pode ser considerada como um eixo de variabilidade, estando correlacionada com outras variáveis. PCA transforma os dados a fim de descrever a mesma variabilidade total

existente, com o mesmo número de eixos originais, porém com as características não mais correlacionados entre si.

A análise de componentes principais considera a matriz de covariância dos padrões a fim de obter covariância nula entre as novas características.

Considere um padrão de treinamento representado pelo vetor \mathbf{a} , de dimensão $m \times n$. A matriz de covariância S é calculada a partir dos N vetores de treinamento $A = [a_1, a_2, \dots, a_N]$, e é dada pela Eq. (23).

$$S = \frac{1}{N}(A - M)(A - M)^T \quad (23)$$

Em que M é a matriz cujas colunas contêm as médias dos valores das colunas de A , ou seja, sua i -ésima coluna é dada pela Eq. (24):

$$M_i = \frac{1}{N} \sum_{j=1}^N a_j \quad (24)$$

A matriz de covariância é expressa pela Eq.

(25):

$$S = \begin{bmatrix} c_{11} & \cdots & c_{1m} \\ \vdots & \ddots & \vdots \\ c_{m1} & \cdots & c_{mm} \end{bmatrix} \quad (25)$$

Em que c_{ij} corresponde a covariância entre a variável i e a variável j .

Na diagonal principal essa matriz possui as variâncias de cada característica, enquanto que, nas demais posições, a covariância cruzada entre as características. Essa matriz é simétrica e real, de modo que é possível encontrar um conjunto de autovetores ortonormais.

O próximo passo é encontrar a matriz de transformação X formada pelos q autovetores $[u_1, u_2, \dots, u_q]$ associados aos q maiores autovalores da matriz de covariância S . Assim, para a redução de dimensão, a matriz de transformação é da forma dada pela Eq. (26).

$$X = [u_1, u_2, \dots, u_q] \quad (26)$$

Dado um conjunto de treinamento A , o algoritmo PCA encontra um transformador linear X de forma a representar A em outro espaço, Y , dado pela Eq. (27) a qual possui matriz de covariância diagonal.

$$Y = X^T A \quad (27)$$

3.7.4 ANÁLISE DISCRIMINANTE LINEAR (LDA)

A Análise Discriminante Linear, ou Análise de Fisher tem sido empregada para redução da dimensionalidade de dados pertencentes a várias classes, procurando escolher uma direção ou um conjunto de direções no espaço que possibilitem otimizar um critério de separabilidade das classes.

LDA consiste em uma técnica empregada para redução de dimensionalidade de dados pertencente a várias classes. O seu princípio é procurar uma direção ou um conjunto de direções no espaço que possibilitem otimizar um critério de separabilidade das classes, ou seja, a projeção com LDA aumenta a distância entre as entre classes e reduz a distância interclasse.

Para uma amostra de treinamento na forma de vetores, a matriz de dispersão interclasse S_w , é dada pela Eq. (28).

$$S_w = \sum_{i=1}^c \sum_{j=1}^{N_i} (a_j^i - \bar{a}_i)(a_j^i - \bar{a}_i)^T \quad (28)$$

Em que: a_j^i é o i -ésimo vetor de treinamento (elemento) da classe i , \bar{a}_i é a média dos elementos da classe i , c é o número total de classes e N_i é o número de elementos da classe i . A matriz de dispersão entre as classes S_b , é dada pela Eq. (29). Em que \bar{a} é média global das amostras de treinamento.

$$S_b = \sum_{i=1}^c N_i (\bar{a}_i - \bar{a})(\bar{a}_i - \bar{a})^T \quad (29)$$

Na técnica LDA, procura-se maximizar o critério de separação de classes dado pela Eq. (30). Pode-se mostrar que a matriz de projeção X que maximiza o esse critério é aquela formada pelos autovetores que correspondem aos maiores autovalores da matriz $S_w^{-1}S_b$. A dimensão da matriz X depende do número de dimensões que se deseja no vetor projetado. Um vetor projetado a_p é obtido a partir da Eq. (31).

$$j(X) = \frac{|X^T S_b X|}{|X^T S_w X|} \quad (30)$$

$$a_p = X^T a \quad (31)$$

Em que a matriz de projeção $X = [u_1, u_2, \dots, u_q]$ é composta dos q autovetores associados aos q maiores autovalores $\lambda_1, \lambda_2 \dots \lambda_q$ não nulos da matriz $S_w^{-1} S_b$ e a é imagem original.

CONSIDERAÇÕES FINAIS

Este capítulo apresentou conceitos essenciais para uma melhor compreensão desta dissertação. Os vários assuntos abordados, imagens ópticas, redes neurais, critérios para extração de características, são essenciais para o entendimento deste trabalho de dissertação de mestrado. O próximo capítulo explicará detalhadamente os métodos utilizados na dissertação, que fazem uso dos conceitos ora apresentados.

4 MATERIAIS E MÉTODOS

O desenvolvimento deste trabalho teve como objetivo o reconhecimento de desmatamento em imagens ópticas LANDSAT-8/OLI. Neste Capítulo são apresentados os materiais utilizados e a metodologia empregada.

4.1 Materiais

4.1.1 Ambiente de desenvolvimento

Neste trabalho foram utilizados os seguintes *softwares*: MATLAB ® 2015, QGIS 2.16 e ENVI 5.3.

O *software* ENVI foi utilizado para realizar o mosaico das imagens utilizadas. O *software* QGIS foi empregado para realizar o recorte das amostras de dados utilizados no classificador. O *software* MATLAB foi empregado no desenvolvimento do classificador da rede neural.

A infraestrutura de *hardware* utilizada para o desenvolvimento da dissertação foi um PC com características descritas na Tabela 7.

Tabela 7 – Características do computador utilizado

Sistema Operacional	<i>Windows 7 Professional</i>
Processador	Intel ® Core™ i7-4770 3,40 GHz
Memória RAM	16,0 GB
Tipo de Sistema	Sistema Operacional de 64 Bits

4.1.2 Imagens utilizadas

As imagens utilizadas foram imagens LANDSAT-8/OLI da região do sul do Amazonas, dos municípios de Apuí, Humaitá, Lábrea. As imagens foram baixadas do site: <http://earthexplorer.usgs.gov/>. As bandas utilizadas foram: Azul (B2), Verde (B3), Vermelho (B4), Infravermelho próximo (B5), Infravermelho de ondas curtas 1 (B6), Infravermelho de ondas curtas 2 (B7). A Figura 7 ilustra o estado do Amazonas, destacando os municípios analisados.

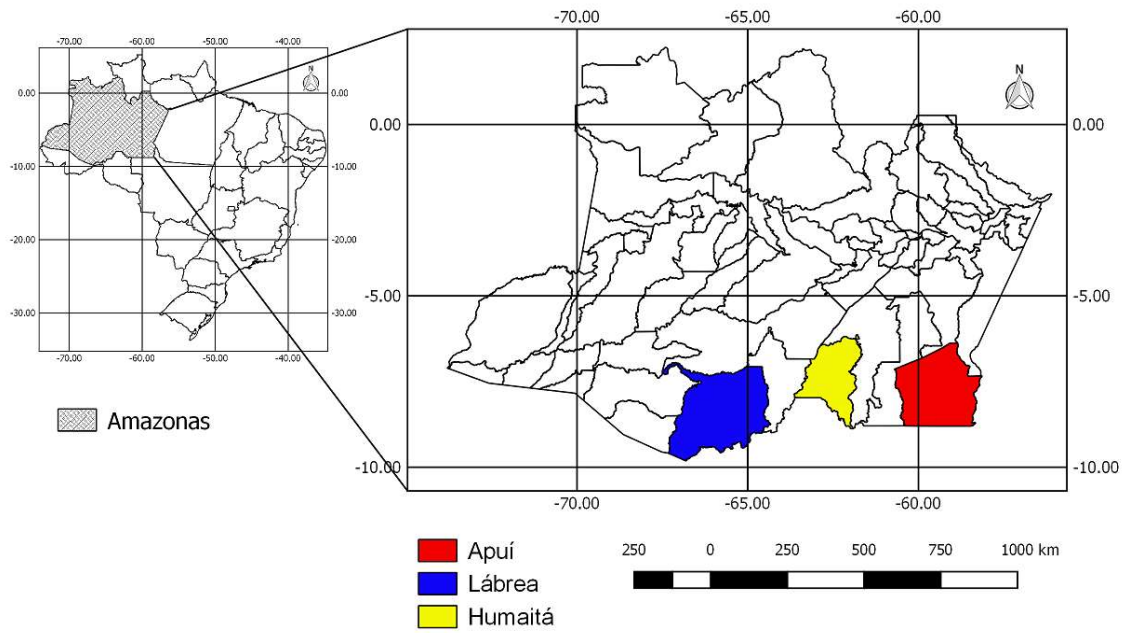


Figura 7 – Estado do Amazonas com as regiões analisadas

A Tabela 8 mostra uma lista das imagens LANDSAT-8/OLI utilizadas e as Figura 8, 9 e 10 mostram, em um mapa do estado do Amazonas, as áreas listadas na Tabela 8.

Tabela 8 – Lista de Imagens utilizadas

Região	Órbita/Ponto	Data da Imagem
Apuí	229/65	09/08/2014
	229/66	27/07/2015
	230/64	16/08/2014
	230/65	03/08/2015
	230/66	03/08/2015
Humaitá	231/64	08/09/2014
	231/65	25/07/2015
	231/66	25/07/2015
	232/64	30/08/2014
	232/65	01/08/2015
Lábrea	232/66	01/08/2015
	001/65	30/07/2015
	001/66	30/07/2015
	001/67	30/07/2015
	002/65	22/08/2015
	233/65	24/08/2015
	233/66	09/09/2015
233/67	09/09/2015	

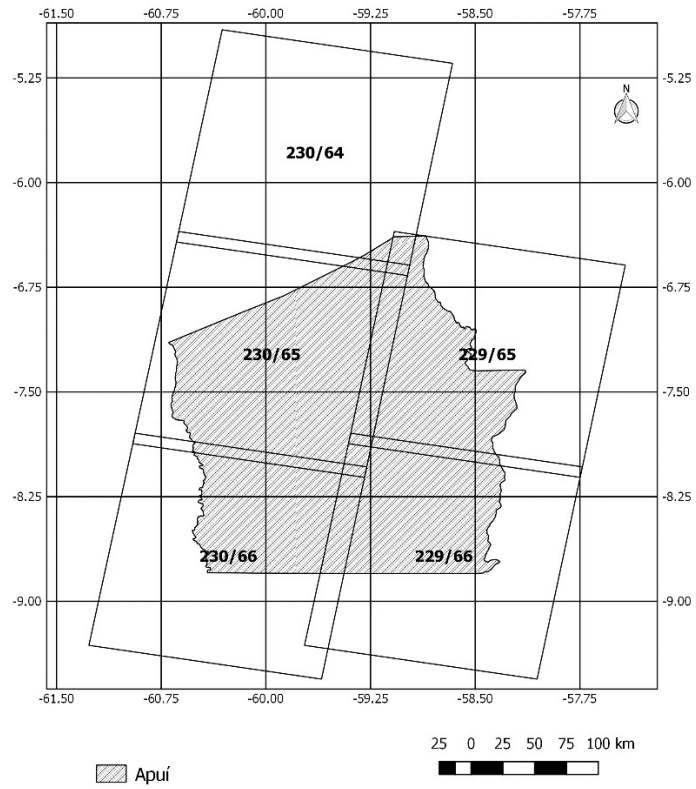


Figura 8 – Identificação das imagens LANDSAT-8/OLI para a Região de Apuí

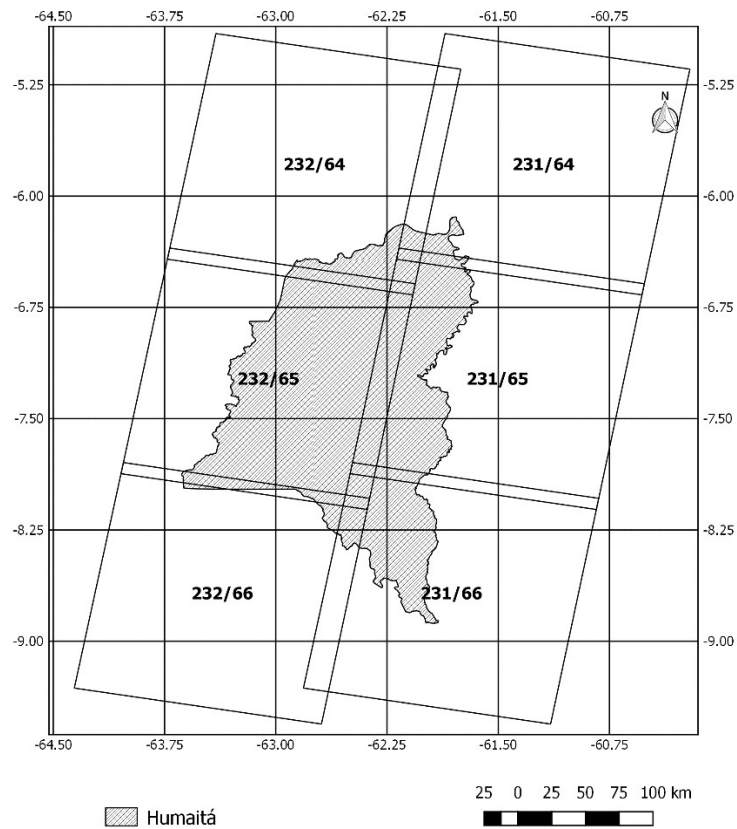


Figura 9 – Identificação das imagens LANDSAT-8/OLI para a Região de Humaitá

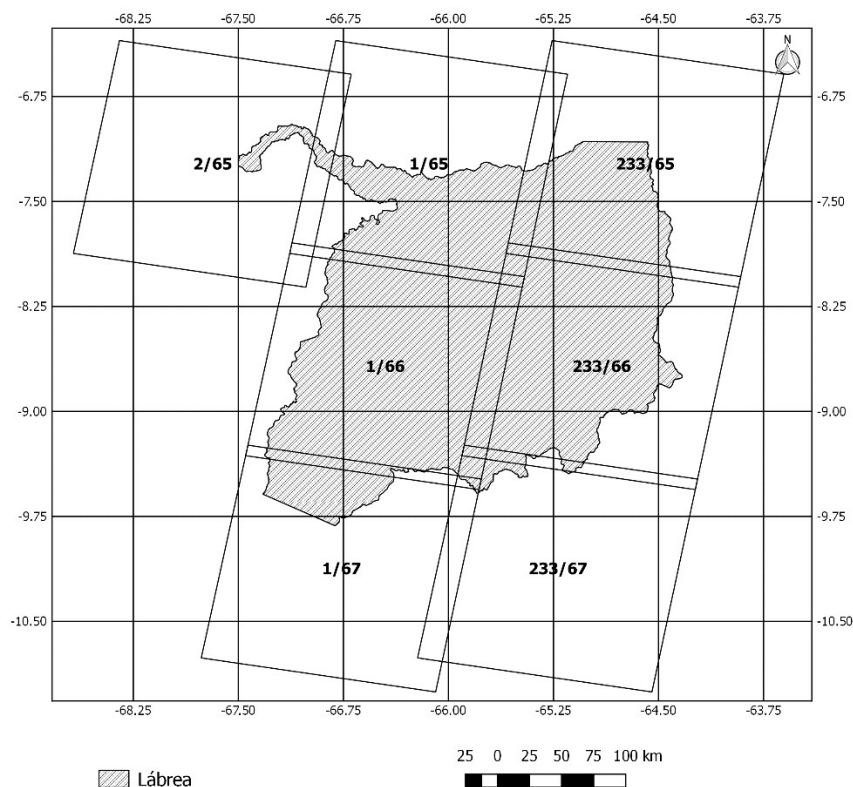


Figura 10 – Identificação das imagens LANDSAT-8/OLI para a Região de Lábrea

No Apêndice 1 são mostradas as imagens LANDSAT-8/OLI, das bandas B3-B7, correspondentes a região de Apuí, Humaitá e Lábrea utilizadas neste trabalho.

4.1.3 Dados vetoriais utilizados

Os dados vetoriais utilizados foram os dados do PRODES para os municípios de Apuí, Humaitá e Lábrea do ano de 2015. Os dados foram adquiridos do site do INPE no endereço: <http://www.dpi.inpe.br/prodesdigital/prodes.php>.

4.2 Métodos

A Figura 11 ilustra um diagrama em blocos das etapas de um método de reconhecimento de padrões, quais sejam: Aquisição de dados, Extração de características e Classificação. As seções a seguir explicam como foram realizadas cada uma dessas etapas.



Figura 11 – Diagrama em blocos com as etapas de reconhecimento de padrões

4.2.1 Aquisição de dados

Essa fase consistiu em obter as imagens para utilização no classificador. Conforme explicado na seção 4.1.2, foram adquiridas imagens da região do sul do Amazonas, dos municípios de Apuí, Humaitá, Lábrea.

Para a geração dos dados de entrada do classificador, foram utilizados os resultados do projeto PRODES, descrito na seção 3.4, do ano de 2015. Foram extraídas informações sobre áreas de floresta, áreas desmatadas e áreas correspondentes a outros tipos de vegetação, que também são denominadas pelo INPE de áreas de não floresta. No nosso problema essas áreas foram classificadas em duas classes. A primeira classe corresponde a áreas de floresta e áreas de não floresta, enquanto que a segunda classe corresponde a áreas desmatadas. Assim sendo, foi definido nesse trabalho um problema de reconhecimento de padrões binário. Ao se apresentar, na entrada do classificador, um vetor de características corresponde a pixels extraídos de regiões de floresta ou de não floresta, a saída do mesmo deve produzir um valor 1. Por outro lado, ao se apresentar, na entrada do classificador, um vetor de características corresponde a pixels extraídos de regiões de desmatamento, a saída do mesmo deve produzir um valor 0.

A Tabela 9 mostra as imagens utilizadas, com as respectivas quantidades de amostras retiradas para compor o conjunto de dados que foi utilizado para o treinamento e teste do classificador.

Tabela 9 – Lista de Regiões

Região	Órbita/Ponto	Data da Imagem	Quant. Amostras Desmatamento	Quant. Amostras Floresta	Quant. Amostras Não-Floresta
Apuí	229/65	09/08/2014	1000	500	500
	229/66	27/07/2015			
	230/64	16/08/2014			
	230/65	03/08/2015			
	230/66	03/08/2015			
Humaitá	231/64	08/09/2014	1000	500	500
	231/65	25/07/2015			
	231/66	25/07/2015			
	232/64	30/08/2014			
	232/65	01/08/2015			
	232/66	01/08/2015			
Lábrea	001/65	30/07/2015	1000	500	500
	001/66	30/07/2015			
	001/67	30/07/2015			
	002/65	22/08/2015			
	233/65	24/08/2015			
	233/66	09/09/2015			
	233/67	09/09/2015			

As Figura 12, 13 e 14 mostram, respectivamente, as regiões de Apuí, Humaitá e Lábrea com as amostras de cada classe utilizadas no classificador.

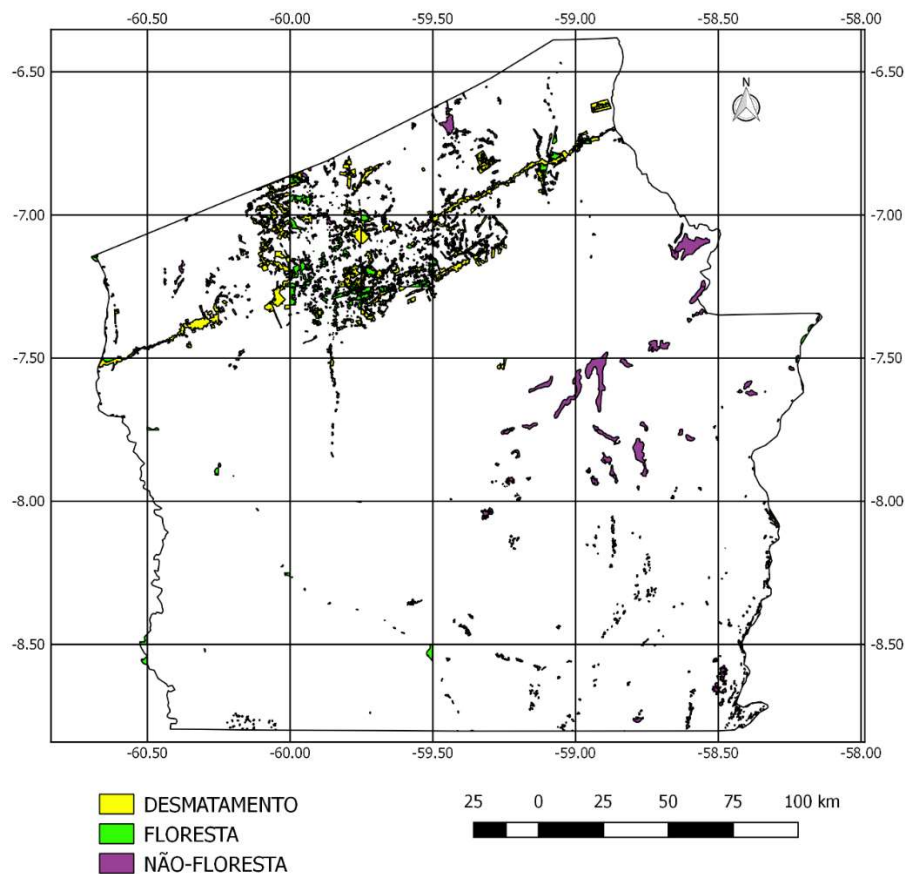


Figura 12 – Região de Apuí

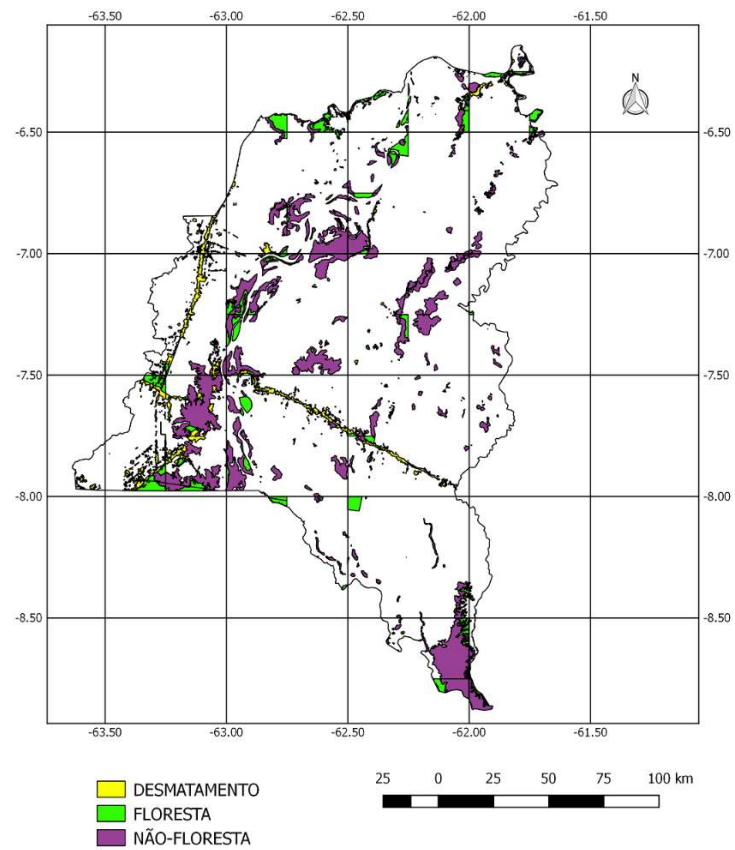


Figura 13 – Região de Humaitá

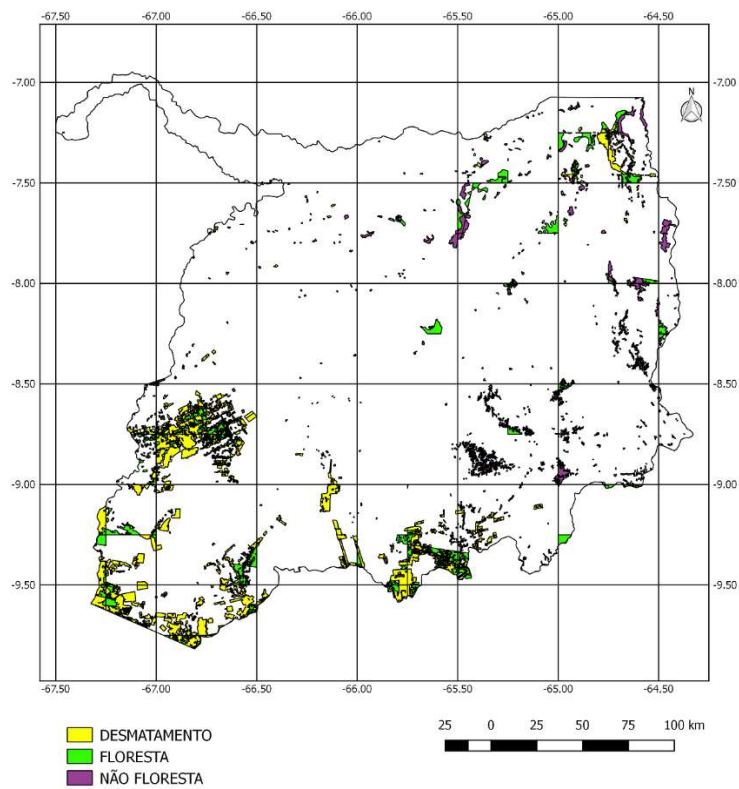


Figura 14 – Região de Lábrea

Para efeitos de comparação, a metodologia proposta nesta dissertação foi também aplicada para detecção de desmatamento às imagens utilizadas nos trabalhos desenvolvidos por Reis (2014), Pimentel (2014) e Zanotta (2014). Estes trabalhos utilizaram imagens ópticas para detecção de regiões desmatadas na região da Amazônia. A Tabela 10 apresenta as imagens utilizadas nos trabalhos citados.

Tabela 10 – Imagens LANDSAT-5/TM utilizadas

Autor	Orbita/Ponto	Data
Reis (2014)	227/062	29/06/2010
Pimentel (2014)	224/64	27/06/2011
Zanotta (2014)	225/64	20/07/2011

A Figura 15 demarca as regiões utilizadas por Reis (2014) e Zanotta (2014) no estado do Pará. As Figuras 16 e 17 mostram ampliações das regiões utilizada por Reis (2014) e Zanotta (2014), respectivamente, destacando as amostras de áreas desmatadas e de floresta utilizadas no classificador.

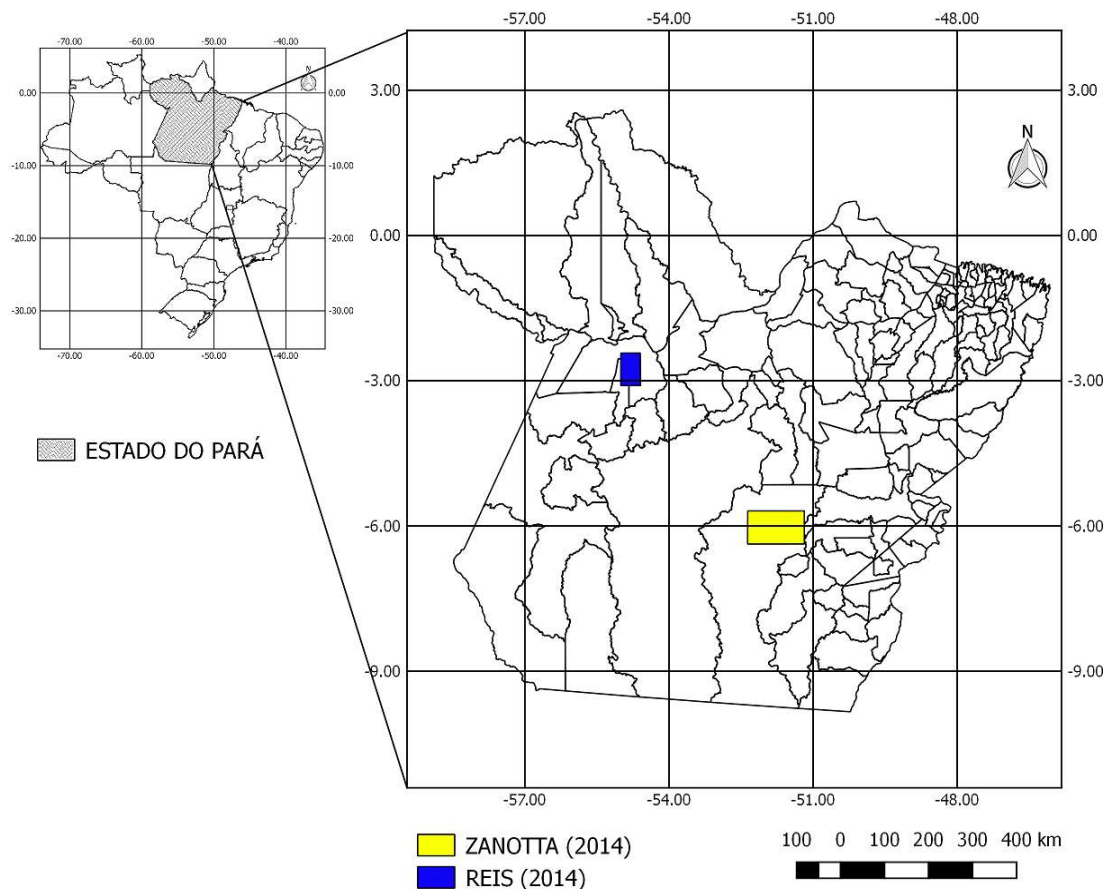


Figura 15 – Regiões utilizadas por Reis (2014) e Zanotta (2014) no estado do Pará.

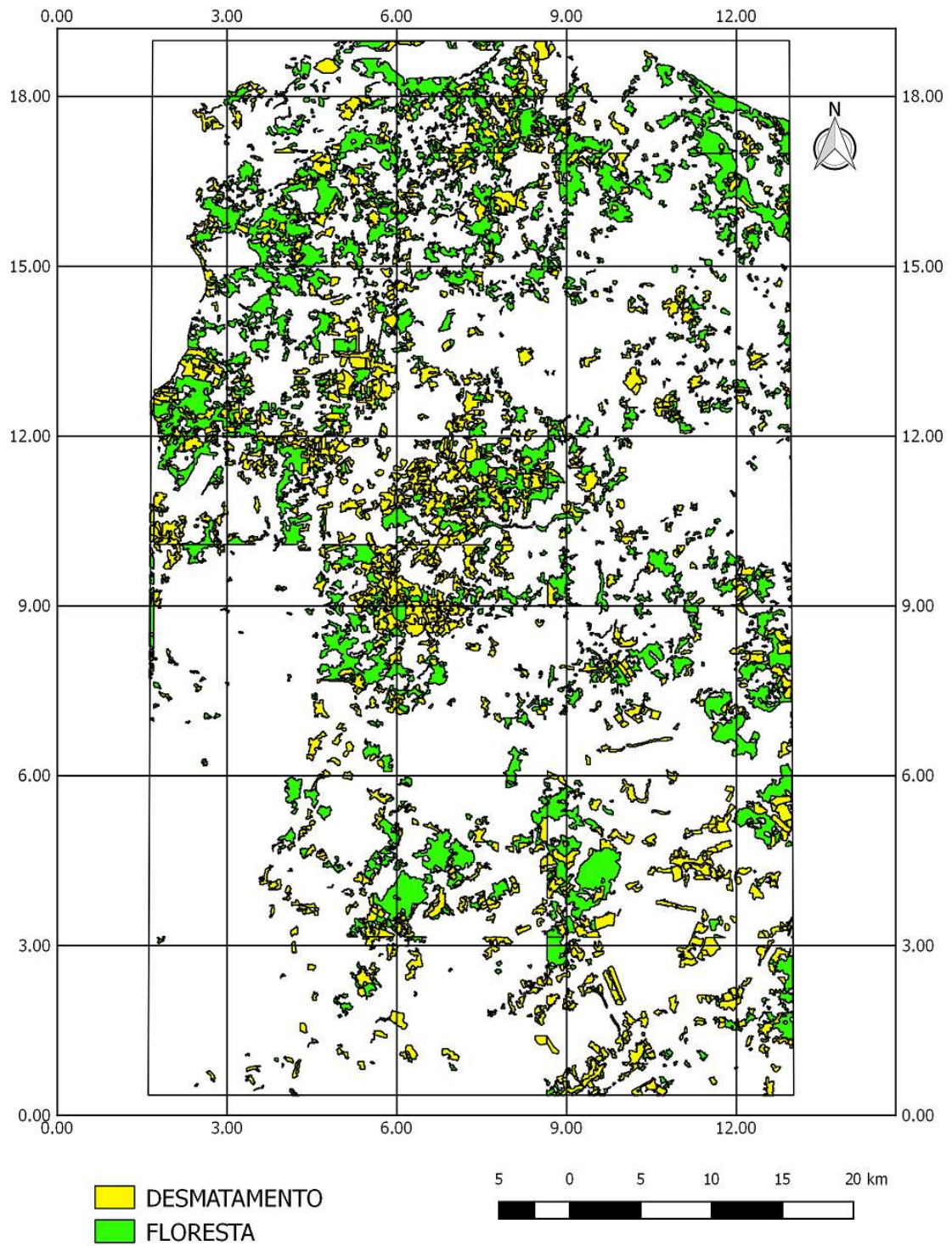


Figura 16 – Ampliação da região utilizada por Reis (2014)

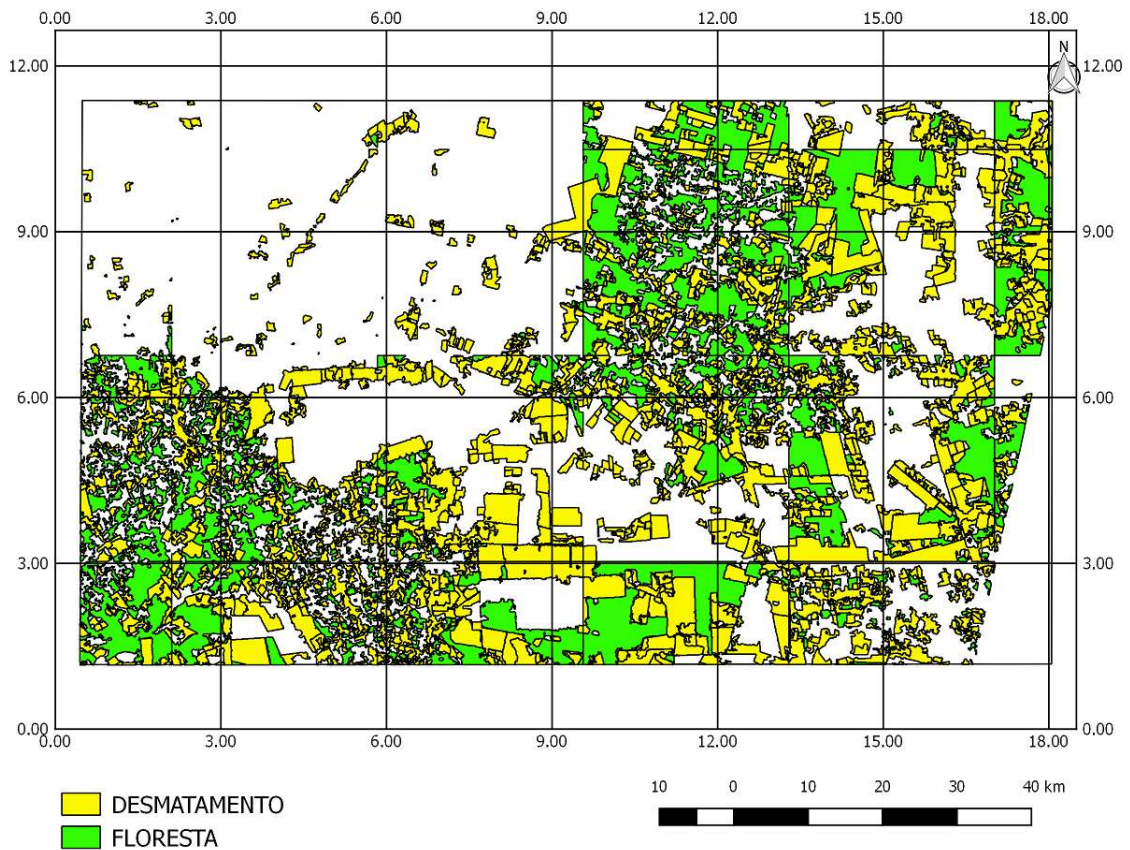


Figura 17 – Ampliação da região utilizada por Zanotta (2014)

A Figura 18, demarca a região analisada por Pimentel (2014), localizada no estado do Mato Grosso. A Figura 19 mostra uma ampliação dessa região, destacando as amostras das regiões desmatadas e de floresta utilizadas no classificador.

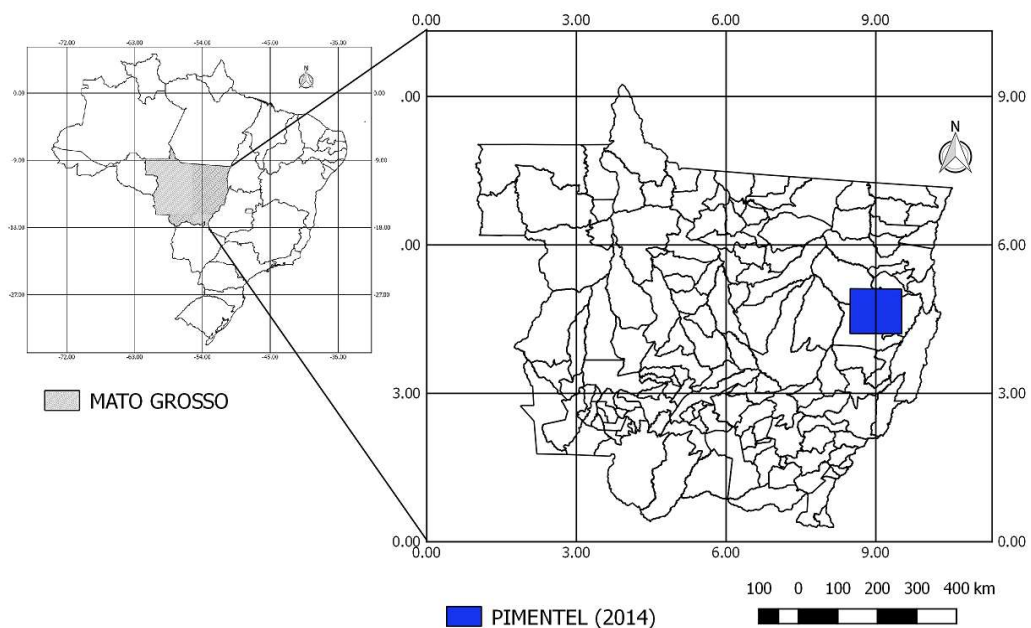


Figura 18 – Estado de Mato Grosso (Região utilizada por Pimentel, 2014)

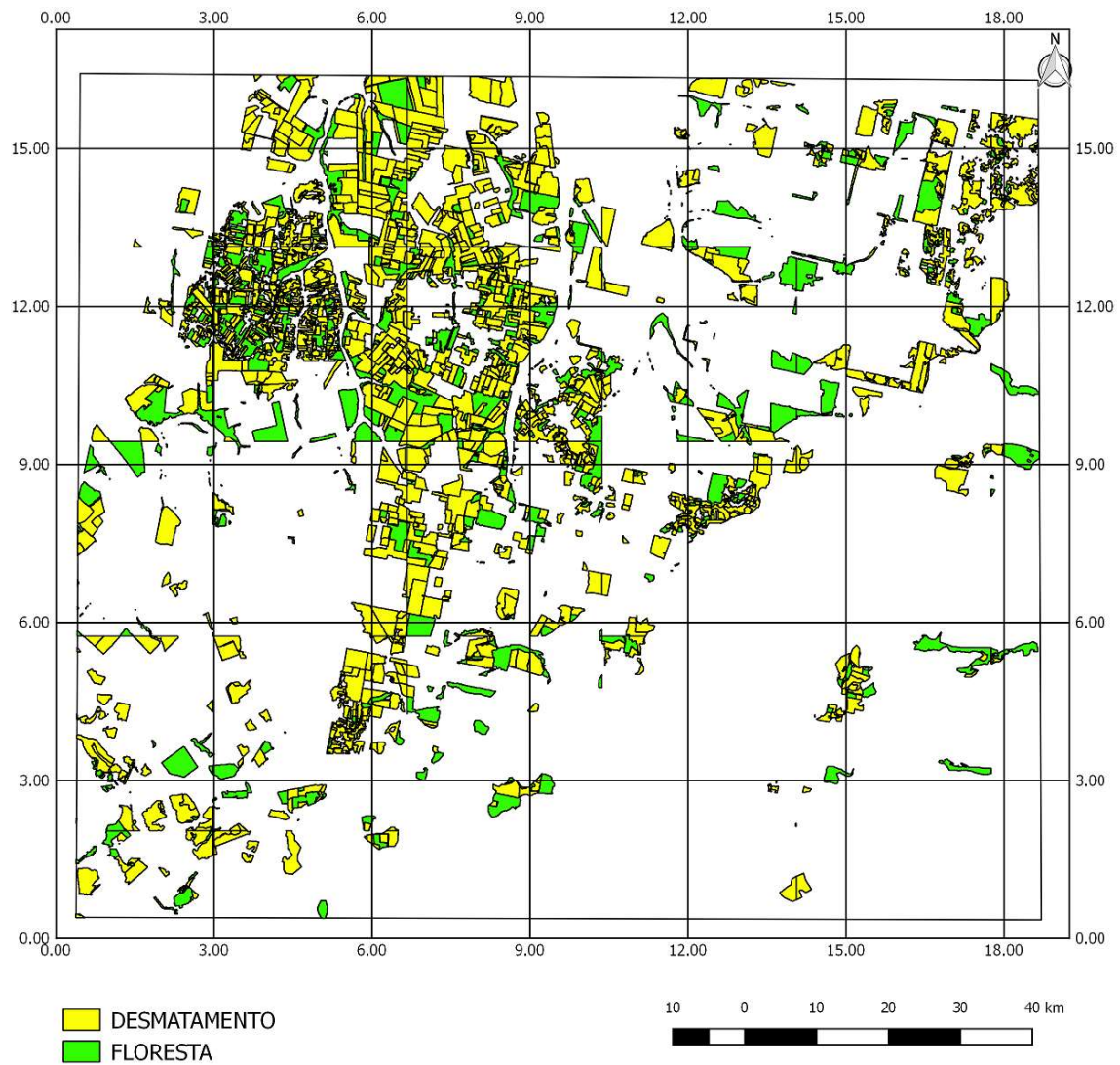


Figura 19 – Região utilizada por Pimentel (2014)

4.2.1 Extração de Características

O conjunto de características utilizados neste trabalho compreende características de textura. Com base na revisão da literatura, as características de textura mais utilizadas e que foram usadas neste trabalho foram: média (μ) → Eq. (32), variância (σ^2) → Eq. (33), desvio padrão (σ) → Eq. (34), homogeneidade (g) → Eq. (35), Energia (e) → Eq. (36), Entropia – (h) → Eq. (37), Correlação – (c) → Eq. (38) e contraste (co) → Eq. (39).

$$\mu = \frac{\sum_{ij}^n x(i,j)}{n} \quad (32)$$

Onde:

(i, j) – coordenadas do pixel da amostra;

n – número de pixels da amostra;

$x(i, j)$ – intensidade do pixels localizado nas coordenadas (i, j) .

$$\sigma^2 = \frac{\sum_{ij}^n ((x(i,j) - \mu)^2)}{(n-1)} \quad (33)$$

$$\sigma = \sqrt{\frac{\sum_{ij}^n ((x(i,j) - \mu)^2)}{(n-1)}} \quad (34)$$

$$g = \sum_{ij} \frac{x(i,j)}{1+|i-j|} \quad (35)$$

$$e = \sum_{ij} x(i,j)^2 \quad (36)$$

$$h = - \sum_{k=0}^{255} p(k) \log_2 p(k) \quad (37)$$

Onde:

$p(k)$ – histograma normalizado de uma amostra;

$$c = \sum_{ij} \frac{(i-\mu_i)(j-\mu_j)x(i,j)}{\sigma_i \sigma_j} \quad (38)$$

Onde:

μ_i – valor médio dos pixels da coordenada i de uma amostra;

μ_j - valor médio dos pixels da coordenada j de uma amostra;

σ_i – desvio padrão dos pixels da coordenada i de uma amostra;

σ_j - desvio padrão dos pixels da coordenada j de uma amostra;

$$co = \sum_{ij} |i - j|^2 x(i, j) \quad (39)$$

Com respeito as bandas da imagem LANDSAT-8/OLI utilizadas, segundo Usgs (2017), para detecção da vegetação, as bandas de imagem mais adequadas são: Azul (B2), Verde (B3), Vermelho (B4), Infravermelho próximo (B5), Infravermelho de ondas curtas 1 (B6), Infravermelho de ondas curtas 2 (B7). A essas bandas citadas anteriormente, foram acrescentadas mais duas bandas calculadas a partir dessas primeiras, conforme sugerido por Huete *et al.*, (2002): NDVI e EVI. Dessa forma, considerando um total de 8 características e 8 bandas, resultam em 64 características. Essas características são mostradas na Tabela 11.

Tabela 11 – Lista de Características Utilizadas

Índice	Característica	Banda	Índice	Característica	Banda
1	Média	B2	33	Média	B6
2	Variância	B2	34	Variância	B6
3	Desvio padrão	B2	35	Desvio padrão	B6
4	Homogeneidade	B2	36	Homogeneidade	B6
5	Energia	B2	37	Energia	B6
6	Correlação	B2	38	Correlação	B6
7	Contraste	B2	39	Contraste	B6
8	Entropia	B2	40	Entropia	B6
9	Média	B3	41	Média	B7
10	Variância	B3	42	Variância	B7
11	Desvio padrão	B3	43	Desvio padrão	B7
12	Homogeneidade	B3	44	Homogeneidade	B7
13	Energia	B3	45	Energia	B7
14	Correlação	B3	46	Correlação	B7
15	Contraste	B3	47	Contraste	B7
16	Entropia	B3	48	Entropia	B7
17	Média	B4	49	Média	NDVI
18	Variância	B4	50	Variância	NDVI
19	Desvio padrão	B4	51	Desvio padrão	NDVI
20	Homogeneidade	B4	52	Homogeneidade	NDVI
21	Energia	B4	53	Energia	NDVI
22	Correlação	B4	54	Correlação	NDVI
23	Contraste	B4	55	Contraste	NDVI
24	Entropia	B5	56	Entropia	NDVI
25	Média	B5	57	Média	EVI
26	Variância	B5	58	Variância	EVI
27	Desvio padrão	B5	59	Desvio padrão	EVI
28	Homogeneidade	B5	60	Homogeneidade	EVI
29	Energia	B5	61	Energia	EVI
30	Correlação	B5	62	Correlação	EVI
31	Contraste	B5	63	Contraste	EVI
32	Entropia	B5	64	Entropia	EVI

Para selecionar as melhores características foram utilizadas a técnica de seleção escalar de características, PCA e LDA, descritas no Capítulo 3 deste trabalho. Os conjuntos de características selecionadas foram compostos por 10, 20, 30, 40, 50 e 64 características. Na seção 5.1 é mostrado o conjunto de características selecionado pela técnica de seleção escalar para cada uma das regiões analisadas nesta dissertação.

4.2.2 Classificação

Essa fase consistiu em implementar o classificador de rede neural para classificar as regiões obtidas na fase de aquisição de dados.

Os classificadores utilizados nesse trabalho para o reconhecimento de áreas desmatadas foram as redes neurais de propagação direta, com três critérios de parada: Erro Médio Quadrático, Regularização e Parada Antecipada. Os dois últimos são métodos utilizados para melhoria da generalização da rede neural. Inicialmente, foram treinados e testados classificadores distintos, com os conjuntos de treinamento e teste extraídos de cada uma das regiões descritas na Tabela 9, separadamente. A seguir foram treinados e testados classificadores com os conjuntos de treinamento e teste extraídos das três regiões, simultaneamente (Apuí, Humaitá e Lábrea). Por fim, para comparar a metodologia proposta, foram treinadas e testados classificadores para as mesmas regiões utilizadas por Reis (2014), Pimentel (2014) e Zanotta (2014). Esses trabalhos foram escolhidos pois também utilizaram imagens ópticas LANDSAT e abordaram a detecção de áreas de desmatamento na região Amazônica.

A Figura 20 ilustra a associação entre técnica e seleção de características e tamanho do conjunto de treinamento e teste.

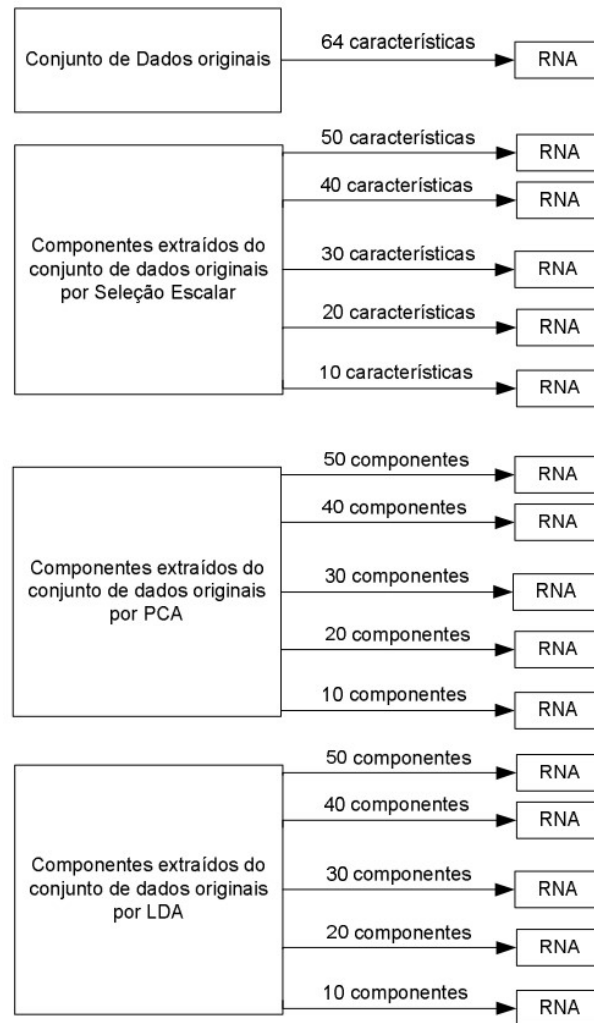


Figura 20 – Estrutura utilizada

Cada RNA foi treinada com os métodos de parada Erro Médio Quadrático, Parada Antecipada e Regularização. O primeiro e o último critério mencionados são divididos e testados com as metodologias de treinamento de validação cruzada com 5 pastas (*5-fold cross validation*). Para o critério de Parada Antecipada, o conjunto de dados foi dividido em três partes, treinamento, validação e teste. Cada etapa será detalhada neste capítulo.

A Tabela 12 apresenta um sumário dos passos metodológicos implementados. Esses passos adotados totalizam 48 cenários distintos de treinamento e teste de uma rede neural para cada região analisada, a saber: Apuí, Humaitá, Lábrea, dados das três regiões (Apuí, Humaitá, Lábrea) juntas, região utilizada por Reis (2014), região utilizada por Pimentel (2014) e região utilizada por Zanotta (2014).

Tabela 12 – Sumário dos passos metodológicos realizados

Método de parada de treinamento da Rede Neural	Metodologias de treinamento e teste	Método de extração de características / número de características
Erro Médio Quadrático	Validação cruzada (5-fold)	Conjunto de dados completo
		Seleção escalar / 50 características
		Seleção escalar / 40 características
		Seleção escalar / 30 características
		Seleção escalar / 20 características
		Seleção escalar / 10 características
		PCA / 50 componentes
		PCA / 40 componentes
		PCA / 30 componentes
		PCA / 20 componentes
		PCA / 10 componentes
		LDA / 50 componentes
		LDA / 40 componentes
		LDA / 30 componentes
		LDA / 20 componentes
LDA / 10 componentes		
Parada Antecipada	Treinamento: 3/5 Validação: 1/5 Teste: 1/5	Conjunto de dados completo
		Seleção escalar / 50 características
		Seleção escalar / 40 características
		Seleção escalar / 30 características
		Seleção escalar / 20 características
		Seleção escalar / 10 características
		PCA / 50 componentes
		PCA / 40 componentes
		PCA / 30 componentes
		PCA / 20 componentes
		PCA / 10 componentes
		LDA / 50 componentes
		LDA / 40 componentes
		LDA / 30 componentes
		LDA / 20 componentes
LDA / 10 componentes		
Regularização	Validação cruzada (5-fold)	Conjunto de dados completo
		Seleção escalar / 50 características
		Seleção escalar / 40 características
		Seleção escalar / 30 características
		Seleção escalar / 20 características
		Seleção escalar / 10 características
		PCA / 50 componentes
		PCA / 40 componentes
		PCA / 30 componentes

Método de parada de treinamento da Rede Neural	Metodologias de treinamento e teste	Método de extração de características / número de características
		PCA / 20 componentes
		PCA / 10 componentes
		LDA / 50 componentes
		LDA / 40 componentes
		LDA / 30 componentes
		LDA / 20 componentes
		LDA / 10 componentes

Para cada uma dessas combinações foi gerada uma tabela contendo os seguintes resultados: acurácia, sensibilidade, especificidade, AUC e desvio padrão da AUC. Com o objetivo de comparar os métodos SEC, LDA, PCA e os critérios de parada da rede neural utilizados, foram calculados testes de significância estatística proposto por (HANLEY e MCNEIL, 1982) utilizados para avaliar as diferenças entre os valores para AUC. Essas informações são apresentadas no capítulo 5 (Resultados).

4.2.2.1 O algoritmo de treinamento: Levenberg Marquardt

O algoritmo de treinamento da rede neural utilizado nesse trabalho foi o Levenberg Marquardt. Segundo Hagan, Demuth e Beale (1995) *apud* Gil (2014) este algoritmo converge em menos iterações do que a maioria dos métodos conhecidos na literatura. Uma desvantagem do mesmo é apresentar um custo computacional maior, dado que, em cada iteração, é necessário calcular a matriz inversa.

Segundo Marquardt (1963) *apud* Almeida (2015), este algoritmo de treinamento realiza uma estimação de parâmetros não lineares e possui uma boa velocidade de treinamento quando é respeitada suas restrições de uso. O treinamento baseia-se na aproximação do Hessiano (H) do conjunto de treinamento pelo produto de Jacobianos (J) deste mesmo conjunto:

$$H = J^T J \quad (40)$$

Nesse caso, o gradiente é calculado como:

$$g = J^T e \quad (41)$$

J é a matriz Jacobiana. Ela possui as primeiras derivadas dos erros da rede em relação aos pesos, e e é corresponde a um vetor de erros da rede neural.

O processo de atualização dos pesos que ocorre durante o treinamento da rede é realizada por meio da Eq. (42) (LERA e PINZOLAS, 2002 apud ALMEIDA, 2015). Neste caso, o termo x_{k+1} é o elemento na próxima iteração, x_k é o valor atual do elemento, η é a taxa de treinamento da rede e e , o Erro Médio Quadrático do problema:

$$x_{k+1} = x_k - [J^T J + \mu I]^{-1} J^T e \quad (42)$$

O gradiente é calculado na Eq. (42) por meio do termo $J^T e$. O termo $\mu.I$ foi acrescentado em (42) para assegurar que a matriz $(J^T.J + \mu.I)$ seja positiva e definida para qualquer dimensão dos vetores de entrada.

Como o método de Newton converge otimamente quando o erro é pequeno, busca-se diminuir μ com a evolução do treinamento para aproximar-se da solução desejada.

Uma restrição ao uso deste algoritmo, segundo Moré (1978) apud Almeida (2015) é que o mesmo possui uma intensa utilização de memória do computador durante a fase de treinamento. Isso cria uma limitação, fazendo com que seja aplicável apenas em redes pequenas, normalmente apresentando algumas centenas de pesos a serem ajustados.

4.2.2.2 Critérios de parada da RNA

Nesta dissertação foram utilizados três critérios para a parada do treinamento: Erro Médio Quadrático, Parada Antecipada e Regularização.

Segundo Almeida (2015), ao utilizar-se o critério de treinamento do Erro Médio Quadrático, o treinamento é finalizado quando uma das condições listadas a seguir ocorrer:

1. O número máximo de 1000 iterações é alcançado;
2. O mínimo erro médio quadrático de 10^{-6} é alcançado.

Segundo WEI (1994) apud Almeida (2015), o Erro Médio Quadrático ou Erro Quadrático Médio (EQM) é uma maneira de realizar uma avaliação da diferença entre um valor estimado e o verdadeiro valor da quantidade estimada. O EQM fornece a média do quadrado do erro. Esse erro corresponde ao montante pelo qual o valor estimado difere da quantidade a ser estimada.

EQM é empregado em modelagem estatística para determinar a medida com que um modelo não se adequou aos dados. O EQM fornece uma maneira de escolher o melhor

estimador: um EQM mínimo frequentemente, mas nem sempre, indica variação mínima e, portanto, um bom estimador (WEI, 1994 apud ALMEIDA, 2015).

Obter um Erro Quadrático Médio igual a zero (0) é a situação ideal, mas na maioria das circunstâncias nunca é possível. O EQM igual a zero significa que o estimador prevê observações de maneira exata (ALMEIDA, 2015).

Em uma rede neural treinada com N padrões, o Erro Médio Quadrático é dado pela Eq. (43).

$$E_{QM} = \frac{\sum_{n=1}^N (S_{on} - S_{en})^2}{N}. \quad (43)$$

Em que:

S_{on} = Valor obtido na saída para o padrão de treinamento n .

S_{en} = Valor esperado na saída para o padrão de treinamento n

Segundo Almeida (2015), ao utilizar-se o critério de parada de Parada Antecipada, o treinamento é encerrado quando uma dentre as seguintes condições for alcançada:

1. O número máximo de 1000 iterações é alcançado.
2. O erro médio quadrático de 10^{-6} é alcançado.
3. O erro do conjunto de Validação aumentou mais de 6 vezes desde a última vez que diminuiu.

O processo de treinamento com o auxílio da metodologia Parada Antecipada é detalhado no Capítulo 3.

Segundo Almeida (2015), ao utilizar-se o critério de parada de Regularização, o treinamento é encerrado quando uma dentre as seguintes condições for alcançada:

1. O número máximo de 1000 iterações é alcançado;
2. O mínimo erro médio quadrático regularizado de 10^{-6} é alcançado;

A expressão do erro regularizado, conforme mostrado na Equação 17 do capítulo 3, é uma soma ponderada do erro médio quadrático, E_{QM} , com a soma dos quadrados dos pesos da rede, E_W .

4.2.2.3 Divisão do banco de dados em conjuntos de treinamento e teste

Para as redes Neurais que utilizaram o Erro Médio Quadrático e a Regularização como critério de parada, foi empregado o método de treinamento e teste de validação cruzada com 5

pastas. O método de validação cruzada consiste em dividir o conjunto total de dados em k subconjuntos mutuamente exclusivos do mesmo tamanho. A partir disto um subconjunto é utilizado para teste e os $k-1$ restantes são utilizados para estimação dos parâmetros. Este processo é realizado k vezes, mudando-se o subconjunto de teste (KOHAVI, 1995). Em cada vez calcula-se uma acurácia. A acurácia média é a média dessas k acurácias. Neste trabalho, k assumiu o valor 5.

Na metodologia de validação cruzada com 5 pastas, o conjunto de 2000 amostras (1000 regiões de desmatamento e 1000 amostras de regiões não-desmatadas que correspondem a 500 amostras de regiões de floresta e 500 amostras de regiões não floresta) é dividido em 5 grupos de 400 amostras cada (200 de desmatamentos, 200 de área não desmatada). A rede é treinada com os dados de quatro grupos e testado com o quinto grupo. Como existem 5 grupos, esse processo é repetido 5 vezes e uma média do desempenho é calculada. A Tabela 13 detalha como foi realizada a divisão dos grupos nos 5 experimentos, segundo a técnica de validação cruzada com 5 pastas.

Tabela 13 – Detalhamento da divisão dos grupos nos cinco experimentos segundo a técnica de validação cruzada com 5 pastas

Experimento	Conjunto de dados	Grupos				
		G1	G2	G3	G4	G5
1	Treinamento		X	X	X	X
	Teste	X				
2	Treinamento	X		X	X	X
	Teste		X			
3	Treinamento	X	X		X	X
	Teste			X		
4	Treinamento	X	X	X		X
	Teste				X	
5	Treinamento	X	X	X	X	
	Teste					X

No treinamento da rede neural utilizando o critério de Parada Antecipada, o conjunto de dados é separado igualmente em três conjuntos: treinamento, validação e teste. Essa separação deve ser feita de tal forma a cada conjunto ter uma boa representação de amostras do conjunto de dados utilizado.

4.2.2.4 Arquitetura da rede

A arquitetura da rede utilizada possui três camadas. O número de nós na camada de entrada é igual ao número de atributos de entrada ($i = 64, 50, 40, 30, 20, 10$). O número de nós

da camada de saída é igual ao número de classes (2). Há apenas uma camada escondida. Algumas fórmulas heurísticas sugerem valores para o número de neurônios na camada escondida da rede MLP. Algumas regras heurísticas são apresentadas a seguir.

Regra da Raiz Quadrada (KOVÁCS, 2002 *apud* ALMEIDA 2015): o número de neurônios da camada escondida é igual a média geométrica do número de entradas (E) e do número de saídas (S) da rede. Por exemplo, para um problema com 64 entradas e 2 duas saídas, o número de neurônios na camada escondida será 11, conforme demonstrado na Eq. (44).

$$n_{H1} = \sqrt{E * S} = \sqrt{64 * 2} \cong 11 \quad (44)$$

Regra do Valor Médio: o número de neurônios da camada escondida é a média aritmética do número de entradas (E) e do número de saídas (S) da rede. Por exemplo, para um problema com 64 entradas e 2 duas saídas, o número de neurônios na camada escondida será 33, conforme demonstrado na Eq. (45).

$$n_{H2} = \frac{E + S}{2} \cong 33 \quad (45)$$

Regra de Kolmogorov (KOLMOGOROV, 1956 *apud* ALMEIDA, 2015): o número de neurônios da camada escondida é igual a duas vezes o número de entradas (E) somado de um. Por exemplo, para um problema com 64 entradas e 2 duas saídas, o número de neurônios na camada escondida será 129, conforme demonstrado na Eq. (46).

$$n_{H3} = 2E + 1 \cong 129 \quad (46)$$

Para este trabalho, foi utilizado a Regra da Raiz Quadrada, desta forma, a quantidade de neurônios da camada escondida foi de 11 neurônios.

4.2.2.5 Aplicação das métricas

Várias métricas têm sido utilizadas para avaliar o desempenho de métodos de segmentação. As mais comuns incluem taxas de sucesso e métricas de similaridade. (SAFARZADEH, 2013).

Neste trabalho foram calculadas a sensibilidade, especificidade, acurácia e curva ROC (*Receiver Operating Characteristic*), que são métricas que medem a taxa de sucesso; e os valores de Dice e Jaccard, que são métricas de similaridade.

A seguir apresentam-se as definições dessas métricas. Nas definições a seguir, G é o conjunto de pixels que pertence ao objeto segmentado manualmente pelos especialistas, chamado de padrão ouro. A é o conjunto de pixels que pertence ao objeto segmentado por um método de segmentação automático. Quando um pixel pertence ao objeto de interesse, atribui-se o valor 1 ao mesmo. Caso contrário, atribui-se o valor 0 ao mesmo. O resultado da classificação de um pixel i é considerado como:

- Verdadeiro positivo (TP, do inglês *True Positive*) quando seu valor no método automático A_i e no padrão ouro G_i são iguais a 1.
- Falso positivo (FP, do inglês *False Positive*) quando A_i é 1 e G_i é 0.
- Falso negativo (FN, do inglês *False Negative*) quando A_i é 0 e G_i é 1.
- Verdadeiro negativo (TN, do inglês *True Negative*) quando A_i e G_i são iguais a 0.

A sensibilidade, especificidade e acurácia são definidas como:

$$\text{Sensibilidade} = \frac{TP}{TP + FN} \quad (47)$$

$$\text{Especificidade} = \frac{TN}{TN + FP} \quad (48)$$

$$\text{Acurácia} = \frac{TP + TF}{N} \quad (49)$$

A sensibilidade reflete o quanto o classificador é eficaz em identificar corretamente regiões desmatadas. A especificidade reflete o quanto o classificador é eficaz em identificar corretamente regiões que não apresentam desmatamento. Por fim, a acurácia reflete o acerto total do classificador.

A curva ROC é uma ferramenta que permite avaliar o desempenho de sistemas de reconhecimento de padrões. Um exemplo prático da aplicação da curva ROC é mostrado na Figura 21. A Figura 21(a) mostra duas funções de densidade de probabilidade sobrepostas, juntamente com um limiar, que descrevem as distribuições de uma característica em duas classes representadas por w_1 e w_2 . Um determinado valor pertence a classe w_1 com uma probabilidade de erro α e probabilidade de acerto $(1-\alpha)$. De modo similar, β é a probabilidade de erro e $(1-\beta)$ é a probabilidade de acerto em relação à classe w_2 . Quando o limiar é variado em todas as direções possíveis, obtêm-se diferentes valores de α e β . Se as duas distribuições possuírem sobreposição completa, em qualquer posição do limiar $\alpha=1-\beta$. Esse caso corresponde à diagonal na Figura 21(b), no qual os eixos são α e $(1-\beta)$. Conforme as distribuições se separam,

a curva correspondente se afasta da reta, conforme mostrado na Figura 21(b). Portanto, a sobreposição das classes é menor quanto maior for a área sob a curva. O valor da área varia de zero a um. Quando duas classes são extremamente separáveis, o valor da área corresponde a um. Em problemas de reconhecimento de padrões, ao se aplicar a curva ROC para avaliar o desempenho do classificador, a saída do classificador é variada e são calculadas as porcentagens de erro e acerto de classificação de vetores de características das classes definidas (THEODORIDIS e KOUTROUMBAS, 2009).

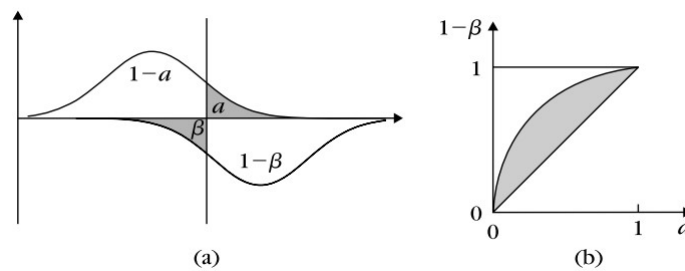


Figura 21 – Exemplo de (a) sobreposição de duas funções de densidade de probabilidade de uma característica em duas classes (b) Curva ROC.
Fonte: Theodoridis e Koutroumbas (2009).

A Dice e Jaccard são duas métricas muito comuns utilizadas para medir a fração de sobreposição espacial entre duas imagens binárias (SAFARZADEH, 2013).

O coeficiente Dice (DICE, 1945) é definido como a interseção entre duas regiões rotuladas r em G e A sobre a área média dessas duas regiões, conforme mostrado na Eq. (50).

$$Dice = \frac{2|G_r \cap A_r|}{|G_r| + |A_r|} = \frac{2|TP|}{(|TP| + |FP| + |TP| + |FN|)} \quad (50)$$

O coeficiente Jaccard (JACCARD, 1912) é definido como a interseção entre duas regiões rotuladas similarmente, r , sobre suas uniões:

$$Jaccard = \frac{|G_r \cap A_r|}{|G_r \cup A_r|} = \frac{|TP|}{(|TP| + |FP| + |FN|)} \quad (51)$$

Esses coeficientes de similaridade variam na faixa de 0 a 1, donde o valor 0 indica que os conjuntos não têm regiões comuns e 1 indica que os conjuntos são equivalentes.

CONSIDERAÇÕES FINAIS

Este capítulo apresentou a metodologia utilizada nessa dissertação. Em relação aos materiais, foram apresentados o ambiente de desenvolvimento, as imagens utilizadas e as regiões analisadas. Além disso, foi explicado o método de reconhecimento de padrões empregado, o conjunto de características escolhido e as técnicas empregadas para a seleção de características. Apresentou-se também a metodologia utilizada para a validação dos resultados. O próximo capítulo apresentará os resultados obtidos e uma discussão desses resultados.

5 RESULTADOS

Neste Capítulo, apresenta-se inicialmente o conjunto de características escolhido utilizando o critério de seleção escalar. Em seguida, são apresentados os resultados obtidos para cada uma das regiões descritas nas Tabela 9 e 10. Essa última tabela contém os trabalhos de comparação. Para cada método de parada da RNA (Erro médio quadrático, Parada Antecipada e Regularização) são apresentados dados de desempenho dos classificadores combinados com os diferentes métodos de extração de características descritos no Capítulo 4, que estão sumarizados na Tabela 12. Esses dados incluem a acurácia, a sensibilidade, a especificidade e AUC.

5.1 Seleção escalar: conjunto de características selecionado

Nesse trabalho, conforme já salientado, a técnica de seleção escalar de características está associada ao critério de separação entre classes FDR (*Fisher's Discriminant Ratio*).

5.1.1 Região de Apuí

Para a região de Apuí, a característica variância (Banda B4) foi a que apresentou o maior valor de FDR, igual a 1,3178. Na Tabela 14 são listados todos os valores de FDR, em ordem decrescente, para o conjunto de 64 características. A coluna índice representa a ordem de aparecimento da característica na Tabela 11.

Tabela 14 – Valores de FDR para o conjunto de variáveis - Apuí

Índice	Característica	Banda	Valor - FDR	Índice	Característica	Banda	Valor - FDR
18	Variância	B4	1,3178	40	Entropia	B6	0,0166
19	Desvio padrão	B4	0,9681	25	Média	B5	0,0125
2	Variância	B2	0,8781	52	Homogeneidade	NDVI	0,0122
26	Variância	B5	0,4825	44	Homogeneidade	B7	0,0114
3	Desvio padrão	B2	0,4407	36	Homogeneidade	B6	0,0113
42	Variância	B7	0,2561	20	Homogeneidade	B4	0,0111
50	Variância	NDVI	0,2561	12	Homogeneidade	B3	0,0110
58	Variância	EVI	0,2561	28	Homogeneidade	B5	0,0109
27	Desvio padrão	B5	0,2456	4	Homogeneidade	B2	0,0108
43	Desvio padrão	B7	0,2283	60	Homogeneidade	EVI	0,0102
51	Desvio padrão	NDVI	0,2283	55	Contraste	NDVI	0,0033
59	Desvio padrão	EVI	0,2283	31	Contraste	B5	0,0026
33	Média	B6	0,2214	15	Contraste	B3	0,0021

Índice	Característica	Banda	Valor - FDR	Índice	Característica	Banda	Valor - FDR
41	Média	B7	0,2139	7	Contraste	B2	0,0020
49	Média	NDVI	0,2139	23	Contraste	B4	0,0019
57	Média	EVI	0,2139	39	Contraste	B6	0,0017
17	Média	B4	0,1829	47	Contraste	B7	0,0016
34	Variância	B6	0,0941	53	Energia	NDVI	0,0013
35	Desvio padrão	B6	0,0840	29	Energia	B5	0,0012
61	Energia	EVI	0,0702	37	Energia	B6	0,0012
9	Média	B3	0,0621	45	Energia	B7	0,0011
10	Variância	B3	0,0589	21	Energia	B4	0,0011
63	Contraste	EVI	0,0483	13	Energia	B3	0,0011
1	Média	B2	0,0211	5	Energia	B2	0,0011
11	Desvio padrão	B3	0,0168	6	Correlação	B2	0,0003
32	Entropia	B5	0,0166	14	Correlação	B3	0,0003
8	Entropia	B2	0,0166	22	Correlação	B4	0,0002
24	Entropia	B5	0,0166	38	Correlação	B6	0,0002
16	Entropia	B4	0,0166	46	Correlação	B7	0,0002
48	Entropia	B7	0,0166	30	Correlação	B5	0,0002
56	Entropia	NDVI	0,0166	54	Correlação	NDVI	0,0002
64	Entropia	EVI	0,0166	62	Correlação	EVI	0,0001

A Tabela 15 mostra as melhores características selecionadas para os conjuntos com 50, 40, 30, 20 e 10 características, com a técnica de seleção escalar de características, para a região de Apuí.

Tabela 15 – Conjunto de Características selecionados - Apuí

Número de Características	Conjunto de Características Selecionado (Índice)
50	18,19,2,63,5,62,26,61,3,53,55,30,13,42,31,6,50,21,58,15,14,29,27,7,54,33,45,43,23,37,22,51,59,39,46,57,47,49,38,41,17,34,60,35,52,9,44,36,20,10
40	18,19,2,63,5,62,26,61,3,53,55,30,13,42,31,6,50,21,58,15,14,29,27,7,54,33,45,43,23,37,22,51,59,39,46,57,47,49,38,41
30	18,19,2,63,5,62,26,61,3,53,55,30,13,42,31,6,50,21,58,15,14,29,27,7,54,33,45,43,23,37
20	18,19,2,63,5,62,26,61,3,53,55,30,13,42,31,6,50,21,58,15
10	18,19,2,63,5,62,26,61,3,53

5.1.2 Região de Humaitá

Para a região de Humaitá, a característica variância (Banda B5) foi a que apresentou o maior valor de FDR, igual a 2,2467. Na Tabela 16 são listados todos os valores de FDR, em ordem decrescente, para o conjunto de 64 características. A coluna índice representa a ordem de aparecimento da característica na Tabela 11.

Tabela 16 – Valores de FDR para o conjunto de variáveis - Humaitá

Índice	Característica	Banda	Valor - FDR	Índice	Característica	Banda	Valor - FDR
26	Variância	B5	2,2467	50	Variância	NDVI	0,0670
27	Desvio padrão	B5	1,4351	58	Variância	EVI	0,0670
25	Média	B5	0,2812	36	Homogeneidade	B6	0,0599
35	Desvio padrão	B6	0,2074	44	Homogeneidade	B7	0,0599
19	Desvio padrão	B4	0,1792	20	Homogeneidade	B4	0,0598
34	Variância	B6	0,1753	12	Homogeneidade	B3	0,0597
33	Média	B6	0,1608	4	Homogeneidade	B2	0,0596
11	Desvio padrão	B3	0,1507	28	Homogeneidade	B5	0,0596
40	Entropia	B6	0,1244	60	Homogeneidade	EVI	0,0578
48	Entropia	B7	0,1244	53	Energia	NDVI	0,0482
56	Entropia	NDVI	0,1244	29	Energia	B5	0,0481
64	Entropia	EVI	0,1244	37	Energia	B6	0,0478
16	Entropia	B3	0,1243	5	Energia	B2	0,0478
24	Entropia	B5	0,1243	13	Energia	B3	0,0478
32	Entropia	B5	0,1243	21	Energia	B4	0,0478
8	Entropia	B2	0,1243	45	Energia	B7	0,0477
43	Desvio padrão	B7	0,1229	1	Média	B2	0,0444
51	Desvio padrão	NDVI	0,1229	9	Média	B3	0,0393
59	Desvio padrão	EVI	0,1229	41	Média	B7	0,0268
55	Contraste	NDVI	0,1068	49	Média	NDVI	0,0268
31	Contraste	B5	0,1063	57	Média	EVI	0,0268
7	Contraste	B2	0,1058	61	Energia	EVI	0,0266
15	Contraste	B3	0,1057	2	Variância	B2	0,0215
23	Contraste	B4	0,1056	17	Média	B4	0,0117
47	Contraste	B7	0,1048	54	Correlação	NDVI	0,0096
63	Contraste	EVI	0,1046	30	Correlação	B5	0,0064
39	Contraste	B6	0,1045	38	Correlação	B6	0,0062
18	Variância	B4	0,1041	62	Correlação	EVI	0,0059
3	Desvio padrão	B2	0,0826	46	Correlação	B7	0,0059
10	Variância	B3	0,0746	6	Correlação	B2	0,0055
52	Homogeneidade	NDVI	0,0681	14	Correlação	B3	0,0055
42	Variância	B7	0,0670	22	Correlação	B4	0,0054

A Tabela 17 mostra as melhores características selecionadas para os conjuntos com 50, 40, 30, 20 e 10 características, com a técnica de seleção escalar de características, para a região de Humaitá.

Tabela 17 – Conjunto de Características selecionados - Humaitá

Número de Características	Conjunto de Características Selecionado
50	26,27,63,53,62,55,25,5,31,54,13,7,35,15,21,34,6,23,29,19,47,14,39,45,33,30,37,11,22,18,46,61,38,40,56,48,64,51,52,59,16,24,43,32,8,60,44,36,10,20
40	26,27,63,53,62,55,25,5,31,54,13,7,35,15,21,34,6,23,29,19,47,14,39,45,33,30,37,11,22,18,46,61,38,40,56,48,64,51,52,59

Número de Características	Conjunto de Características Selecionado
30	26,27,63,53,62,55,25,5,31,54,13,7,35,15,21,34,6,23,29,19,47,14,39,45,33,30,37,11,22,18
20	26,27,63,53,62,55,25,5,31,54,13,7,35,15,21,34,6,23,29,19
10	26,27,63,53,62,55,25,5,31,54

5.1.3 Região de Lábrea

Para a região de Lábrea, a característica média (Banda B2) foi a que apresentou o maior valor de FDR, igual a 0,6355. Na Tabela 18 são listados todos os valores de FDR, em ordem decrescente, para o conjunto de 64 características. A coluna índice representa a ordem de aparecimento da característica na Tabela 11.

Tabela 18 – Valores de FDR para o conjunto de variáveis - Lábrea

Índice	Característica	Banda	Valor - FDR	Índice	Característica	Banda	Valor - FDR
1	Média	B2	0,6355	11	Desvio padrão	B3	0,0180
17	Média	B4	0,5770	62	Correlação	EVI	0,0160
9	Média	B3	0,5737	52	Homogeneidade	NDVI	0,0156
41	Média	B7	0,2962	28	Homogeneidade	B5	0,0155
49	Média	NDVI	0,2962	60	Homogeneidade	EVI	0,0154
57	Média	EVI	0,2962	36	Homogeneidade	B6	0,0152
33	Média	B6	0,2652	44	Homogeneidade	B7	0,0152
25	Média	B5	0,0978	12	Homogeneidade	B3	0,0152
63	Contraste	EVI	0,0859	20	Homogeneidade	B4	0,0151
47	Contraste	B7	0,0830	4	Homogeneidade	B2	0,0151
39	Contraste	B6	0,0825	54	Correlação	NDVI	0,0114
23	Contraste	B4	0,0813	30	Correlação	B5	0,0112
15	Contraste	B3	0,0796	14	Correlação	B3	0,0101
7	Contraste	B2	0,0790	6	Correlação	B2	0,0101
31	Contraste	B5	0,0747	22	Correlação	B4	0,0098
2	Variância	B2	0,0710	38	Correlação	B6	0,0094
55	Contraste	NDVI	0,0696	46	Correlação	B7	0,0092
27	Desvio padrão	B5	0,0552	34	Variância	B6	0,0014
26	Variância	B5	0,0498	45	Energia	B7	0,0005
3	Desvio padrão	B2	0,0493	53	Energia	NDVI	0,0005
40	Entropia	B6	0,0461	37	Energia	B6	0,0005
48	Entropia	B7	0,0461	21	Energia	B4	0,0005
56	Entropia	NDVI	0,0461	29	Energia	B5	0,0005
64	Entropia	EVI	0,0461	13	Energia	B3	0,0005
16	Entropia	B3	0,0461	5	Energia	B2	0,0005
24	Entropia	B5	0,0461	35	Desvio padrão	B6	0,0003
32	Entropia	B5	0,0461	42	Variância	B7	0,0001
8	Entropia	B2	0,0461	50	Variância	NDVI	0,0001

Índice	Característica	Banda	Valor - FDR
61	Energia	EVI	0,0412
18	Variância	B4	0,0304
10	Variância	B3	0,0289
19	Desvio padrão	B4	0,0284

Índice	Característica	Banda	Valor - FDR
58	Variância	EVI	0,0001
43	Desvio padrão	B7	0,0000
51	Desvio padrão	NDVI	0,0000
59	Desvio padrão	EVI	0,0000

A Tabela 19 mostra as melhores características selecionadas para os conjuntos com 50, 40, 30, 20 e 10 características, com a técnica de seleção escalar de características, para a região de Lábrea.

Tabela 19 – Conjunto de Características selecionados - Lábrea

Número de Características	Conjunto de Características Selecionado
50	1,5,17,62,9,63,61,54,47,41,13,49,30,57,21,39,33,14,29,23,6,15,37,7,22,45,31,38,53,55,60,46,44,2,36,25,28,20,52,12,26,18,4,27,40,56,3,48,64,16
40	1,5,17,62,9,63,61,54,47,41,13,49,30,57,21,39,33,14,29,23,6,15,37,7,22,45,31,38,53,55,60,46,44,2,36,25,28,20,52,12
30	1,5,17,62,9,63,61,54,47,41,13,49,30,57,21,39,33,14,29,23,6,15,37,7,22,45,31,38,53,55
20	1,5,17,62,9,63,61,54,47,41,13,49,30,57,21,39,33,14,29,23
10	1,5,17,62,9,63,61,54,47,41

5.1.4 Três Regiões (Apuí, Humaitá, Lábrea)

Utilizando dados das três regiões (Apuí, Humaitá e Lábrea) juntas, a característica variância (Banda B4) foi a que apresentou o maior valor de FDR, igual a 0,2062. Na Tabela 20 são listados todos os valores de FDR, em ordem decrescente, para o conjunto de 64 características. A coluna índice representa a ordem de aparecimento da característica na Tabela 11.

Tabela 20 – Valores de FDR para o conjunto de variáveis

Índice	Característica	Banda	Valor - FDR	Índice	Característica	Banda	Valor - FDR
18	Variância	B4	0,2062	44	Homogeneidade	B7	0,0574
19	Desvio padrão	B4	0,2052	12	Homogeneidade	B3	0,0574
63	Contraste	EVI	0,1958	20	Homogeneidade	B4	0,0574
27	Desvio padrão	B5	0,1625	4	Homogeneidade	B2	0,0573
55	Contraste	NDVI	0,1496	10	Variância	B3	0,0542
31	Contraste	B5	0,1495	60	Homogeneidade	EVI	0,0540
15	Contraste	B3	0,1491	3	Desvio padrão	B2	0,0410
23	Contraste	B4	0,1489	53	Energia	NDVI	0,0303
7	Contraste	B2	0,1488	29	Energia	B5	0,0280
39	Contraste	B6	0,1486	37	Energia	B6	0,0278
47	Contraste	B7	0,1482	21	Energia	B4	0,0277

Índice	Característica	Banda	Valor - FDR	Índice	Característica	Banda	Valor - FDR
48	Entropia	B7	0,1439	45	Energia	B7	0,0277
56	Entropia	NDVI	0,1439	13	Energia	B3	0,0277
64	Entropia	EVI	0,1439	5	Energia	B2	0,0277
32	Entropia	B5	0,1439	41	Média	B7	0,0204
16	Entropia	B4	0,1439	49	Média	NDVI	0,0204
24	Entropia	B5	0,1439	57	Média	EVI	0,0204
40	Entropia	B6	0,1439	25	Média	B5	0,0189
8	Entropia	B2	0,1439	17	Média	B4	0,0118
43	Desvio padrão	B7	0,1294	2	Variância	B2	0,0111
51	Desvio padrão	NDVI	0,1294	61	Energia	EVI	0,0030
59	Desvio padrão	EVI	0,1294	54	Correlação	NDVI	0,0025
35	Desvio padrão	B6	0,1119	33	Média	B6	0,0024
42	Variância	B7	0,1090	38	Correlação	B6	0,0015
50	Variância	NDVI	0,1090	46	Correlação	B7	0,0014
58	Variância	EVI	0,1090	30	Correlação	B5	0,0014
26	Variância	B5	0,1085	6	Correlação	B2	0,0012
11	Desvio padrão	B3	0,0879	14	Correlação	B3	0,0012
34	Variância	B6	0,0867	22	Correlação	B4	0,0012
52	Homogeneidade	NDVI	0,0678	62	Correlação	EVI	0,0007
28	Homogeneidade	B5	0,0577	1	Média	B2	0,0005
36	Homogeneidade	B6	0,0576	9	Média	B3	0,0003

A Tabela 21 mostra as melhores características selecionadas para os conjuntos com 50, 40, 30, 20 e 10 características, com a técnica de seleção escalar de características, para as três regiões.

Tabela 21 – Conjunto de Características selecionados

Número de Características	Conjunto de Características Selecionado
50	18,63,53,62,55,19,5,31,54,27,13,15,6,23,21,7,48,22,29,39,56,47,45,14,64,37,32,30,16,46,24,61,38,40,50,8,58,42,52,26,60,51,59,43,28,36,44,35,20,34
40	18,63,53,62,55,19,5,31,54,27,13,15,6,23,21,7,48,22,29,39,56,47,45,14,64,37,32,30,16,46,24,61,38,40,50,8,58,42,52,26
30	18,63,53,62,55,19,5,31,54,27,13,15,6,23,21,7,48,22,29,39,56,47,45,14,64,37,32,30,16,46
20	18,63,53,62,55,19,5,31,54,27,13,15,6,23,21,7,48,22,29,39
10	18,63,53,62,55,19,5,31,54,27

Em relação ao critério de seleção escalar, os resultados mostraram que as características Contraste (Banda EVI), Correlação (Banda EVI) e Energia (Banda B2) sempre estão presentes em todos os conjuntos de características selecionados. Para conjuntos de características de tamanho 20, além das características comuns já mencionadas, foi constatada a presença das

características Energia (Banda B3) e Energia (Banda B4). A presença das características Contraste e Correlação da banda EVI pode ser explicada pelo fato da banda EVI ser formada a partir do cálculo do índice de vegetação EVI. Índices de vegetação foram desenvolvidos para explorar as propriedades espectrais de vegetação. Dessa forma, índices de vegetação ressaltam o comportamento espectral da vegetação em relação ao solo e a outros alvos da superfície terrestre.

5.2 Resultados dos classificadores

Nessa sessão são apresentados inicialmente os resultados dos classificadores para todas as regiões analisadas nesta dissertação, que foram Apuí, Humaitá e Lábrea. Em seguida, são apresentados os resultados de classificação para as mesmas regiões analisadas por Zanotta (2014), Reis (2014) e Pimentel (2014). Para aquilatar a significância das diferenças observadas entre os resultados dos métodos de parada do treinamento da rede neural e entre as técnicas de seleção de características utilizadas, foram empregados testes de significância estatística proposto em (HANLEY e MCNEIL, 1982).

5.2.1 Região de Apuí

A Tabela 22 apresenta os resultados obtidos para a classificação feita com dados da região de Apuí. Para a região de Apuí, observa-se que as técnicas SEC e LDA apresentaram melhores desempenhos em relação a técnica PCA com respeito a área AUC. As Tabela 23 e 24 apresentam valores para os testes de significância estatística proposto em (HANLEY e MCNEIL, 1982) utilizados para avaliar as diferenças entre os valores para AUC, para os métodos SEC, LDA e PCA. A partir dessas tabelas observa-se que as técnicas SEC e LDA apresentaram diferenças estatisticamente significantes em relação à técnica PCA. Por exemplo, comparando as técnicas SEC e PCA, quando utilizado o método de parada MSE, a Tabela 24 mostra que o teste de significância estatística resultou em um valor p médio de $0,000023 < 0,0001$. Então, a hipótese nula deve ser rejeitada (isto é, o desempenho da técnica SEC é superior ao da técnica PCA, ao nível de significância de 0,01%).

Comparando as técnicas LDA e PCA, quando utilizado o método de parada MSE, a Tabela 23 mostra que o teste de significância estatística resultou em um valor de p médio de $0,000015 < 0,0001$. Então, a hipótese nula deve ser rejeitada (isto é, o desempenho da técnica LDA é superior ao da técnica PCA, ao nível de significância de 0,01%). Por fim, a comparação entre as técnicas SEC e LDA mostra resultados muito próximos, sendo o desempenho da técnica

LDA levemente superior. Comparando as técnicas SEC e LDA, quando utilizado o método de parada Regularização, a Tabela 23 mostra que o teste de significância estatística resultou em um valor de p médio de $0,28255 > 0,01$. Então, a hipótese nula não deve ser rejeitada (isto é, o desempenho da técnica LDA não é superior ao da técnica SEC, ao nível de significância de 1%).

Com respeito aos métodos de parada da rede neural, observa-se, na Tabela 22, uma pequena superioridade no desempenho do método parada antecipada frente aos métodos MSE e Regularização, em relação à AUC. Porém os testes de significância estatística apresentados nas Tabela 24 e Tabela 25 mostraram que nenhum método de parada foi superior ao outro. Por exemplo, comparando os métodos de parada MSE e parada antecipada quando utilizada a técnica PCA, a Tabela 26 mostra que o teste de significância estatística resultou em um valor p médio de $0,357892 > 0,01$. Então, a hipótese nula não deve ser rejeitada (isto é, o desempenho do método parada antecipada não é superior ao do método MSE, ao nível de significância de 1%).

Tabela 22 – Resultados médios para acurácia, AUC, sensibilidade e especificidade, para os métodos de parada de treinamento MSE, parada antecipada e regularização, combinados as técnicas de extração de características SEC, PCA e LDA, para a região de Apuí.

Região	Métodos de parada do treinamento	N. de Caract.	Técnica de Seleção de Características											
			SEC				PCA				LDA			
			Acur.	AUC	Sens	Esp	Acur.	AUC	Sens	Esp	Acur.	AUC	Sens	Esp
Apuí	MSE	10	99,80	0,998 ± 0,0022	99,80	99,80	95,90	0,959 ± 0,0102	95,60	96,20	99,95	0,999 ± 0,0011	99,90	100,00
		20	99,85	0,998 ± 0,0019	99,70	99,90	94,95	0,950 ± 0,0114	94,30	95,60	99,85	0,998 ± 0,0019	99,90	99,80
		30	99,80	0,998 ± 0,0022	99,70	99,90	93,15	0,931 ± 0,0132	92,40	93,90	99,95	0,998 ± 0,0019	99,90	100,00
		40	99,85	0,998 ± 0,0019	99,80	99,90	93,70	0,937 ± 0,0127	94,30	93,10	99,95	0,999 ± 0,0011	99,90	100,00
		50	99,85	0,998 ± 0,0019	99,70	100,00	92,60	0,926 ± 0,0138	91,20	94,00	99,95	0,999 ± 0,0011	99,90	100,00
		64	99,85	0,998 ± 0,0019	99,80	99,90	99,85	0,998 ± 0,0019	99,80	99,90	99,85	0,998 ± 0,0011	99,80	99,90
	Parada Antecipada	10	99,90	0,999 ± 0,0016	99,85	99,95	97,33	0,973 ± 0,0082	97,35	97,30	99,93	0,999 ± 0,0014	99,85	99,99
		20	99,93	0,999 ± 0,0014	99,85	100,00	98,63	0,986 ± 0,0059	98,70	98,55	99,90	0,999 ± 0,0016	99,85	100,00
		30	99,93	0,999 ± 0,0014	99,85	100,00	95,78	0,958 ± 0,0104	95,70	95,85	99,93	0,999 ± 0,0014	99,85	99,95
		40	99,93	0,999 ± 0,0014	99,85	100,00	93,78	0,938 ± 0,0126	93,60	93,95	99,93	0,999 ± 0,0014	99,85	100,00
		50	99,90	0,999 ± 0,0016	99,80	100,00	95,48	0,955 ± 0,0107	95,80	95,15	99,93	0,999 ± 0,0014	99,85	100,00
		64	99,93	0,999 ± 0,0014	99,85	100,00	99,93	0,999 ± 0,0014	99,85	100,00	99,93	0,999 ± 0,0014	99,85	100,00
	Regularização	10	99,80	0,998 ± 0,0022	99,70	99,90	96,40	0,964 ± 0,0096	96,10	96,70	99,95	0,999 ± 0,0011	99,90	100,00
		20	99,90	0,999 ± 0,0016	99,80	100,00	84,25	0,843 ± 0,0199	81,30	87,20	99,95	0,999 ± 0,0011	99,90	100,00
		30	99,85	0,988 ± 0,0019	99,80	99,90	85,20	0,852 ± 0,0170	87,30	83,10	99,95	0,999 ± 0,0011	99,90	100,00
		40	99,85	0,988 ± 0,0019	99,80	99,90	86,10	0,861 ± 0,0162	81,30	90,90	99,95	0,999 ± 0,0011	99,90	100,00
		50	99,95	0,999 ± 0,0011	99,90	100,00	87,20	0,872 ± 0,0155	82,30	92,10	99,95	0,999 ± 0,0011	99,90	100,00
		64	99,85	0,999 ± 0,0011	99,90	100,00	99,95	0,995 ± 0,0014	99,90	99,90	99,95	0,999 ± 0,0011	99,90	100,00

Tabela 23 – Valores para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para as técnicas SEC, PCA e LDA, para a região de Apuí.

Métodos de parada do treinamento	N. de Características	Teste de Significância Estatística		
		SEC-PCA	SEC-LDA	PCA-LDA
MSE	10	0,000097	0,357892	0,000055
	20	0,000015	0,500000	0,000015
	30	0,000001	0,500000	0,000001
	40	0,000001	0,357892	0,000001
	50	0,000001	0,357892	0,000001
Parada Antecipada	10	0,001020	0,500000	0,001020
	20	0,017361	0,500000	0,017361
	30	0,000045	0,500000	0,000045
	40	0,000001	0,500000	0,000001
	50	0,000024	0,500000	0,000024
Regularização	10	0,000272	0,357892	0,000155
	20	0,000001	0,500000	0,000001
	30	0,000001	0,027429	0,000001
	40	0,000001	0,027429	0,000001
	50	0,000001	0,500000	0,000001

Tabela 24 – Valores médios para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para as técnicas SEC, PCA e LDA, para a região de Apuí.

Métodos de parada do treinamento	Teste de Significância Estatística Médio		
	SEC-PCA	SEC-LDA	PCA-LDA
MSE	0,000023	0,414735	0,000015
Parada Antecipada	0,006142	0,500000	0,006142
Regularização	0,0000552	0,282550	0,0000318

Tabela 25 – Valores para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para os métodos de parada MSE, Parada Antecipada e Regularização, para a região de Apuí.

Técnica de Seleção de Característica	N. de Caract.	Teste de Significância Estatística		
		MSE-PAN	MSE-REG	PAN-REG
SEC	10	0,357892	0,500000	0,357892
	20	0,357892	0,357892	0,500000
	30	0,357892	0,046265	0,027429
	40	0,357892	0,046265	0,027429
	50	0,357892	0,357892	0,500000
PCA	10	0,143574	0,360508	0,23848
	20	0,002400	0,000001	0,000001
	30	0,054357	0,000373	0,000001
	40	0,054357	0,000392	0,000323
	50	0,357892	0,008591	0,151693
LDA	10	0,500000	0,500000	0,500000
	20	0,357892	0,357892	0,500000
	30	0,357892	0,357892	0,500000
	40	0,500000	0,500000	0,500000
	50	0,500000	0,500000	0,500000

Tabela 26 – Valores médios para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para os métodos de parada MSE, Parada Antecipada e Regularização, para a região de Apuí.

Técnica de Seleção de Característica	Teste de Significância Estatística Médio		
	MSE-PAN	MSE-REG	PAN-REG
SEC	0,3578920	0,261663	0,2825500
PCA	0,1225160	0,073973	0,0780996
LDA	0,4431568	0,443156	0,5000000

5.2.2 Região de Humaitá

Para a região de Humaitá, observa-se na Tabela 27 que a técnica LDA apresentou desempenho levemente superior com respeito a AUC em relação as técnicas SEC e PCA. As Tabela 28 e Tabela 29 apresentam valores para os testes de significância estatística utilizados para avaliar as diferenças entre os valores para AUC, para os métodos SEC, LDA e PCA. Comparando as técnicas SEC e PCA, quando utilizado o método de parada MSE com 50 características, a Tabela 28 mostra que o teste de significância estatística resultou em um valor de p de $0,002144 < 0,01$. Nesse caso, a hipótese nula deve ser rejeitada (isto é, o desempenho da técnica SEC é superior ao da técnica PCA, ao nível de significância de 1%). Por outro lado, comparando as técnicas SEC e PCA, quando utilizado o método de parada MSE e um número de características menor (40, 30, 20 ou 10), os valores dos testes de significância estatística apresentados na Tabela 28 mostram que a técnica SEC não é superior a PCA. Comparando as técnicas LDA e SEC quando utilizado o número de características iguais a 10 ou 20 independentemente do método de parada (MSE, PAN ou REG), os valores dos testes de significância estatística apresentados na Tabela 28 mostram que a técnica LDA foi superior a SEC. Por fim, comparando as técnicas LDA e SEC quando utilizado o número de características iguais 40 ou 50 e o método de parada MSE ou REG, os valores dos testes de significância estatística apresentados na Tabela 28 mostram que a técnica LDA foi superior a técnica SEC.

Em relação aos métodos de parada de rede neural, observa-se na Tabela 27, uma pequena superioridade no desempenho do método parada antecipada frente a outros métodos de parada, em relação à AUC. Porém os testes de significância estatística apresentados nas Tabela 30 e 31 mostraram que nenhum método de parada foi superior ao outro. Por exemplo, comparando os métodos de parada MSE e parada antecipada quando utilizado a técnica PCA, a Tabela 31 mostra que o teste de significância estatística resultou em um valor p médio de $0,205951 > 0,01$. Então, a hipótese nula não deve ser rejeitada (isto é, o desempenho do método parada antecipada não é superior ao do método MSE, ao nível de significância de 1%).

Tabela 27 – Resultados médios para acurácia, AUC, sensibilidade e especificidade, para os métodos de treinamento MSE, parada antecipada e regularização, combinadas as técnicas de extração de características SEC, PCA e LDA, para a região de Humaitá

Região	Método de parada do treinamento	N. de Caract.	Técnica de Seleção de Características											
			SEC				PCA				LDA			
			Acur.	AUC	Sens	Esp	Acur.	AUC	Sens	Esp	Acur.	AUC	Sens	Esp
Humaitá	MSE	10	94,25	0,943 ± 0,0121	94,00	94,50	97,05	0,971 ± 0,0087	97,50	96,60	99,55	0,996 ± 0,0034	99,20	99,90
		20	95,05	0,950 ± 0,0112	94,20	95,90	97,75	0,977 ± 0,0076	97,90	97,60	99,35	0,994 ± 0,0040	99,00	99,70
		30	98,85	0,988 ± 0,0054	98,60	99,10	98,00	0,980 ± 0,0071	97,80	98,20	99,60	0,996 ± 0,0032	99,30	99,90
		40	98,70	0,987 ± 0,0057	98,10	99,30	96,85	0,969 ± 0,0089	96,80	96,90	99,60	0,996 ± 0,0032	99,30	99,90
		50	98,80	0,988 ± 0,0055	98,40	99,20	95,25	0,953 ± 0,0110	95,40	95,10	99,50	0,995 ± 0,0035	99,30	99,70
		64	99,25	0,993 ± 0,0043	98,60	99,90	99,25	0,993 ± 0,0043	98,60	99,90	99,25	0,993 ± 0,0043	98,60	99,90
	Parada Antecipada	10	94,10	0,941 ± 0,0122	94,30	93,95	99,43	0,994 ± 0,0038	99,20	99,65	99,65	0,996 ± 0,0030	99,45	99,85
		20	97,00	0,970 ± 0,0087	96,25	97,75	98,98	0,990 ± 0,0051	98,95	99,00	99,68	0,997 ± 0,0029	99,45	99,90
		30	99,45	0,995 ± 0,0037	99,20	99,70	98,88	0,989 ± 0,0053	98,70	99,05	99,73	0,997 ± 0,0026	99,50	99,95
		40	99,53	0,995 ± 0,0035	99,30	99,75	99,08	0,991 ± 0,0048	99,15	99,00	99,73	0,997 ± 0,0026	99,50	99,95
		50	99,33	0,993 ± 0,0041	99,05	99,60	98,48	0,985 ± 0,0062	98,45	98,50	99,68	0,997 ± 0,0029	99,45	99,90
		64	99,60	0,996 ± 0,0032	99,40	99,80	99,60	0,986 ± 0,0032	98,40	99,80	99,60	0,996 ± 0,0032	99,40	99,80
	Regularização	10	93,10	0,931 ± 0,0133	92,00	94,20	98,25	0,982 ± 0,0067	98,40	98,10	99,50	0,995 ± 0,0035	99,30	99,70
		20	95,50	0,955 ± 0,0107	94,80	96,20	98,20	0,982 ± 0,0068	97,60	98,80	99,50	0,995 ± 0,0035	99,10	99,90
		30	99,15	0,991 ± 0,0046	98,70	99,60	97,20	0,977 ± 0,0076	97,70	97,70	99,60	0,996 ± 0,0032	99,30	99,90
		40	98,95	0,990 ± 0,0051	98,40	99,50	96,60	0,966 ± 0,0093	96,80	96,80	99,65	0,996 ± 0,0030	99,30	100,00
		50	98,75	0,988 ± 0,0056	98,10	99,40	96,25	0,963 ± 0,0098	96,10	96,40	99,65	0,996 ± 0,0030	99,30	100,00
		64	99,45	0,995 ± 0,0037	99,00	99,90	99,45	0,995 ± 0,0037	99,00	99,00	99,45	0,995 ± 0,0037	99,00	99,90

Tabela 28 – Valores para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para as técnicas SEC, PCA e LDA, para a região de Humaitá.

Métodos de parada do treinamento	N. de Características	Teste de Significância Estatística		
		SEC-PCA	SEC-LDA	PCA-LDA
MSE	10	0,029306	0,000011	0,003147
	20	0,023835	0,000115	0,023639
	30	0,187002	0,104015	0,020032
	40	0,044228	0,084752	0,00209
	50	0,002144	0,358041	0,000132
Parada Antecipada	10	0,453747	0,000007	0,327907
	20	0,080607	0,001589	0,110759
	30	0,142602	0,327907	0,110759
	40	0,117653	0,327907	0,137658
	50	0,235147	0,327907	0,037587
Regularização	10	0,25178	0,000002	0,044125
	20	0,374029	0,000196	0,044125
	30	0,34017	0,191223	0,010795
	40	0,346938	0,156296	0,001132
	50	0,5	0,5	0,000614

Tabela 29 – Valores médios para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para as técnicas SEC, PCA e LDA, para a região de Humaitá.

Métodos de parada do treinamento	Teste de Significância Estatística Médio		
	SEC-PCA	SEC-LDA	PCA-LDA
MSE	0,057303	0,109387	0,009808
Parada Antecipada	0,205951	0,197063	0,144934
Regularização	0,362583	0,169543	0,020158

Tabela 30 – Valores para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para os métodos de parada MSE, Parada Antecipada e Regularização, para a região de Humaitá.

Técnica de Seleção de Característica	N. de Caract.	Teste de Significância Estatística		
		MSE-PAN	MSE-REG	PAN-REG
SEC	10	0,453747	0,25178	0,290132
	20	0,080607	0,374029	0,13885
	30	0,142602	0,34017	0,250378
	40	0,117653	0,346938	0,208138
	50	0,235147	0,5	0,235147
PCA	10	0,007323	0,156871	0,061731
	20	0,077484	0,311855	0,170878
	30	0,077484	0,311855	0,097954
	40	0,014492	0,407725	0,008367
	50	0,005425	0,247146	0,027804
LDA	10	0,5	0,416795	0,416795
	20	0,264315	0,424655	0,327907
	30	0,405863	0,5	0,405863
	40	0,405863	0,5	0,405863
	50	0,405863	0,5	0,405863

Tabela 31 – Valores médios para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para os métodos de parada MSE, Parada Antecipada e Regularização, para a região de Humaitá.

Técnica de Seleção de Característica	Teste de Significância Estatística Médio		
	MSE-PAN	MSE-REG	PAN-REG
SEC	0,205951	0,362583	0,224529
PCA	0,0364416	0,2870904	0,0733468
LDA	0,3963808	0,46829	0,3924582

5.2.3 Região de Lábrea

Para a região de Lábrea, observa-se na Tabela 32 que a técnica LDA apresentou desempenho levemente superior com respeito a AUC em relação a SEC e PCA. As Tabela 33 e 34 apresentam valores para os testes de significância estatística para avaliar as diferenças entre os valores para AUC, para os métodos SEC, LDA e PCA. A partir dessas tabelas, observa-se que comparando as técnicas SEC e PCA, a técnica PCA foi superior em relação a técnica SEC quando utilizado 10 características independentemente do método de parada da rede neural. Por outro lado, comparando as técnicas SEC e PCA quando utilizado 50 características, os valores dos testes de significância estatística apresentados na Tabela 33 mostram que a técnica SEC foi superior a técnica PCA. Comparando as técnicas LDA, PCA e SEC, os valores dos testes de significância estatística apresentados na Tabela 33 mostram que a técnica LDA foi superior a PCA e também foi superior a técnica SEC quando utilizado um número de características de 10 ou 20.

Com respeito aos métodos de parada da rede neural, observa-se na Tabela 32, uma pequena superioridade no desempenho do método parada antecipada frente os outros métodos de parada da rede neural (MSE e Regularização), em relação a AUC. Porém os testes de significância estatística apresentados nas Tabela 35 e Tabela 36 mostraram que nenhum método de parada da rede neural foi superior ao outro. Por exemplo, comparando os métodos de parada MSE e parada antecipada quando utilizado a técnica SEC, a Tabela 36 mostra que o teste de significância estatística resultou em um valor p médio de $0,157553 > 0,01$. Então, a hipótese nula não deve ser rejeitada (isto é, o desempenho do método parada antecipada não é superior ao desempenho do método MSE, ao nível de significância de 1%).

Tabela 32 – Resultados médios para acurácia, AUC, sensibilidade e especificidade, para os métodos de treinamento MSE, parada antecipada e regularização, combinadas as técnicas de extração de características SEC, PCA e LDA, para a região de Lábrea

Região	Métodos de parada do treinamento	N. de Caract.	Técnica de Seleção de Características											
			SEC				PCA				LDA			
			Acur.	AUC	Sens	Esp	Acur.	AUC	Sens	Esp	Acur.	AUC	Sens	Esp
Lábrea	MSE	10	92,10	0,921 ± 0,0142	92,30	91,90	98,50	0,985 ± 0,0062	98,40	98,60	99,50	0,995 ± 0,0035	99,50	99,50
		20	93,80	0,938 ± 0,0126	94,20	93,40	97,20	0,972 ± 0,0084	97,00	97,40	99,60	0,996 ± 0,0032	99,70	99,50
		30	96,40	0,964 ± 0,0096	95,70	97,10	95,85	0,958 ± 0,0103	96,20	95,50	99,50	0,996 ± 0,0034	99,50	99,60
		40	97,75	0,977 ± 0,0076	97,20	98,30	94,75	0,948 ± 0,0116	94,90	94,60	99,65	0,996 ± 0,0030	99,60	99,70
		50	99,50	0,995 ± 0,0035	99,30	99,70	93,90	0,939 ± 0,0125	93,70	94,10	99,70	0,997 ± 0,0027	99,60	99,80
		64	99,60	0,996 ± 0,0032	99,30	99,90	99,60	0,996 ± 0,0032	99,30	99,90	99,60	0,996 ± 0,0032	99,30	99,90
	Parada Antecipada	10	95,13	0,951 ± 0,0112	95,10	95,15	98,88	0,989 ± 0,0053	98,80	98,95	99,70	0,997 ± 0,0027	99,60	99,80
		20	95,85	0,959 ± 0,0103	96,00	95,70	98,75	0,987 ± 0,0056	98,90	98,60	99,75	0,998 ± 0,0025	99,65	99,85
		30	97,90	0,979 ± 0,0073	97,55	98,25	97,65	0,976 ± 0,0077	97,25	98,05	99,75	0,997 ± 0,0025	99,65	99,85
		40	98,53	0,985 ± 0,0061	98,35	98,70	98,03	0,980 ± 0,0071	97,95	98,10	99,73	0,997 ± 0,0026	99,70	99,75
		50	99,68	0,997 ± 0,0029	99,60	99,75	95,55	0,956 ± 0,0107	94,95	96,15	99,80	0,998 ± 0,0022	99,80	99,80
		64	99,68	0,997 ± 0,0029	99,60	99,75	99,68	0,997 ± 0,0029	99,60	99,75	99,68	0,997 ± 0,0029	99,60	99,75
	Regularização	10	93,50	0,935 ± 0,0129	92,50	94,50	94,50	0,985 ± 0,0062	95,00	94,00	99,50	0,995 ± 0,0035	99,60	99,40
		20	93,45	0,935 ± 0,0129	94,00	92,90	95,50	0,955 ± 0,0104	95,50	95,50	99,85	0,998 ± 0,0019	99,80	88,90
		30	95,65	0,956 ± 0,0105	95,80	95,50	96,70	0,967 ± 0,0083	97,00	96,40	99,75	0,998 ± 0,0025	99,70	99,80
		40	97,60	0,976 ± 0,0078	97,30	97,90	96,80	0,968 ± 0,0081	96,20	97,40	99,70	0,997 ± 0,0027	99,60	99,80
		50	99,65	0,996 ± 0,0030	99,40	99,90	97,50	0,975 ± 0,0076	97,10	97,90	99,70	0,997 ± 0,0027	99,80	99,60
		64	99,55	0,996 ± 0,0034	99,60	99,50	99,55	0,996 ± 0,0034	99,60	99,50	99,55	0,996 ± 0,0034	99,60	99,50

Tabela 33 – Valores para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para as técnicas SEC, PCA e LDA, para a região de Lábrea.

Métodos de parada do treinamento	N. de Características	Teste de Significância Estatística		
		SEC-PCA	SEC-LDA	PCA-LDA
MSE	10	0,0000180	0,0000002	0,0000001
	20	0,0124034	0,0000040	0,0000001
	30	0,3351523	0,0007532	0,0015514
	40	0,0179430	0,0107951	0,0559044
	50	0,0000080	0,3279227	0,4378282
Parada Antecipada	10	0,0010545	0,0000324	0,0000001
	20	0,0084326	0,0000969	0,0000001
	30	0,3894078	0,0104679	0,0147763
	40	0,2976372	0,0375864	0,0578830
	50	0,0000894	0,3889854	0,4631988
Regularização	10	0,0002321	0,0000036	0,0000001
	20	0,1163291	0,0000008	0,0000001
	30	0,2160576	0,0000524	0,0000043
	40	0,2511727	0,0055630	0,0129659
	50	0,0071533	0,4058437	0,4527589

Tabela 34 – Valores médios para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para as técnicas SEC, PCA e LDA, para a região de Lábrea.

Métodos de parada do treinamento	Teste de Significância Estatística Médio		
	SEC-PCA	SEC-LDA	PCA-LDA
MSE	0,073105	0,067895	0,099057
Parada Antecipada	0,139324	0,087434	0,107172
Regularização	0,118189	0,082293	0,093146

Tabela 35 – Valores para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para os métodos MSE, Parada Antecipada e Regularização, para a região de Lábrea.

Técnica de Seleção de Característica	N. de Caract.	Teste de Significância Estatística		
		MSE-PAN	MSE-REG	PAN-REG
SEC	10	0,0485228	0,2327377	0,1741677
	20	0,0976007	0,4338620	0,0722560
	30	0,1063035	0,2875903	0,0368431
	40	0,2074144	0,4635106	0,1826225
	50	0,3279227	0,4168067	0,4058437
PCA	10	0,3108381	0,5000000	0,3108381
	20	0,0706200	0,1061843	0,0042164
	30	0,0823887	0,2574151	0,2272470
	40	0,0090609	0,0858147	0,1481508
	50	0,1495174	0,0075227	0,0758040
LDA	10	0,3279227	0,5000000	0,3279227
	20	0,3033788	0,3033788	0,5000000
	30	0,4058437	0,2649131	0,3889854
	40	0,4058437	0,4058437	0,5000000
	50	0,3889854	0,5000000	0,3889854

Tabela 36 – Valores médios para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para os métodos MSE, Parada Antecipada e Regularização, para a região de Lábrea.

Técnica de Seleção de Característica	Teste de Significância Estatística Médio		
	MSE-PAN	MSE-REG	PAN-REG
SEC	0,1575530	0,3669010	0,1743470
PCA	0,1244850	0,1913874	0,1532513
LDA	0,3663948	0,3948271	0,4211787

5.2.4 Avaliação dos resultados relativos a utilização das três regiões separadamente

Em primeiro lugar, foi observado que a técnica LDA, para as três regiões, apresentou o melhor desempenho em relação a AUC. Por exemplo, para a região de Lábrea, para o método de parada regularização e para um número de características iguais a 30, a AUC e o desvio padrão obtidos para as técnicas SEC, PCA e LDA foram, respectivamente, $0,956 \pm 0,0105$, $0,967 \pm 0,0083$ e $0,998 \pm 0,0025$. Aplicando-se o teste de significância estatística proposto em (HANLEY e MCNEIL, 1982) para avaliar os valores apresentados pelas técnicas LDA e SEC resultou em um valor de $p = 0,0000524 < 0,00001$. Então, a hipótese nula deve ser rejeitada (isto é, o desempenho da técnica LDA foi superior ao da técnica SEC, ao nível de significância de 0,001%). Aplicando-se o teste de significância estatística (HANLEY e MCNEIL, 1982) para avaliar os valores apresentados pelas técnicas LDA e PCA, resultou em um valor de $p = 0,0000043 < 0,0001$. Então, a hipótese nula deve ser rejeitada (isto é, o desempenho da técnica LDA foi superior ao da técnica PCA, ao nível de significância de 0,01%).

Em segundo lugar, com respeito aos métodos de parada da rede neural, foi observada, nas três regiões, uma leve superioridade no desempenho do método parada antecipada frente os outros dois, em relação à AUC. Por exemplo, na região de Humaitá, para a técnica PCA e para um número de características iguais a 50, a AUC e o desvio padrão obtidos para os métodos MSE, parada antecipada e regularização foram, respectivamente, $0,953 \pm 0,0110$, $0,985 \pm 0,0062$ e $0,963 \pm 0,0062$. Aplicando-se o teste de significância estatística proposto em (HANLEY e MCNEIL, 1982) para avaliar os valores apresentados pelos métodos de parada antecipada e MSE, resultou em um valor de $p = 0,005425 < 0,01$. Então, a hipótese nula deve ser rejeitada (isto é, o desempenho do método parada antecipada foi superior ao do método MSE, ao nível de significância de 1%). Aplicando-se o teste de significância estatística proposto em (HANLEY e MCNEIL, 1982) para avaliar os valores apresentados pelos métodos de parada antecipada e regularização, resultou em um valor de $p = 0,027804 > 0,01$. Então, a hipótese nula não deve ser rejeitada (isto é, o desempenho do método parada antecipada não foi superior ao do método regularização, ao nível de significância de 1%).

Em terceiro lugar, com respeito ao número de características, para os métodos SEC e PCA, via de regra o desempenho com um número de características iguais a 64 foi melhor que o desempenho com um número menor de características, em relação a AUC. Por exemplo, para a região de Apuí, quando utilizado a técnica PCA e o método parada antecipada, a AUC e o desvio padrão obtidos para um número de características iguais a 10 e 64 foram, respectivamente, $0,973 \pm 0,0082$ e $0,999 \pm 0,0014$. Aplicando-se o teste de significância estatística proposto em (HANLEY e MCNEIL, 1982) para avaliar os valores apresentados com 10 e 64 variáveis, resultou em um valor de $p = 0,00102 < 0,01$. Então, a hipótese nula deve ser rejeitada (isto é, o desempenho com 64 características é superior ao desempenho com 10 características, ao nível de significância de 1%).

5.2.5 Dados das três Regiões (Apuí, Humaitá e Lábrea) juntas

A Tabela 37 apresenta os resultados obtidos para a classificação feita utilizando dados das três regiões (Apuí, Humaitá e Lábrea) juntas. Pode-se observar nessa tabela que a técnica LDA apresentou desempenho levemente superior com respeito a AUC em relação as técnicas SEC e PCA. As Tabelas 38 e 39 apresentam valores para os testes de significância estatística para avaliar as diferenças entre os valores para AUC, para os métodos SEC, LDA e PCA. A partir dessas tabelas, observa-se que comparando as técnicas SEC e PCA, a técnica PCA foi superior em relação a técnica SEC quando utilizado 30 características e o método de parada MSE. Comparando as técnicas LDA, PCA e SEC, os valores dos testes de significância estatística apresentados na Tabela 38 mostram que a técnica LDA foi superior a PCA e também superior a técnica SEC quando utilizado um número de características iguais a 30 e o método de parada MSE.

Com respeito aos métodos de parada da rede neural, observa-se na Tabela 37, uma pequena superioridade no desempenho do método parada antecipada frente os outros métodos de parada da rede neural (MSE e Regularização), em relação a AUC. Porém os testes de significância estatística apresentados nas Tabela 40 e 41 mostraram que nenhum método de parada da rede neural foi superior ao outro. Por exemplo, comparando os métodos de parada antecipada e regularização quando utilizado a técnica LDA, a Tabela 41 mostra que o teste de significância estatística resultou em um valor de p médio de $0,2371741 > 0,01$. Então, a hipótese nula não deve ser rejeitada (isto é, o desempenho do método parada antecipada não é superior ao do método regularização, ao nível de significância de 1%).

Com respeito ao número de características, quando são utilizados os dados das três regiões simultaneamente, para os métodos SEC e PCA, a Tabela 37 mostra que via de regra o desempenho com 64 características foi melhor que o desempenho com um número menor de características. Por exemplo, para a técnica SEC e parada antecipada, a AUC e o desvio padrão obtidos para 10 e 64 variáveis foram, respectivamente, $0,946 \pm 0,0118$ e $0,993 \pm 0,0042$. Aplicando-se o teste de significância estatística proposto em (HANLEY e MCNEIL, 1982) para avaliar os valores apresentados com 10 e 64 variáveis, resultou em um valor de $p = 0,000082 < 0,01$. Então, a hipótese nula deve ser rejeitada (isto é, o desempenho com 64 variáveis é superior ao desempenho com 10 variáveis, ao nível de significância de 1%).

Tabela 37 – Resultados médios para acurácia, AUC, sensibilidade e especificidade, para os métodos de treinamento MSE, parada antecipada e regularização, combinadas as técnicas de extração de características SEC, PCA e LDA, para as três regiões (Apuí, Humaitá e Lábrea)

Região	Métodos de parada do treinamento	N. de Caract.	Técnica de Seleção de Características											
			SEC				PCA				LDA			
			Acur.	AUC	Sens	Esp	Acur.	AUC	Sens	Esp	Acur.	AUC	Sens	Esp
Geral	MSE	10	93,00	0,946 ± 0,0134	92,90	91,90	96,65	0,966 ± 0,0092	97,40	95,90	95,40	0,954 ± 0,0108	94,90	95,90
		20	93,55	0,936 ± 0,0128	93,00	94,10	98,25	0,982 ± 0,0067	98,40	98,10	96,60	0,966 ± 0,0093	95,90	97,30
		30	90,65	0,906 ± 0,0154	91,30	90,00	96,25	0,970 ± 0,0088	97,20	96,70	97,50	0,975 ± 0,0080	97,40	97,60
		40	95,60	0,956 ± 0,0106	95,00	96,20	95,75	0,957 ± 0,0104	95,80	95,70	97,60	0,976 ± 0,0078	97,20	98,00
		50	96,25	0,963 ± 0,0098	96,50	96,00	95,50	0,955 ± 0,0107	95,20	95,80	97,75	0,978 ± 0,0076	97,60	97,90
		64	98,85	0,988 ± 0,0054	98,70	99,00	98,85	0,988 ± 0,0054	98,70	99,00	98,85	0,988 ± 0,0054	98,70	99,00
	Parada Antecipada	10	94,55	0,946 ± 0,0118	94,05	95,05	98,05	0,981 ± 0,0070	98,20	97,20	96,38	0,964 ± 0,0096	95,65	97,10
		20	95,78	0,958 ± 0,0104	95,75	95,80	99,10	0,991 ± 0,0048	99,25	98,95	98,25	0,982 ± 0,0062	98,00	98,50
		30	96,63	0,966 ± 0,0093	96,45	96,80	98,75	0,988 ± 0,0056	98,60	98,90	98,23	0,982 ± 0,0067	98,30	98,15
		40	97,13	0,971 ± 0,0085	96,65	97,60	98,80	0,988 ± 0,0055	98,50	99,10	98,75	0,988 ± 0,0056	98,60	98,90
		50	97,28	0,973 ± 0,0083	97,25	97,30	98,00	0,980 ± 0,0071	98,00	98,00	98,75	0,987 ± 0,0056	98,75	98,75
		64	99,30	0,993 ± 0,0042	99,05	99,55	99,30	0,993 ± 0,0042	99,05	99,55	99,30	0,993 ± 0,0042	99,05	99,55
	Regularização	10	92,05	0,920 ± 0,0143	90,70	93,40	96,90	0,969 ± 0,0089	96,50	97,30	96,20	0,962 ± 0,0098	95,30	97,10
		20	94,55	0,946 ± 0,0118	94,30	94,80	97,70	0,977 ± 0,0076	97,20	98,20	97,20	0,972 ± 0,0084	96,80	97,60
		30	94,65	0,947 ± 0,0117	94,10	95,20	97,70	0,977 ± 0,0076	97,70	97,70	97,45	0,974 ± 0,0080	97,60	97,30
		40	95,65	0,957 ± 0,0105	95,50	95,80	96,90	0,969 ± 0,0089	97,10	96,70	97,55	0,975 ± 0,0079	97,70	97,40
		50	96,40	0,964 ± 0,0096	96,60	96,20	96,05	0,961 ± 0,0100	97,00	95,10	98,10	0,981 ± 0,0069	97,90	98,30
		64	98,75	0,988 ± 0,0056	98,70	98,80	98,75	0,988 ± 0,0056	98,70	98,80	98,75	0,988 ± 0,0056	98,70	98,80

Tabela 38 – Valores para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para as técnicas SEC, PCA e LDA, para as três regiões (Apuí, Humaitá e Lábrea).

Métodos de parada do treinamento	N. de Características	Teste de Significância Estatística		
		SEC-PCA	SEC-LDA	PCA-LDA
MSE	10	0,0908821	0,3082387	0,2875939
	20	0,0007330	0,0288873	0,0045178
	30	0,0001596	0,0000374	0,0000001
	40	0,4732290	0,0641991	0,0627782
	50	0,2899582	0,1102932	0,1253197
Parada Antecipada	10	0,0051347	0,1173583	0,0639081
	20	0,0018811	0,0260114	0,0018382
	30	0,0208890	0,0819046	0,0331181
	40	0,0477398	0,0477398	0,0145013
	50	0,2607232	0,0822143	0,0627642
Regularização	10	0,0017973	0,0077559	0,0007620
	20	0,0134366	0,0360353	0,0111453
	30	0,0155372	0,0285083	0,0077254
	40	0,1909713	0,0855849	0,0656052
	50	0,4140586	0,0751926	0,0807399

Tabela 39 – Valores médios para os testes de significância estatística relativos as diferenças entre as áreas sob a curva ROC, para as técnicas SEC, PCA e LDA, para as três regiões (Apuí, Humaitá e Lábrea).

Métodos de parada do treinamento	Teste de Significância Estatística Médio		
	SEC-PCA	SEC-LDA	PCA-LDA
MSE	0,170992	0,102331	0,096042
Parada Antecipada	0,067274	0,071046	0,035226
Regularização	0,127160	0,046615	0,033196

Tabela 40 – Valores para os testes de significância estatística relativos as diferenças entre as áreas sob a curva ROC, para os métodos MSE, Parada Antecipada e Regularização, para as três regiões (Apuí, Humaitá e Lábrea).

Técnica de Seleção de Característica	N. de Caract.	Teste de Significância Estatística		
		MSE-PAN	MSE-REG	PAN-REG
SEC	10	0,5000000	0,0799526	0,0799526
	20	0,0905163	0,2822894	0,2215781
	30	0,0004485	0,0171449	0,1009951
	40	0,1355812	0,4732290	0,1505206
	50	0,2165429	0,4707578	0,2384926
PCA	10	0,0980169	0,4077415	0,1433866
	20	0,1380236	0,3118707	0,0599366
	30	0,0405886	0,2730975	0,1213183
	40	0,0043837	0,1909713	0,0344528
	50	0,0259491	0,3408584	0,0603943
LDA	10	0,2445069	0,2963327	0,4420683
	20	0,0819046	0,3379523	0,1772768
	30	0,2512754	0,4692710	0,2243958
	40	0,1044386	0,4708841	0,0896858
	50	0,1695206	0,3974469	0,2524440

Tabela 41 – Valores médios para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para os métodos MSE, Parada Antecipada e Regularização, para as três regiões (Apuí, Humaitá e Lábrea).

Técnica de Seleção de Característica	Teste de Significância Estatística Médio		
	MSE-PAN	MSE-REG	PAN-REG
SEC	0,1886178	0,2646747	0,1583078
PCA	0,0613924	0,3049079	0,0838977
LDA	0,1703292	0,3943774	0,2371741

5.2.6 Avaliação dos resultados relativos a utilização das três regiões conjuntamente

Em relação ao desempenho da rede neural utilizando os dados das três regiões (Apuí, Humaitá e Lábrea) juntas foi observada uma piora no desempenho, em relação à AUC, em comparação com o desempenho da rede neural utilizando dados das regiões separadamente. Por exemplo, para a técnica LDA, método parada antecipada e um número de características iguais a 10, a AUC e desvio padrão obtidos para as regiões de Apuí, Humaitá, Lábrea e as três regiões juntas, foram, respectivamente, $0,999 \pm 0,0014$, $0,996 \pm 0,0030$, $0,997 \pm 0,0027$, $0,964 \pm 0,0096$. Aplicando-se o teste de significância estatística proposto em (HANLEY e MCNEIL, 1982) para avaliar os valores apresentados da rede utilizando dados das três regiões juntas e utilizando dados das três regiões separadamente, resultou, respectivamente em $p = 0,000753 < 0,01$, $p = 0,000155 < 0,01$ e $p = 0,00046 < 0,01$. Então, a hipótese nula deve ser rejeitada (isto é, o desempenho do método utilizando a técnica LDA, 10 características, método parada antecipada e dados das regiões em separado é superior ao do método utilizando dados das três regiões juntas, ao nível de significância de 1%). Por outro lado, quando o número de características utilizado é superior a 10, observa-se pouca melhora no desempenho. Por exemplo, para a técnica LDA, método parada antecipada e um número de variáveis iguais a 20, a AUC e desvio padrão obtidos para as regiões de Apuí, Humaitá, Lábrea e as três regiões juntas, foram, respectivamente, $0,999 \pm 0,0016$, $0,997 \pm 0,0029$, $0,998 \pm 0,0025$, $0,982 \pm 0,0067$. Aplicando-se o teste de significância estatística proposto em (HANLEY e MCNEIL, 1982) para avaliar os valores apresentados da rede utilizando dados das três regiões juntas e utilizando dados das três regiões separadamente, resultou, respectivamente em $p = 0,007 < 0,01$, $p = 0,02 > 0,01$ e $p = 0,012 > 0,01$. Então, a hipótese nula não deve ser rejeitada (isto é, o desempenho do método utilizando a técnica LDA, método parada antecipada, 20 características e dados das regiões em separado não é superior ao do método utilizando dados das três regiões, ao nível de significância de 1%).

Com respeito aos métodos de parada da rede neural, foi observada uma leve superioridade no desempenho do método parada antecipada frente as outras duas, quando utilizado dados das três regiões (Apuí, Humaitá e Lábrea) juntas, em relação à AUC, da mesma maneira como ocorreu quando os dados das regiões foram utilizados separadamente. Por exemplo, quando utilizada a técnica PCA e 50 características, a AUC e o desvio padrão obtidos para os métodos MSE, parada antecipada e regularização foram, respectivamente, $0,955 \pm 0,0107$, $0,980 \pm 0,0071$ e $0,961 \pm 0,0100$. Aplicando-se o teste de significância estatística proposto em (HANLEY e MCNEIL, 1982) para avaliar os valores apresentados pelos métodos de parada antecipada e MSE, resultou em um valor de $p = 0,026 > 0,01$. Então, a hipótese nula deve não ser rejeitada (isto é, o desempenho do método parada antecipada não é superior ao do método MSE, ao nível de significância de 1%). Aplicando-se o teste de significância estatística proposto em (HANLEY e MCNEIL, 1982) para avaliar os valores apresentados pelos métodos de parada antecipada e regularização, resultou em um valor de $p = 0,06 > 0,01$. Então, a hipótese nula não deve ser rejeitada (isto é, o desempenho do método parada antecipada não é superior ao do método regularização, ao nível de significância de 1%).

5.2.7 Região analisada por Reis (2014)

A Tabela 42 apresenta os resultados obtidos para a classificação feita com dados da mesma região analisada por Reis (2014), região do município de Belterra, no estado do Pará. Utilizando todo o conjunto de 64 características, o melhor desempenho foi obtido utilizando-se o método de parada MSE, com uma acurácia de 97,20%. Quando aplicado as técnicas de extração de características, o melhor desempenho foi obtido utilizando-se o método de parada regularização e a técnica LDA e um número de características iguais a 50, com uma acurácia de 97,16%. Os resultados foram superiores ao obtido pelo referido autor, que foi uma acurácia de 84,2%.

As Tabela 43 e 44 apresentam valores para os testes de significância estatística para avaliar as diferenças entre os valores para AUC, para os métodos SEC, LDA e PCA. A partir dessas tabelas, observa-se que comparando as técnicas SEC e PCA, a técnica SEC foi estatisticamente superior em relação a técnica PCA apenas quando utilizado um número de características iguais a 20, 30 ou 40 e o método de parada MSE. Comparando as técnicas SEC e LDA, os valores dos testes de significância estatística apresentados nas Tabela 43 e 44 mostram que nenhuma técnica foi superior a outra. Por exemplo, comparando as técnicas SEC e LDA quando utilizado método de parada regularização, a Tabela 44 mostra que o teste de

significância estatística resultou em um valor de p médio de $0,399235 > 0,01$. Então, a hipótese nula não deve ser rejeitada (isto é, o desempenho da técnica LDA não é superior a técnica SEC, ao nível de significância de 1%).

Com respeito aos métodos de parada da rede neural, observa-se na Tabela 42, uma pequena superioridade no desempenho do método parada antecipada frente os outros métodos de parada da rede neural, em relação a AUC. Porém os testes de significância estatística apresentados nas Tabela 45 e 46 mostraram que nenhum método de parada da rede neural foi superior ao outro. Por exemplo, comparando os métodos MSE e parada antecipada quando utilizado a técnica SEC, a Tabela 46 mostra que o teste de significância estatística resultou em um valor de p médio de $0,4060146 > 0,01$. Então, a hipótese nula não deve ser rejeitada (isto é, o desempenho do método parada antecipada não é superior ao do método MSE, ao nível de significância de 1%).

Tabela 42 – Resultados médios para acurácia, AUC, sensibilidade e especificidade, para os métodos de treinamento MSE, parada antecipada e regularização, combinadas as técnicas de extração de características SEC, PCA e LDA, para a região analisada por Reis (2014)

Região	Métodos de parada do treinamento	N. de Caract.	Técnica de Seleção de Características											
			SEC				PCA				LDA			
			Acur.	AUC	Sens	Esp	Acur.	AUC	Sens	Esp	Acur.	AUC	Sens	Esp
Geral	MSE	10	95,35	0,954 ± 0,0109	95,45	95,26	91,42	0,914 ± 0,0148	91,47	91,37	96,69	0,967 ± 0,0104	96,62	96,75
		20	96,28	0,963 ± 0,0097	96,23	96,33	91,13	0,911 ± 0,0151	90,69	91,56	96,80	0,968 ± 0,0090	96,80	96,80
		30	96,62	0,966 ± 0,0093	96,52	96,71	91,47	0,915 ± 0,0148	92,15	90,78	96,88	0,969 ± 0,0089	96,90	96,85
		40	96,96	0,970 ± 0,0088	97,01	96,91	90,06	0,901 ± 0,0159	90,98	89,13	97,04	0,970 ± 0,0088	97,50	97,03
		50	97,15	0,972 ± 0,0085	97,20	96,10	90,73	0,973 ± 0,0084	90,69	90,76	97,15	0,971 ± 0,0086	97,16	97,14
		64	97,20	0,972 ± 0,0084	97,20	97,20	97,20	0,972 ± 0,0084	97,20	97,20	97,20	0,972 ± 0,0084	97,20	97,20
	Parada Antecipada	10	95,79	0,958 ± 0,0104	95,89	95,69	93,63	0,936 ± 0,0128	93,41	93,85	96,68	0,966 ± 0,0091	96,61	96,75
		20	96,88	0,969 ± 0,0089	97,01	97,76	94,26	0,943 ± 0,0121	94,58	93,94	96,78	0,967 ± 0,0089	96,80	96,76
		30	96,96	0,970 ± 0,0088	96,96	96,96	95,11	0,951 ± 0,0112	95,35	94,87	96,78	0,967 ± 0,0089	96,70	96,85
		40	97,13	0,971 ± 0,0085	97,15	97,10	94,19	0,942 ± 0,0122	94,96	93,41	97,02	0,970 ± 0,0088	97,01	97,03
		50	97,10	0,971 ± 0,0086	97,05	97,15	93,63	0,936 ± 0,0128	93,21	94,04	97,08	0,970 ± 0,0088	97,11	97,05
		64	97,18	0,972 ± 0,0085	97,20	97,15	97,18	0,972 ± 0,0085	97,20	97,15	97,18	0,972 ± 0,0085	97,20	97,15
	Regularização	10	95,40	0,954 ± 0,0108	95,45	95,35	91,27	0,913 ± 0,0149	90,01	92,53	96,77	0,967 ± 0,0104	96,79	96,75
		20	96,42	0,964 ± 0,0095	96,81	96,03	94,26	0,943 ± 0,0121	94,58	93,94	96,80	0,968 ± 0,0090	96,80	96,80
		30	96,81	0,968 ± 0,0090	96,91	96,71	95,11	0,951 ± 0,0112	95,35	94,87	96,88	0,969 ± 0,0089	96,90	96,85
		40	97,10	0,971 ± 0,0086	97,01	97,20	94,19	0,942 ± 0,0122	94,96	93,41	97,04	0,970 ± 0,0088	97,05	97,03
		50	97,10	0,971 ± 0,0086	97,01	97,10	93,63	0,936 ± 0,0128	93,21	94,04	97,16	0,971 ± 0,0086	97,17	97,14
		64	97,20	0,972 ± 0,0084	97,20	97,20	97,20	0,972 ± 0,0084	97,20	97,20	97,20	0,972 ± 0,0084	97,20	97,20

Tabela 43 – Valores para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para as técnicas SEC, PCA e LDA, para a região analisada por Reis (2014).

Métodos de parada do treinamento	N. de Características	Teste de Significância Estatística		
		SEC-PCA	SEC-LDA	PCA-LDA
MSE	10	0,0146639	0,1797079	0,2278026
	20	0,0018641	0,3529404	0,3879724
	30	0,0017117	0,4077415	0,4307782
	40	0,0000709	0,5000000	0,5000000
	50	0,4662819	0,4668820	0,4665861
Parada Antecipada	10	0,0905163	0,2826066	0,3064370
	20	0,0412916	0,4377039	0,4474803
	30	0,0902194	0,4063036	0,4177875
	40	0,0257490	0,4674514	0,4733848
	50	0,0115237	0,4674514	0,4742523
Regularização	10	0,0130351	0,1797079	0,2287397
	20	0,0863402	0,3805081	0,3952937
	30	0,1183174	0,4685078	0,4720769
	40	0,0257490	0,4674514	0,4733848
	50	0,0115237	0,5000000	0,5000000

Tabela 44 – Valores médios para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para as técnicas SEC, PCA e LDA, para a região analisada por Reis (2014).

Métodos de parada do treinamento	Teste de Significância Estatística Médio		
	SEC-PCA	SEC-LDA	PCA-LDA
MSE	0,096918	0,381454	0,402628
Parada Antecipada	0,051860	0,412303	0,423868
Regularização	0,050993	0,399235	0,413899

Tabela 45 – Valores para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para os métodos MSE, Parada Antecipada e Regularização, para a região analisada por Reis (2014).

Técnica de Seleção de Característica	N. de Caract.	Teste de Significância Estatística		
		MSE-PAN	MSE-REG	PAN-REG
SEC	10	0,3947038	0,5000000	0,3947038
	20	0,3241176	0,4707578	0,3508482
	30	0,3769182	0,4386552	0,4367073
	40	0,4674514	0,4674514	0,5000000
	50	0,4668820	0,4668820	0,5000000
PCA	10	0,1305323	0,4810291	0,1207809
	20	0,0487322	0,0487322	0,5000000
	30	0,0258107	0,0258107	0,5000000
	40	0,0202633	0,0202633	0,5000000
	50	0,0075729	0,0075729	0,5000000
LDA	10	0,4694682	0,5000000	0,4694682
	20	0,4689991	0,5000000	0,4689991
	30	0,4377039	0,5000000	0,4377039
	40	0,5000000	0,5000000	0,5000000
	50	0,4674514	0,5000000	0,4674514

Tabela 46 – Valores médios para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para os métodos MSE, Parada Antecipada e Regularização, para a região analisada por Reis (2014).

Técnica de Seleção de Característica	Teste de Significância Estatística Médio		
	MSE-PAN	MSE-REG	PAN-REG
SEC	0,4060146	0,4687493	0,4364519
PCA	0,0465823	0,1166817	0,4241562
LDA	0,4687245	0,5000000	0,4687245

5.2.8 Região analisada por Pimentel (2014)

A Tabela 47 apresenta os resultados obtidos para a classificação feita com dados da mesma região analisada por Pimentel (2014), região situada ao leste do Mato Grosso. O melhor desempenho foi obtido utilizando-se o método de regularização associado à técnica LDA e com um número de características iguais a 50, com uma acurácia média de 97,19%. Esse resultado foi ligeiramente superior em relação àquele obtido pelo referido autor, uma acurácia de 96,05%. O autor trabalhou com 1300 amostras e utilizou redes neurais como classificador.

Com respeito as técnicas seleção de características, observa-se na Tabela 47 que a técnica LDA apresentou desempenho levemente superior com respeito a AUC em relação a SEC e PCA. As Tabela 48 e 49 apresentam valores para os testes de significância estatística utilizados para avaliar as diferenças entre os valores para AUC, para os métodos SEC, LDA e PCA. Comparando as técnicas SEC e PCA, quando utilizado o método de parada MSE e o número de características iguais 30, 40 ou 50, os teste de significância estatística apresentados na Tabela 48 mostram que a técnica SEC foi superior a técnica PCA. Por outro lado, comparando as técnicas SEC e PCA, quando utilizado o método de parada MSE e um número de características iguais a 10 ou 20, os valores dos testes de significância estatística apresentados na Tabela 48 mostram que a técnica SEC não foi superior a PCA. Comparando as técnicas LDA e SEC, os valores dos testes de significância estatística apresentados na Tabela 48 mostram que a técnica LDA não foi superior a técnica SEC independentemente do método de parada (MSE, Parada Antecipada ou Regularização). Por fim, comparando as técnicas LDA e PCA os valores dos testes de significância estatística apresentados na Tabela 48 mostram que a técnica LDA não foi superior a técnica PCA independentemente do método de parada (MSE, Parada Antecipada ou Regularização).

Com respeito aos métodos de parada de treinamento da rede neural, observa-se na Tabela 47, uma pequena superioridade no desempenho do método parada antecipada frente a outros métodos de parada (MSE e Regularização), em relação à AUC. Porém os testes de

significância estatística apresentados nas Tabela 50 e 51 mostraram que nenhum método de parada foi superior ao outro. Por exemplo, comparando os métodos de parada regularização e parada antecipada quando utilizado a técnica PCA, a Tabela 51 mostra que o teste de significância estatística resultou em um valor de p médio de $0,0838977 > 0,01$. Então, a hipótese nula não deve ser rejeitada (isto é, o desempenho do método parada antecipada não é superior ao do método regularização, ao nível de significância de 1%).

Tabela 47 – Resultados médios para acurácia, AUC, sensibilidade e especificidade, para os métodos de treinamento MSE, parada antecipada e regularização, combinadas as técnicas de extração de características SEC, PCA e LDA, para a região analisada por Pimentel (2014)

Região	Métodos de parada do treinamento	N. de Caract.	Técnica de Seleção de Características											
			SEC				PCA				LDA			
			Acur.	AUC	Sens	Esp	Acur.	AUC	Sens	Esp	Acur.	AUC	Sens	Esp
Geral	MSE	10	95,30	0,953 ± 0,0109	95,06	95,55	91,76	0,918 ± 0,0145	92,05	91,47	96,63	0,963 ± 0,0104	96,50	96,75
		20	96,96	0,970 ± 0,0088	96,91	97,01	94,14	0,941 ± 0,0122	93,80	94,48	96,80	0,968 ± 0,0090	96,80	96,80
		30	96,96	0,970 ± 0,0088	96,81	97,10	92,29	0,923 ± 0,0140	92,53	92,05	96,88	0,968 ± 0,0089	96,90	96,85
		40	97,05	0,971 ± 0,0087	97,01	97,10	89,18	0,892 ± 0,0166	88,94	89,42	97,04	0,970 ± 0,0088	97,05	97,03
		50	97,01	0,970 ± 0,0087	96,91	97,10	83,88	0,839 ± 0,0201	93,40	84,37	97,11	0,971 ± 0,0086	97,08	97,14
		64	97,15	0,972 ± 0,0085	97,10	97,20	97,15	0,972 ± 0,0085	97,10	97,20	97,15	0,972 ± 0,0085	97,10	97,20
	Parada Antecipada	10	96,16	0,962 ± 0,0099	96,62	96,69	93,00	0,930 ± 0,0134	92,73	93,26	96,68	0,966 ± 0,0091	96,61	96,75
		20	97,15	0,972 ± 0,0085	97,15	97,15	95,00	0,950 ± 0,0113	95,16	94,92	96,78	0,967 ± 0,0089	96,80	96,76
		30	97,08	0,971 ± 0,0086	97,05	97,10	93,77	0,938 ± 0,0126	93,99	93,56	96,78	0,967 ± 0,0089	96,70	96,85
		40	97,10	0,971 ± 0,0086	97,05	97,15	93,41	0,934 ± 0,0130	93,99	92,83	97,02	0,970 ± 0,0088	97,01	97,03
		50	97,10	0,971 ± 0,0086	97,05	97,15	93,56	0,936 ± 0,0128	93,07	94,04	97,07	0,970 ± 0,0088	97,08	97,05
		64	97,18	0,972 ± 0,0085	97,15	97,20	97,18	0,972 ± 0,0085	97,15	97,20	97,18	0,972 ± 0,0085	97,15	97,20
	Regularização	10	95,11	0,951 ± 0,0112	95,06	95,16	91,27	0,913 ± 0,0149	90,01	95,52	96,76	0,967 ± 0,0104	96,79	96,72
		20	96,81	0,968 ± 0,0090	96,71	96,91	94,25	0,943 ± 0,0121	94,57	93,93	96,75	0,967 ± 0,0090	96,80	96,70
		30	97,01	0,970 ± 0,0087	97,01	97,01	95,11	0,951 ± 0,0112	95,35	94,87	96,83	0,968 ± 0,0089	96,90	96,75
		40	97,15	0,972 ± 0,0085	97,10	97,20	94,19	0,942 ± 0,0122	96,96	93,41	97,07	0,977 ± 0,0088	97,04	97,09
		50	97,01	0,970 ± 0,0087	96,81	97,20	93,62	0,936 ± 0,0128	93,20	94,03	97,19	0,971 ± 0,0086	97,19	97,18
		64	97,10	0,971 ± 0,0086	97,01	97,20	97,10	0,971 ± 0,0086	97,01	97,20	97,10	0,971 ± 0,0077	97,01	97,20

Tabela 48 – Valores para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para as técnicas SEC, PCA e LDA, para a região analisada por Pimentel (2014).

Métodos de parada do treinamento	N. de Características	Teste de Significância Estatística		
		SEC-PCA	SEC-LDA	PCA-LDA
MSE	10	0,0268942	0,2471367	0,2830388
	20	0,0271036	0,4367073	0,4477704
	30	0,0022232	0,4367073	0,4522708
	40	0,0000118	0,4674514	0,4787256
	50	0,0000001	0,4674514	0,4817369
Parada Antecipada	10	0,0269592	0,3838030	0,4030146
	20	0,0592833	0,3439966	0,3654905
	30	0,0151428	0,3749984	0,3985999
	40	0,0087278	0,4674514	0,4745229
	50	0,0115237	0,4674514	0,4742523
Regularização	10	0,0207088	0,1341533	0,1802360
	20	0,0484901	0,4689991	0,4736830
	30	0,0902194	0,4367073	0,4446350
	40	0,0213730	0,3301509	0,3639207
	50	0,0140499	0,4674514	0,4741138

Tabela 49 – Valores médios para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para as técnicas SEC, PCA e LDA, para a região analisada por Pimentel (2014).

Métodos de parada do treinamento	Teste de Significância Estatística Médio		
	SEC-PCA	SEC-LDA	PCA-LDA
MSE	0,011247	0,411091	0,428709
Parada Antecipada	0,024327	0,407540	0,423176
Regularização	0,038968	0,367492	0,387318

Tabela 50 – Valores para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para os métodos MSE, Parada Antecipada e Regularização, para a região analisada por Pimentel (2014).

Técnica de Seleção de Característica	N. de Caract.	Teste de Significância Estatística		
		MSE-PAN	MSE-REG	PAN-REG
SEC	10	0,2704278	0,4491487	0,2300441
	20	0,4345627	0,4367073	0,3729866
	30	0,4674514	0,5000000	0,4674514
	40	0,5000000	0,4668820	0,4668820
	50	0,4674514	0,5000000	0,4674514
PCA	10	0,2712936	0,4049168	0,1979844
	20	0,2947562	0,4537461	0,3359394
	30	0,2130296	0,0592675	0,2199819
	40	0,0230833	0,0075386	0,3265347
	50	0,0000231	0,0000231	0,5000000
LDA	10	0,4116901	0,3806130	0,4694682
	20	0,4689991	0,4689991	0,5000000
	30	0,4689991	0,5000000	0,4689991
	40	0,5000000	0,2425167	0,2730975
	50	0,4674514	0,5000000	0,4674514

Tabela 51 – Valores médios para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para os métodos MSE, Parada Antecipada e Regularização, para a região analisada por Pimentel (2014).

Técnica de Seleção de Característica	Teste de Significância Estatística Médio		
	MSE-PAN	MSE-REG	PAN-REG
SEC	0,4279787	0,4705476	0,4009631
PCA	0,1604372	0,1850984	0,3160881
LDA	0,4634279	0,4184258	0,4358032

5.2.9 Região analisada por Zanotta (2014)

A Tabela 52 apresenta os resultados obtidos para a classificação feita com dados da mesma região utilizada por Zanotta (2014), região situada no centro do estado do Pará. Quando aplicado as técnicas de extração de características, o melhor desempenho foi obtido utilizando-se o método parada antecipada, a técnica SEC e um número de características iguais a 50, com uma acurácia de 97,18%. Esse resultado foi bem superior àquele obtido pelo referido autor, uma acurácia de 82%. O autor trabalhou com apenas 3 características e com o classificador baseado na análise do vetor de mudanças.

Com respeito as técnicas seleção de características, observa-se na Tabela 52 que as técnicas SEC e LDA apresentaram melhores desempenhos em relação a técnica PCA com respeito a área AUC. As Tabela 53 e 54 apresentam valores para os testes de significância estatística utilizados para avaliar as diferenças entre os valores para AUC, para os métodos SEC, LDA e PCA. Comparando as técnicas SEC e PCA, quando utilizado o método de parada MSE, o teste de significância estatística apresentado na Tabela 54 resultou em um valor de p médio de $0,000587 < 0,001$. Então, a hipótese nula deve ser rejeitada (isto é, o desempenho da técnica SEC é superior a técnica PCA, ao nível de significância de 0,1%). Comparando as técnicas LDA e SEC, os valores dos testes de significância estatística apresentados na Tabela 53 mostram que a técnica LDA não foi superior a técnica SEC independentemente do método de parada (MSE, Parada Antecipada ou Regularização). Por fim, comparando as técnicas LDA e PCA os valores dos testes de significância estatística apresentados na Tabela 53 mostram que a técnica LDA não foi superior a técnica PCA independentemente do método de parada (MSE, Parada Antecipada ou Regularização).

Com respeito aos métodos de parada de treinamento da rede neural, observa-se na Tabela 52, uma pequena superioridade no desempenho do método parada antecipada frente a

outros métodos de parada (MSE e Regularização), em relação à AUC. Porém os testes de significância estatística apresentados nas Tabela 55 e 56 mostraram que nenhum método de parada foi superior ao outro. Por exemplo, comparando os métodos de parada regularização e parada antecipada quando utilizado a técnica PCA, a Tabela 56 mostra que o teste de significância estatística resultou em um valor de p médio de $0,3172750 > 0,01$. Então, a hipótese nula não deve ser rejeitada (isto é, o desempenho do método parada antecipada não é superior ao do método regularização, ao nível de significância de 1%).

Tabela 52 – Resultados médios para acurácia, AUC, sensibilidade e especificidade, para os métodos de treinamento MSE, parada antecipada e regularização, combinadas as técnicas de extração de características SEC, PCA e LDA, para a região utilizada por Zanotta (2014)

Região	Métodos de parada do treinamento	N. de Caract.	Técnica de Seleção de Características											
			SEC				PCA				LDA			
			Acur.	AUC	Sens	Esp	Acur.	AUC	Sens	Esp	Acur.	AUC	Sens	Esp
Geral	MSE	10	95,21	0,952 ± 0,0111	94,77	95,64	88,70	0,887 ± 0,0170	88,16	89,23	96,69	0,967 ± 0,0104	96,62	96,75
		20	95,99	0,960 ± 0,0101	95,90	96,03	90,74	0,907 ± 0,0154	90,30	91,17	96,73	0,979 ± 0,0090	96,70	96,76
		30	95,45	0,955 ± 0,0108	95,64	95,26	88,55	0,885 ± 0,0171	88,94	88,16	96,87	0,969 ± 0,0089	96,89	96,85
		40	95,55	0,955 ± 0,0107	95,35	95,74	86,02	0,860 ± 0,0188	85,92	86,12	97,04	0,970 ± 0,0088	97,05	97,03
		50	96,96	0,970 ± 0,0088	96,81	97,10	83,40	0,834 ± 0,0204	84,66	92,13	97,16	0,971 ± 0,0086	97,16	97,15
		64	97,15	0,972 ± 0,0085	97,10	97,20	97,15	0,972 ± 0,0085	97,10	97,20	97,15	0,972 ± 0,0085	97,10	97,20
	Parada Antecipada	10	95,96	0,960 ± 0,0101	96,03	95,89	90,64	0,906 ± 0,0155	90,20	91,08	96,68	0,966 ± 0,0091	96,61	96,75
		20	96,42	0,964 ± 0,0095	96,52	96,33	94,96	0,950 ± 0,0113	95,21	94,72	96,78	0,967 ± 0,0089	96,80	96,76
		30	96,28	0,963 ± 0,0097	96,42	96,13	93,68	0,937 ± 0,0127	94,19	93,17	96,79	0,967 ± 0,0089	96,72	96,85
		40	96,50	0,965 ± 0,0094	96,52	96,47	93,29	0,933 ± 0,0131	93,51	93,07	97,02	0,970 ± 0,0088	97,01	97,03
		50	97,18	0,972 ± 0,0085	97,15	97,20	94,24	0,942 ± 0,0121	93,99	94,48	97,06	0,970 ± 0,0088	97,06	97,05
		64	97,20	0,972 ± 0,0085	97,15	97,20	97,18	0,972 ± 0,0085	97,15	97,20	97,18	0,972 ± 0,0085	97,15	97,20
	Regularização	10	94,96	0,950 ± 0,0113	94,77	95,16	91,27	0,913 ± 0,0149	90,01	92,53	96,76	0,967 ± 0,0104	96,79	96,72
		20	95,64	0,956 ± 0,0105	96,03	95,26	94,26	0,943 ± 0,0121	94,58	93,94	96,75	0,968 ± 0,0090	96,80	96,70
		30	95,60	0,956 ± 0,0106	95,94	95,26	95,11	0,951 ± 0,0112	95,35	94,87	96,83	0,969 ± 0,0089	96,90	96,75
		40	95,30	0,953 ± 0,0109	95,64	94,96	94,19	0,942 ± 0,0122	94,96	93,41	94,04	0,970 ± 0,0088	97,05	97,03
		50	97,15	0,972 ± 0,0085	97,10	97,20	93,63	0,936 ± 0,0128	93,21	94,04	97,16	0,971 ± 0,0086	97,17	97,14
		64	97,15	0,972 ± 0,0085	97,10	97,20	97,15	0,972 ± 0,0085	97,10	97,20	97,15	0,972 ± 0,0085	97,10	97,20

Tabela 53 – Valores para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para as técnicas SEC, PCA e LDA, para a região analisada por Zanotta (2014).

Métodos de parada do treinamento	N. de Características	Teste de Significância Estatística		
		SEC-PCA	SEC-LDA	PCA-LDA
MSE	10	0,0006630	0,1481904	0,2181971
	20	0,0020050	0,0635478	0,1325269
	30	0,0002614	0,1571090	0,2337440
	40	0,0000057	0,1388486	0,2346637
	50	0,0000001	0,4674514	0,4819543
Parada Antecipada	10	0,0017456	0,3309910	0,3699070
	20	0,1721642	0,4104304	0,4182813
	30	0,0517842	0,3821899	0,3991209
	40	0,0236487	0,3486649	0,3752921
	50	0,0213730	0,4345627	0,4468878
Regularização	10	0,0239444	0,1212038	0,1656197
	20	0,2090229	0,1941646	0,2128293
	30	0,3727106	0,1734063	0,1812385
	40	0,2508032	0,1123708	0,1281894
	50	0,0093791	0,4668820	0,4741138

Tabela 54 – Valores médios para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para as técnicas SEC, PCA e LDA, para a região analisada por Zanotta (2014).

Métodos de parada do treinamento	Teste de Significância Estatística Médio		
	SEC-PCA	SEC-LDA	PCA-LDA
MSE	0,000587	0,195029	0,260217
Parada Antecipada	0,054143	0,381368	0,401898
Regularização	0,173172	0,213605	0,232398

Tabela 55 – Valores para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para os métodos MSE, Parada Antecipada e Regularização, para a região analisada por Zanotta (2014).

Técnica de Seleção de Característica	N. de Caract.	Teste de Significância Estatística		
		MSE-PAN	MSE-REG	PAN-REG
SEC	10	0,2966165	0,4496713	0,2545702
	20	0,3868409	0,3922670	0,2875903
	30	0,2899582	0,4735334	0,3130142
	40	0,2417947	0,4480530	0,2031997
	50	0,4345627	0,4345627	0,5000000
PCA	10	0,2040355	0,1247440	0,3723563
	20	0,0121989	0,0329116	0,3359394
	30	0,0073028	0,0006186	0,2038579
	40	0,0007192	0,0001257	0,3072512
	50	0,0000027	0,0000112	0,3669703
LDA	10	0,4694682	0,5000000	0,4694682
	20	0,1527439	0,1670968	0,4689991
	30	0,4377039	0,5000000	0,4377039
	40	0,5000000	0,5000000	0,5000000
	50	0,4674514	0,5000000	0,4674514

Tabela 56 – Valores médios para os testes de significância estatística relativos as diferenças entre as áreas sob a curva de ROC, para os métodos MSE, Parada Antecipada e Regularização, para a região analisada por Zanotta (2014).

Técnica de Seleção de Característica	Teste de Significância Estatística Médio		
	MSE-PAN	MSE-REG	PAN-REG
SEC	0,3299546	0,4396175	0,3116749
PCA	0,0448518	0,0316822	0,3172750
LDA	0,4054735	0,4334194	0,4687245

5.3 Resultados: Mapa de classificação

Com o objetivo de obter uma validação mais completa da metodologia proposta nesse trabalho para detecção de áreas desmatadas, a mesma foi aplicada a todas as áreas de desmatamento e áreas não desmatadas (floresta e não-floresta) das regiões de Apuí, Humaitá e Lábrea. Dessa forma, a seguir são apresentadas, para as três regiões, uma classificação de todas as áreas desmatadas e não desmatadas. Nesse caso foi utilizado um classificador com um número de características foram iguais a 50, juntamente com a técnica de seleção de características LDA e o método de parada da rede neural parada antecipada pois, em geral, apresentaram os melhores resultados de classificação em relação as demais técnicas de seleção de características e critérios de parada da rede neural. As Figuras 22 e 23 apresentam mapas para a região de Apuí, sendo a Figura 23 um *zoom* de uma área da Figura 22. As Figuras 24 e 25 apresentam mapas para a região de Humaitá, sendo a Figura 25 um *zoom* de uma área da Figura 24. As Figuras 26 e 27 apresentam mapas para a região de Lábrea, sendo a Figura 27 um *zoom* de uma área da Figura 26.

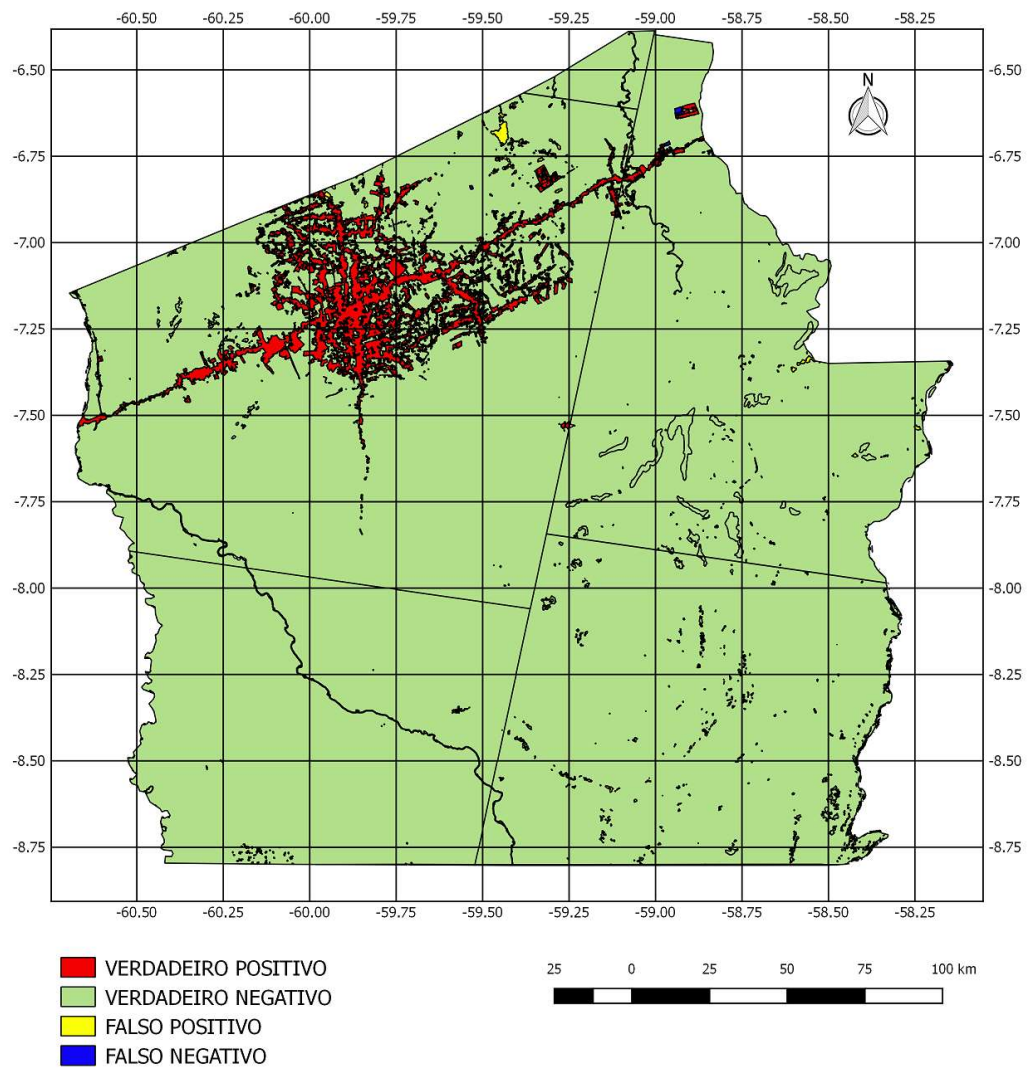


Figura 22 – Mapa de classificação. Apuí

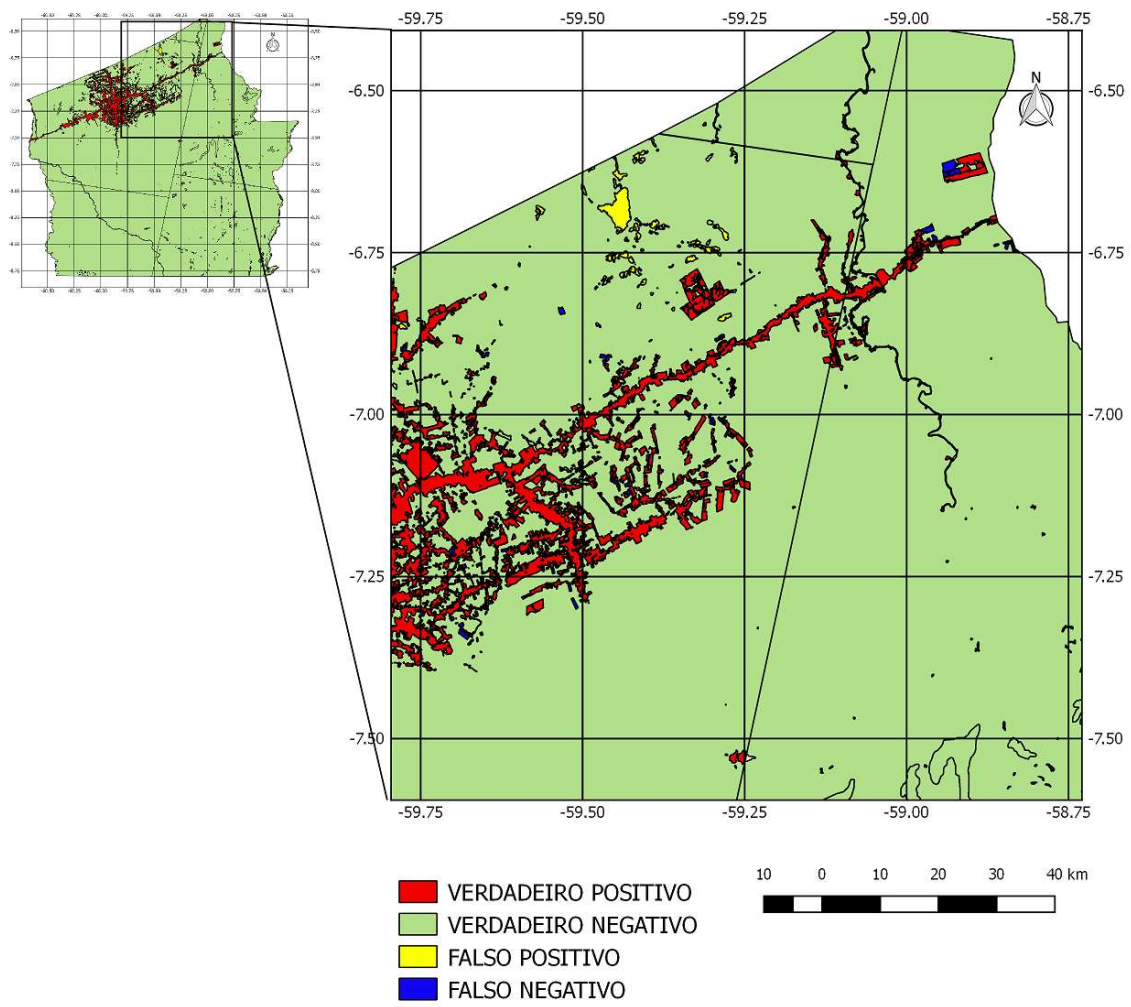


Figura 23 – Mapa classificação. Apuí

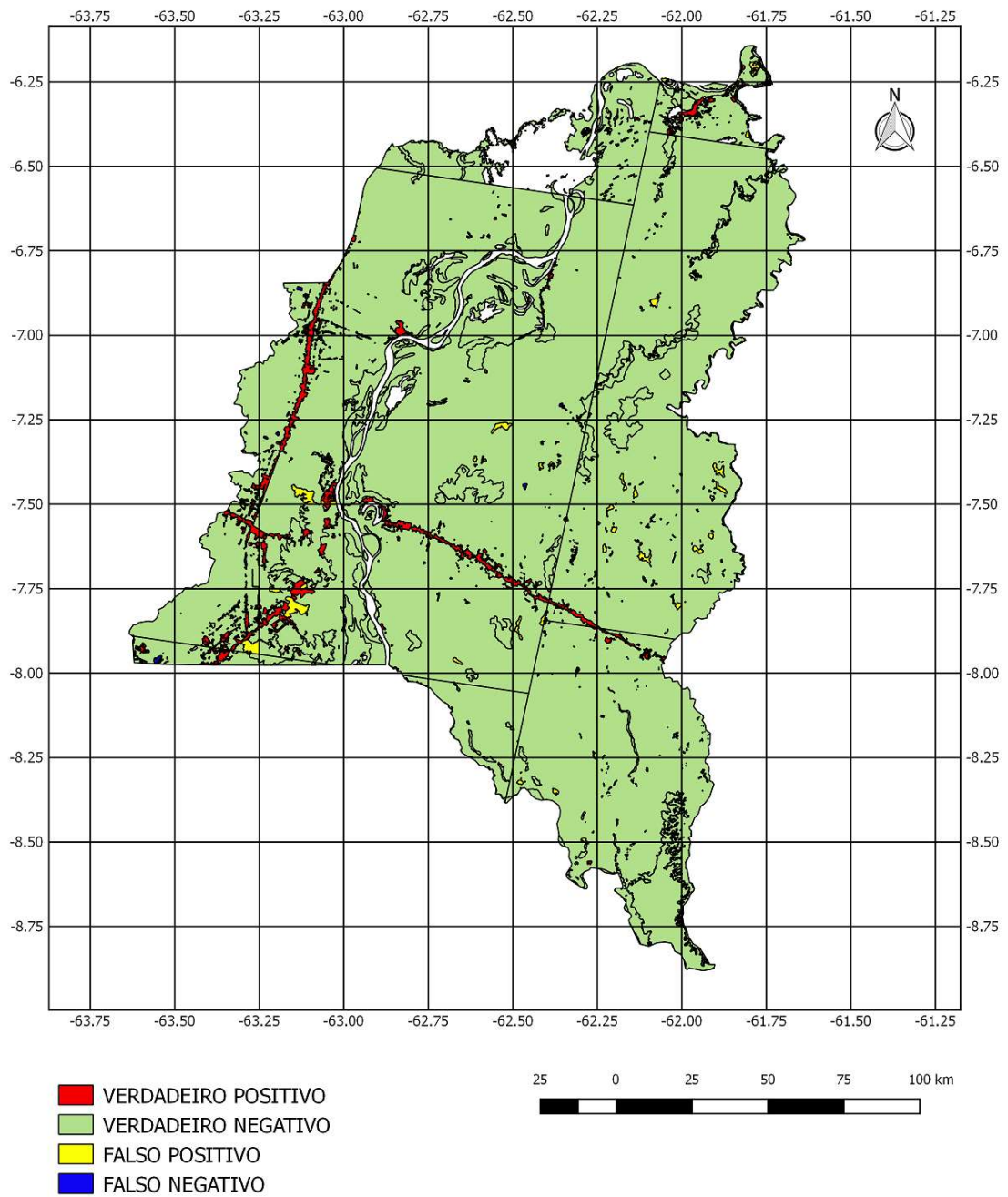


Figura 24 – Mapa classificação Humaitá

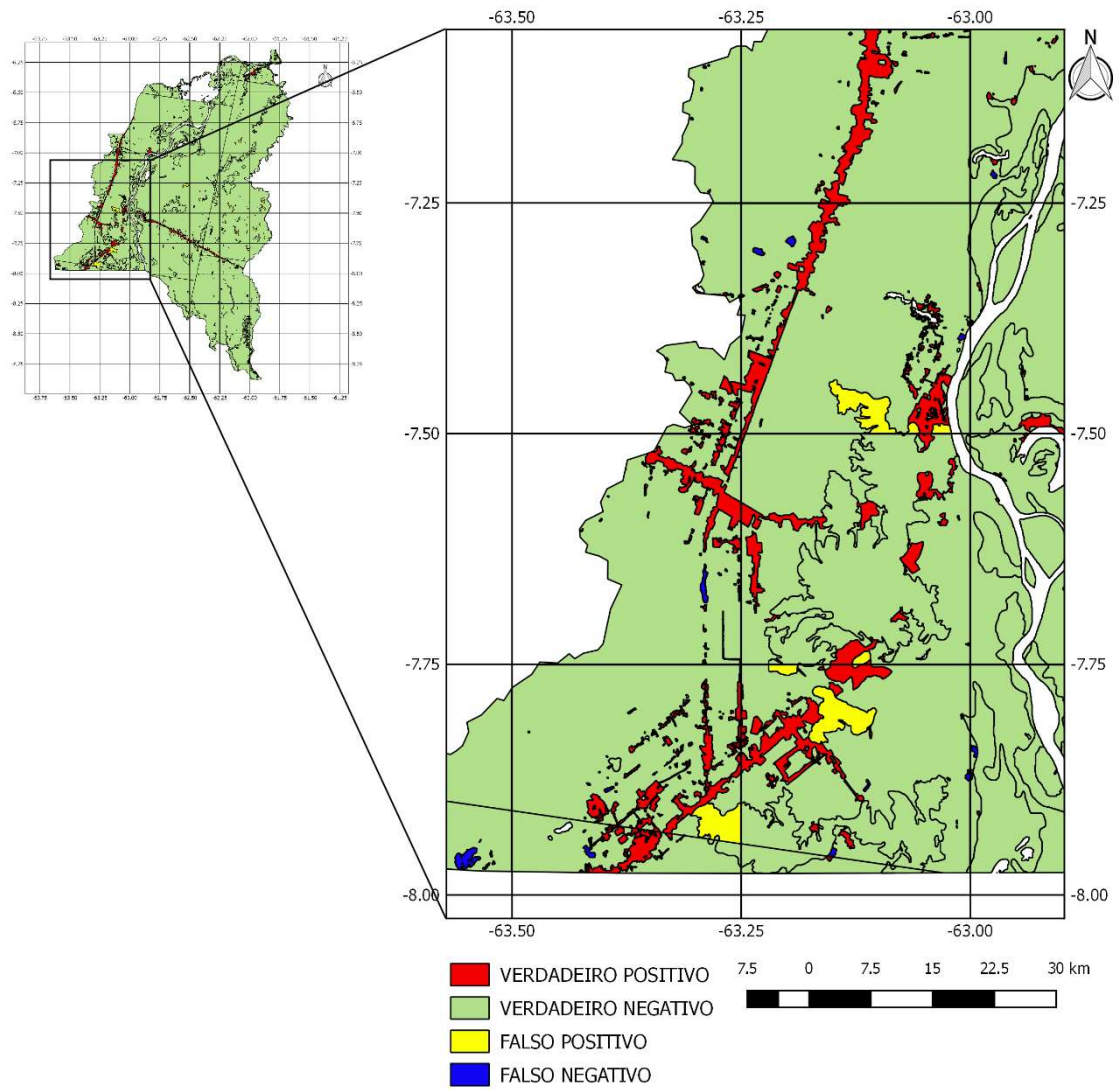


Figura 25 – Mapa classificação Humaitá

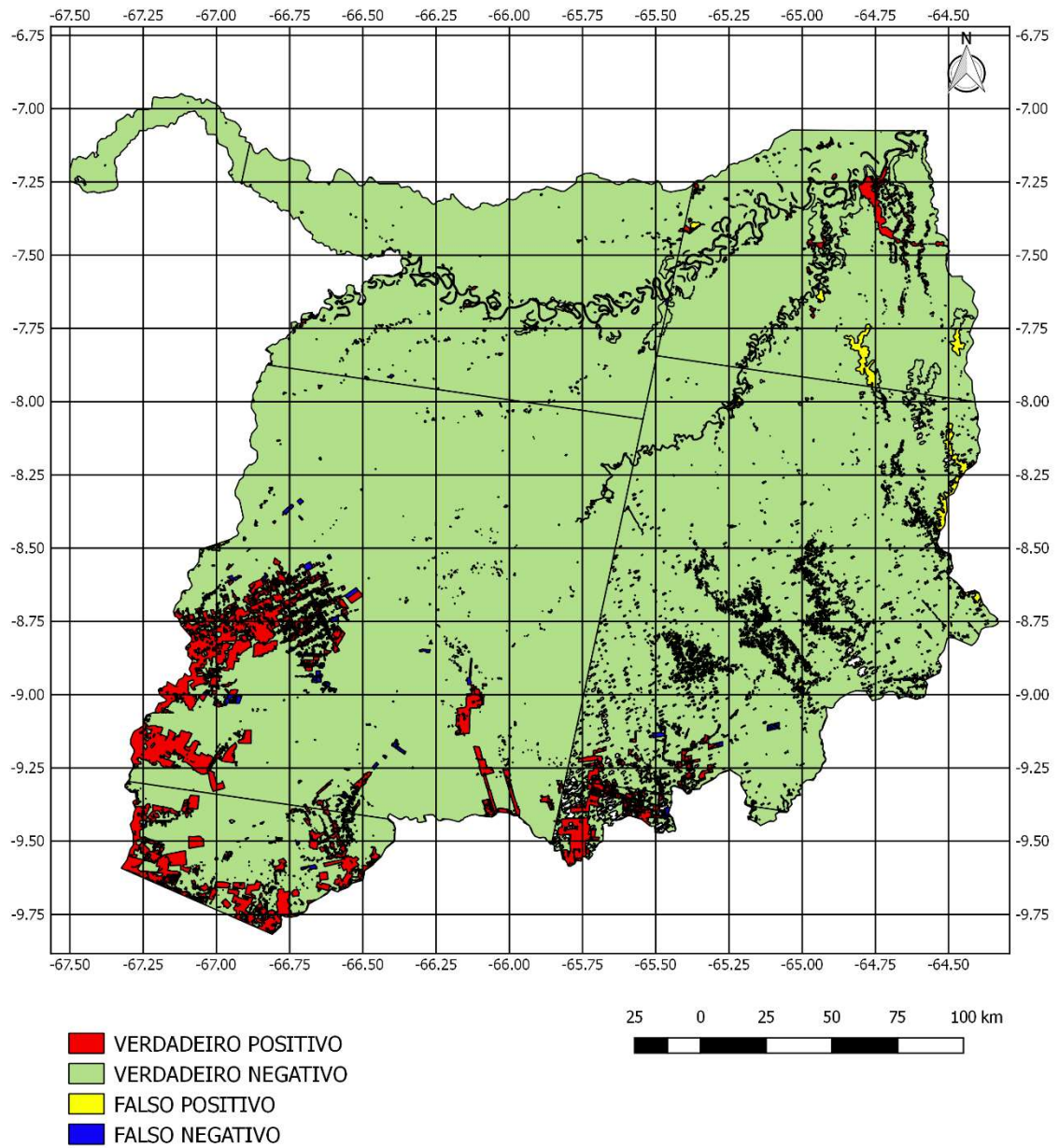


Figura 26 – Mapa classificação Lábrea

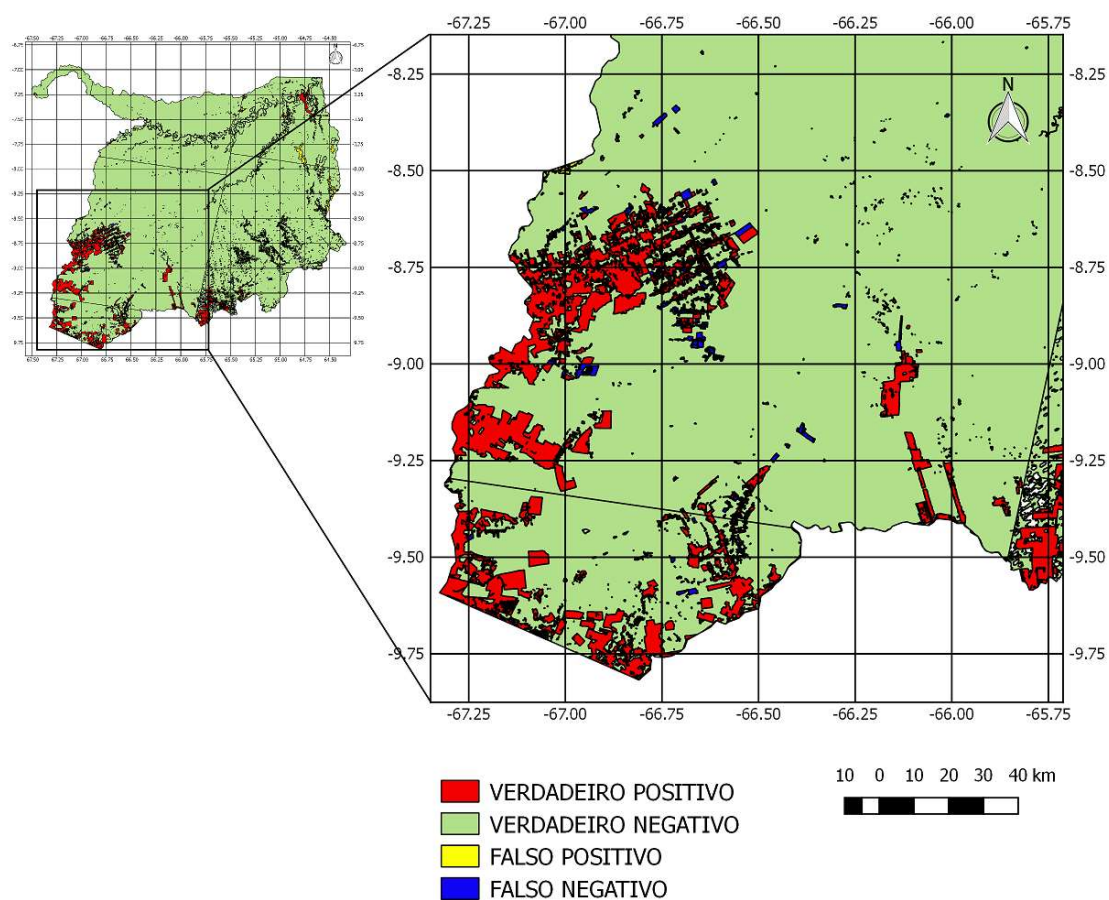


Figura 27 – Mapa classificação Lábrea

A Tabela 57 apresenta os resultados para as métricas sensibilidade, especificidade e coeficientes de Dice e Jaccard para cada um dos mapas das três regiões referentes a Tabela 8, utilizando a Técnica LDA com 50 características e o critério de parada Regularização.

Com respeito as medidas de similaridade Dice e Jaccard apresentadas na Tabela 57, os resultados obtidos foram superiores a 0,9. Isso mostra que a região classificada pelo método proposto é muito similar ao padrão ouro.

Tabela 57 – Resultados de sensibilidade, especificidade e coeficientes de Dice e Jaccard para as três regiões apresentadas na Tabela 8

Região	Sensibilidade (%)	Especificidade (%)	Dice	Jaccard
Apuí	96,08	96,04	0,96	0,94
Humaitá	94,85	95,71	0,94	0,91
Lábrea	96,44	96,02	0,96	0,93

6 CONCLUSÕES E TRABALHOS FUTUROS

Esta dissertação teve como objetivo o reconhecimento de desmatamento em imagens ópticas LANDSAT-8/OLI utilizando redes neurais artificiais. Os municípios de Apuí, Humaitá e Lábrea do estado do Amazonas foram as regiões escolhidas para validação da metodologia de classificação proposta. Para fins de comparação, a metodologia proposta foi comparada com as mesmas regiões analisadas por Reis (2014), Pimentel (2014) e Zanotta (2014). O conjunto de dados de entrada foi constituído de 64 características. Para avaliar a influência da redução de características sobre o desempenho das redes neurais, foram utilizadas três técnicas de seleção de características: SEC, PCA e LDA. Os tamanhos dos conjuntos de características utilizados foram: 50, 40, 30, 20 e 10. Como critérios de parada da rede neural foram adotados o MSE, a regularização e a parada antecipada. Os dois últimos visaram melhorar a capacidade de generalização das redes neurais. Para cada região foram gerados 48 cenários conforme descrito na Tabela 12.

Os principais parâmetros utilizados para avaliação de um método de classificação foram a AUC, a acurácia do método e as medidas de similaridade Dice e Jaccard. Com base nos resultados apresentados no Capítulo 5, com respeito a AUC, foram derivadas algumas conclusões importantes.

Os resultados mostraram a importância de se avaliar o resultado da combinação de novos métodos de generalização com técnicas de seleção de características. A técnica de seleção LDA apresentou um melhor desempenho em relação às técnicas SEC e PCA. O método de parada da rede neural parada antecipada mostrou um desempenho levemente superior em relação aos métodos de parada regularização e MSE. Em relação à técnica de seleção de características LDA, a que apresentou melhores resultados, foi verificado que o número de características utilizadas não possui influência sobre o desempenho do classificador.

A metodologia utilizada nessa dissertação foi comparada com três trabalhos da literatura: Reis (2014), Pimentel (2014) e Zanotta (2014). Na mesma região analisada por Reis (2014), foi obtida uma acurácia de 97,20% quando utilizado 64 características. Quando aplicado as técnicas de extração de características, o melhor desempenho foi obtido utilizando-se o método de parada regularização e a técnica LDA e um número de características iguais a 50, com uma acurácia de 97,16%. Os resultados foram superiores ao obtido pelo referido autor, que obteve uma acurácia de 84,2%. Na mesma região analisada por Zanotta (2014), o melhor desempenho foi obtido quando aplicado o método de parada antecipada, a técnica de extração de características SEC e um número de características iguais a 50. Nesse caso, a acurácia obtida

foi de 97,18%. Esse resultado foi bem superior àquele obtido pelo referido autor, que obteve uma acurácia de 82%. Quanto ao trabalho de Pimentel, os resultados foram bem próximos. A metodologia proposta neste trabalho obteve uma acurácia de 97,19% utilizando o método de regularização associado à técnica LDA e com um número de características iguais a 50. Esse resultado foi ligeiramente superior em relação àquele obtido pelo referido autor, que obteve uma acurácia de 96,05%.

Em relação as medidas de similaridade Dice e Jaccard calculadas para as regiões de Apuí, Humaitá e Lábrea, foram obtidos resultados superiores a 0,9. Dessa forma, o método proposto classificou as regiões de maneira muito similar ao padrão ouro.

Destaca-se, por fim, em função das altas taxas de acurácia e de AUC obtidas, a validade de se utilizar critérios de extração de características juntamente com métodos de generalização de redes neurais para classificação de áreas desmatadas, a partir de imagens ópticas.

Em trabalhos futuros, propõe-se expandir a área geográfica de estudos, selecionando mais regiões do país, e utilizando-se classes de uso do solo mais variadas. O objetivo principal seria investigar se é possível criar um sistema capaz de diferenciar as diversas características das coberturas vegetais presentes no Brasil ou se, para isso, seria necessário criar um sistema especializado para cada região do país. Pretende-se também verificar se as marcações entre a metodologia proposta e de um operador humano são maiores entre as marcações entre dois operadores humanos.

7 REFERÊNCIAS

- ALMEIDA, L. M. Avaliação da gravidade da malária utilizando técnicas de extração de características e redes neurais artificiais. Dissertação de Mestrado do Curso de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Amazonas, Manaus, 2015.
- ARANGO, R. B.; CAMPOS, A. M.; COMBARRO, E. F.; CANAS, E. R.; DÍAZ, I. Mapping cultivable land from satellite imagery with clustering algorithms. In: *International Journal of Applied Earth Observation and Geoinformation*. Volume 49, July 2016, Pages 99–106.
- ARMENTERAS, D.; RUDAS, G; RODRIGUEZ, N; Sua, S; Romerom. Patterns and causes of deforestation in the Colombian Amazon. *Ecological Indicators*, v. 6, p. 353–368, 2006.
- ASSUNÇÃO, E. T. Representação e classificação de texturas da íris baseado na análise discriminante de fisher bi-dimensional. Dissertação de Mestrado do Curso de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Amazonas, Manaus, 2011.
- AUGUSTO-SILVA, P. ; VALÉRIO, L. P. ; SANTOS, T. B. ; ALCÂNTARA, E. H. ; STECH, J. J . Análise de classificadores para mapeamento de uso e cobertura do solo. In: *Anais XVI Simpósio Brasileiro de Sensoriamento Remoto - SBSR*, Foz do Iguaçu, PR, Brasil, 13 a 18 de abril de 2013, INPE.
- BHARATKAR, P.; PATEL, R. . Evaluation of RSI Classification Methods for Effective Land Use Mapping. *International Conference on Communication Systems and Network Technologies*, 2013 IEEE, p.109-113.
- BARBOSA, D. P; NOMA, A.; KÖRTING, T. S.; FONSECA, L. M. G. Um Estudo Experimental com Classificadores baseados em Regiões e Perfis EVI. *Anais XVII Simpósio Brasileiro de Sensoriamento Remoto - SBSR*, João Pessoa-PB, Brasil, 25 a 29 de abril de 2015, INPE
- BECKER, B. K. Geopolítica da Amazônia. *Estudos Avançados*, v. 19, n.53, p. 71-86,2005.
- BRAGA, A DE PADUA; LUDEMIR, T BERNARDA; CARVALHO, ANDRE C. P. DE L. FERREIRA. *Redes Neurais Artificiais – Teoria e aplicações*. Rio de Janeiro – Editora LTC – Livros Técnicos e Científicos – 2000
- CÂMARA, G.; VALERIANO, D. D. M.; SOARES, J. V. Metodologia para o Cálculo da Taxa Anual de Desmatamento na Amazônia Legal. São José dos Campos: Instituto Nacional de Pesquisas Espaciais (INPE), 2006, 22 p. apud ZANOTTA, D. C. Uma metodologia adaptativa baseada em objetos para detecção semi-automática de desmatamento na Amazônia por imagens de sensoriamento remoto. Tese de Doutorado do Curso de Pós-Graduação em Sensoriamento Remoto do Instituto Nacional de Pesquisas Espaciais, São José dos Campos. 2014.
- COSTA, W. S.; FONSECA, L. M. G; KÖRTING, T. S. Classificação de pastagens cultivadas e formações campestres nativas no Cerrado brasileiro a partir da análise de séries temporais extraídas de índices EVI do sensor MODIS. *Anais XVII Simpósio Brasileiro de Sensoriamento Remoto - SBSR*, João Pessoa-PB, Brasil, 25 a 29 de abril de 2015, INPE

DICE, L. R. Measures of the amount of ecologic association between species. *Ecology*, v. 26, n.3, p. 297-302, 1945.

DOAN, C.; LIONG, S. Generalization for multilayer neural network: Bayesian regularization or early stopping. *Proceedings of Asia Pacific Association of Hydrology and Water Resources 2nd Conference*, 2004.

FERREIRA, L. G.; FERREIRA, N. C.; FERREIRA, M. E. Sensoriamento remoto da vegetação: evolução e estado da arte. *Acta Sci. Biol. Sci.*, v. 30, n. 4, p. 379-390, 2008 apud GINCIENE, B. R.; BITENCOURT, M. D. Utilização do EVI (Enhanced Vegetation Index) para maior sensibilidade na detecção de mudanças temporais em fragmentos de floresta estacional semidecidual. *Anais XV Simpósio Brasileiro de Sensoriamento Remoto - SBSR*, Curitiba, PR, Brasil, 30 de abril a 05 de maio de 2011, INPE p.1620.

FEARNSIDE, P. M. Greenhouse gases from deforestation in Brazilian Amazonia: net committed emissions. *Climatic Change*, v. 35, p. 321–360, 1997.

FLOREZANO, T. G. *Iniciação em Sensoriamento Remoto*. 3ª edição. São Paulo: Oficina de textos, 2011.

GOZALEZ, R. C; WOODS, R. E. *Processamento Digital de Imagens*. 3ª edição. São Paulo: Pearson, 2010.

HAGAN, M. T.; DEMUTH, H. B.; BEALE, M. H. *Neural Network Design*. [S.l.]: PWS Pub.,1995 apud GIL, A. M. Reconhecimento de dígitos manuscritos: busca de um classificador com máxima taxa de acerto. *Dissertação de Mestrado do Curso de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Amazonas*, Manaus, 2014.

HANLEY, JA; MCNEIL, BJ. The meaning and Use of the Area under a Receiver Operating Characteristic (ROC) Curve. *Radiology*, 1982; 143(1):29-36.

HAYKIN, S. A comprehensive foundation. *Neural Networks*, v. 2, n. 2004, 2004 apud ALMEIDA, L. M. Avaliação da gravidade da malária utilizando técnicas de extração de características e redes neurais artificiais. *Dissertação de Mestrado do Curso de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Amazonas*, Manaus, 2015.

HOTELLING, H. Analysis of complex statistical variables into principal components. *Journal of Educational Psychology*, v. 24 (6), p. 417-441, 1933 apud FALCÃO, T. A. Comparação do desempenho do classificador de Novidades com o classificador do vizinho mais próximo no reconhecimento facial. *Dissertação de Mestrado do Curso de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Amazonas*, Manaus, 2014.

HERTZ, John, Krogh, Andres and Palmer, Richard G. . "Introduction to the Theory of Neural Computation", Addison-Wesley Publ. Comp. cop., 1991 apud GINCIENE, B. R.; BITENCOURT, M. D. Utilização do EVI (Enhanced Vegetation Index) para maior sensibilidade na detecção de mudanças temporais em fragmentos de floresta estacional semidecidual. *Anais XV Simpósio Brasileiro de Sensoriamento Remoto - SBSR*, Curitiba, PR, Brasil, 30 de abril a 05 de maio de 2011, INPE p.1620.

HUETE, A.; DIDAN, K.; MIURA T.; RODRIGUEZ E. P; GAO, X.; FERREIRA L. G. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment* 83 Elsevier Science Inc 2002 p. 195-213.

INPE – Instituto Nacional de Pesquisas Especiais. Metodologia para o Cálculo da Taxa Anual de Desmatamento na Amazônia Legal. São José dos Campos: INPE, 2013. Disponível online: http://www.obt.inpe.br/prodes/metodologia_TaxaProdes.pdf. Acesso: 27/03/2017.

_____. Projeto PRODES. São José dos Campos: INPE, 2017. Disponível online: <http://www.obt.inpe.br/prodes/index.php>. Acesso: 28/03/2017.

_____. Desflorestamento nos Municípios da Amazônia Legal para o ano de 2015. São José dos Campos: INPE, 2016. Disponível online: <http://www.dpi.inpe.br/prodesdigital/prodesmunicipal.php>. Acesso: 28/03/2017

JACCARD, P. The distribution of the flora in the alpine zone. *New Phytologist*, v. 11, n. 2, p. 37-50, 1912.

JENSEN, J. R. Sensoriamento Remoto do Ambiente. Uma perspectiva em recursos terrestres. Tradução de José Carlos N. Epiphânio, Antonio R. Formaggio, Athos R. Santos, Bernardo F. T. Rudorff, Cláudia M. Almeida, Lênio S. Galvão. São José dos Campos, SP. Editora Parêntese. 2009.

KOHAVI, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. *Ijcai*, 1995. p.1137-1145.

KOLMOGOROV, A. On the representation of continuous functions of several variables by superpositions of continuous functions of a smaller number of variables. *Dokl. Akad. Nauk SSSR*, 1956. p.179-182.

KOUTROUMBAS, K. *Pattern Recognition*. 4th. Academic Press, 2008. 984 ISBN 978-1-59749-272-0 apud ALMEIDA, L. M. Avaliação da gravidade da malária utilizando técnicas de extração de características e redes neurais artificiais. Dissertação de Mestrado do Curso de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Amazonas, Manaus, 2015.

KOVÁCS, Z. L. Redes neurais artificiais. Editora Livraria da Física, 2002. ISBN 8588325144 apud ALMEIDA, L. M. Avaliação da gravidade da malária utilizando técnicas de extração de características e redes neurais artificiais. Dissertação de Mestrado do Curso de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Amazonas, Manaus, 2015.

KUCK, T. N.; KEIZER, E. W. H.; PACHECO, P.; LIRA, R. V. C.; VASCONCELOS, S. A.; ARRUDA, A. N. Mapeamento multitemporal (2001-2009) do uso da terra no bioma Amazônia do estado do Mato Grosso através de imagens MODIS. *Anais XV Simpósio Brasileiro de Sensoriamento Remoto - SBSR*, Curitiba, PR, Brasil, 30 de abril a 05 de maio de 2011, INPE p.7776

LI, M.; MA, L.; BLASHKE, T.; CHENG, L.; TIEDE, D. A systematic comparison of different object-based classification techniques using high spatial resolution imagery in agricultural environments. *International Journal of Applied Earth Observation and Geoinformation*. Volume 49, July 2016, Pages 87–98.

MACKAY, D. J. C. Bayesian interpolation. *Neural Computation* 4, 415-447. 1992 ALMEIDA, L. M. Avaliação da gravidade da malária utilizando técnicas de extração de características e redes neurais artificiais. Dissertação de Mestrado do Curso de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Amazonas, Manaus, 2015.

MORÉ, J. J. The Levenberg-Marquardt algorithm: implementation and theory. In: (Ed.). *Numerical analysis*: Springer, 1978. p.105-116. ISBN 3540085386 apud ALMEIDA, L. M. Avaliação da gravidade da malária utilizando técnicas de extração de características e redes neurais artificiais. Dissertação de Mestrado do Curso de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Amazonas, Manaus, 2015.

MEDEIROS, T. H. Treinamento de Redes Neurais Artificiais com otimização multi-objetivo e regularização Bayesiana: um estudo comparativo. Dissertação de Mestrado do Curso de Pós-Graduação em Engenharia Elétrica da Escola de Engenharia da Universidade Federal de Minas Gerais, Belo Horizonte, 2004.

MENESES, P. R., ALMEIDA, T. Introdução ao Processamento de Imagens de Sensoriamento Remoto. Brasília, 2012. Disponível em: <http://cnpq.br/documents/10157/56b578c4-0fd5-4b9f-b82a-e9693e4f69d8>.

MORAN, E.F. Deforestation and land use in the Brazilian Amazon. *Human Ecology*, v. 21, p. 1–21, 1993.

MYERS, N. Tropical forests: present status and future outlook. *Climate Change*. v. 19, p. 3-32, 1991.

NASA. MODIS - Moderate Resolution Imaging Spectroradiometer, 2016. Disponível em: <http://modis.gsfc.nasa.gov/data/>. Acesso em: maio de 2016.

NETO, J. M. M.; MOITA, G. C. Uma introdução à análise exploratória de dados multivariados. *Química Nova*, v. 21, n. 4, p. 467-469, 1998 apud ALMEIDA, L. M. Avaliação da gravidade da malária utilizando técnicas de extração de características e redes neurais artificiais. Dissertação de Mestrado do Curso de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Amazonas, Manaus, 2015.

NOMA, A.; KÖRTING, T. S.; FONSECA, L. M. G. Uma Comparação entre Classificadores usando Regiões e Perfis EVI para Agricultura. *Anais XVI Simpósio Brasileiro de Sensoriamento Remoto - SBSR*, Foz do Iguaçu, PR, Brasil, 13 a 18 de abril de 2013.

NOVO, E. M. L. M.; *Sensoriamento Remoto: Princípios e Aplicações*. 3ª edição. São Paulo: Editora Edgard Blücher Ltda. 363p. 2008.

PATROCINIO, A. C. Classificador automático de achados mamográficos em imagens digitais de mamas densas utilizando técnicas híbridas. Tese de Doutorado do Curso de Pós-Graduação da Escola de Engenharia de São Carlos da Universidade de São Paulo, São Carlos, 2004.

PEARSON, K. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine Series* 6, v. 2, p. 559-572, 1901 apud FALCÃO, T. A. Comparação do desempenho do classificador de Novidades com o classificador do vizinho mais próximo no reconhecimento

facial. Dissertação de Mestrado do Curso de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Amazonas, Manaus, 2014.

PIMENTEL, T. R. G.; Classificação de padrões temporais de uso do solo e cobertura da terra em séries temporais de índice de vegetação utilizando um sistema neuro-difuso; São José dos Campos. 2014. Dissertação de Mestrado do Curso de Pós-Graduação em Computação Aplicada do Instituto Nacional de Pesquisas Espaciais, São José dos Campos. 2014.

RAIA JUNIOR; Acessibilidade e mobilidade na estimativa de um índice de potencial de viagens utilizando redes neurais artificiais e sistemas de informações Geográficas. Tese de Doutorado do Curso de Pós-graduação da Escola de Engenharia de São Carlos – Universidade de São Paulo. <http://www.teses.usp.br/teses/disponiveis/18/18137/tde-10112001-160812/publico/RaiaJrT.pdf>

REIS, M.; Detecção de mudanças de uso e cobertura da terra utilizando dados óticos e de micro-ondas em uma região da Amazônia Brasileira (2014); São José dos Campos. 2014. Dissertação de Mestrado do Curso de Pós-Graduação em Sensoriamento Remoto do Instituto Nacional de Pesquisas Espaciais, São José dos Campos. 2014.

RUMELHART, D. E. e MCCLELLAND, J. L. (1986). Parallel distributed processing: explorations in the microstructure of cognition, volume 1: foundations. MIT Press, Cambridge, MA, USA. 1986.

SAFARZADEH KHOOSHABI, G. Segmentation Validation Framework. Master Thesis, Linköping University, Department of Biomedical Engineering, Sweden, 2013.

SARLE, W. S. Stopped Training and Other Remedies for Overfitting. To appear in Proceedings of the 27th Symposium on the Interface, 1995.

SYMEONAKIS, E. *Modelling land cover change in a mediterranean environment using random forests and a multi-layer neural network model. Geoscience and Remote Sensing Symposium (IGARSS), 2016 IEEE International.*

THEODORIDIS, S.; KOUTROUMBAS, K. Pattern Recognition. 4. ed. San Diego, California: Elsevier, 2009 apud XAVIER, C. M. Segmentação, classificação e quantificação de bacilos de tuberculose em imagens de baciloscopia de campo claro através do emprego de uma nova técnica de classificação de pixels utilizando máquinas de vetores de suporte. Dissertação de Mestrado do Curso de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Amazonas, Manaus, 2012.

TIKHONOV, A. N.; ARSENIN, V. Y. Solutions of Ill Posed Problems. Vh Winston. 1977.

TUBB, N. A development path to success in neural computing. Expert Systems Applications, v. 9, n. 5, p. 5-9, 1993 apud ALMEIDA, L. M. Avaliação da gravidade da malária utilizando técnicas de extração de características e redes neurais artificiais. Dissertação de Mestrado do Curso de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Amazonas, Manaus, 2015.

TUCKER, C. J.; TOWNSHEND, J. R. G. Strategies for monitoring tropical deforestation using satellite data. International Journal of Remote Sensing, v. 21, n. 2, 3, p. 1461– 1471, 2000.

USGS. Landsat Missions. Disponível em: [http://landsat.usgs.gov/l8handbook_section2 .php](http://landsat.usgs.gov/l8handbook_section2.php). Acesso em: junho de 2016.

USGS - United States Geological Survey. What are the best spectral bands to use for my study?. Disponível em: <https://landsat.usgs.gov/what-are-best-spectral-bands-use-my-study>. Acesso em: maio de 2017.

WANG, W.; HU, H.; HU, J. Land Cover Change Detection Based On MODIS 250m Vegetation Index Time Series Data. 2009 17th International Conference on Geoinformatics. pp. 1-6 12-14 Aug. 2009.

WEI, W. W.-S. Time series analysis. Addison-Wesley publ, 1994. ISBN 0201159112.

XAVIER, C. M. Segmentação, classificação e quantificação de bacilos de tuberculose em imagens de baciloscopia de campo claro através do emprego de uma nova técnica de classificação de pixels utilizando máquinas de vetores de suporte. Dissertação de Mestrado do Curso de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Amazonas, Manaus, 2012.

YIN, H.; PFLUGMACHER, D.; KENNEDY, R. E.; SULLA-MENASHE, D.; HOSTERT, P. Mapping Annual Land Use and Land Cover Changes Using MODIS Time Series. IEEE Journal of selected topics in applied earth observations and remote sensing, vol. 7, no. 8, august 2014.

ZANOTTA, D. C. Uma metodologia adaptativa baseada em objetos para detecção semi-automática de desmatamento na Amazônia por imagens de sensoriamento remoto. Tese de Doutorado do Curso de Pós-Graduação em Sensoriamento Remoto do Instituto Nacional de Pesquisas Espaciais, São José dos Campos. 2014.

APÊNDICE 1

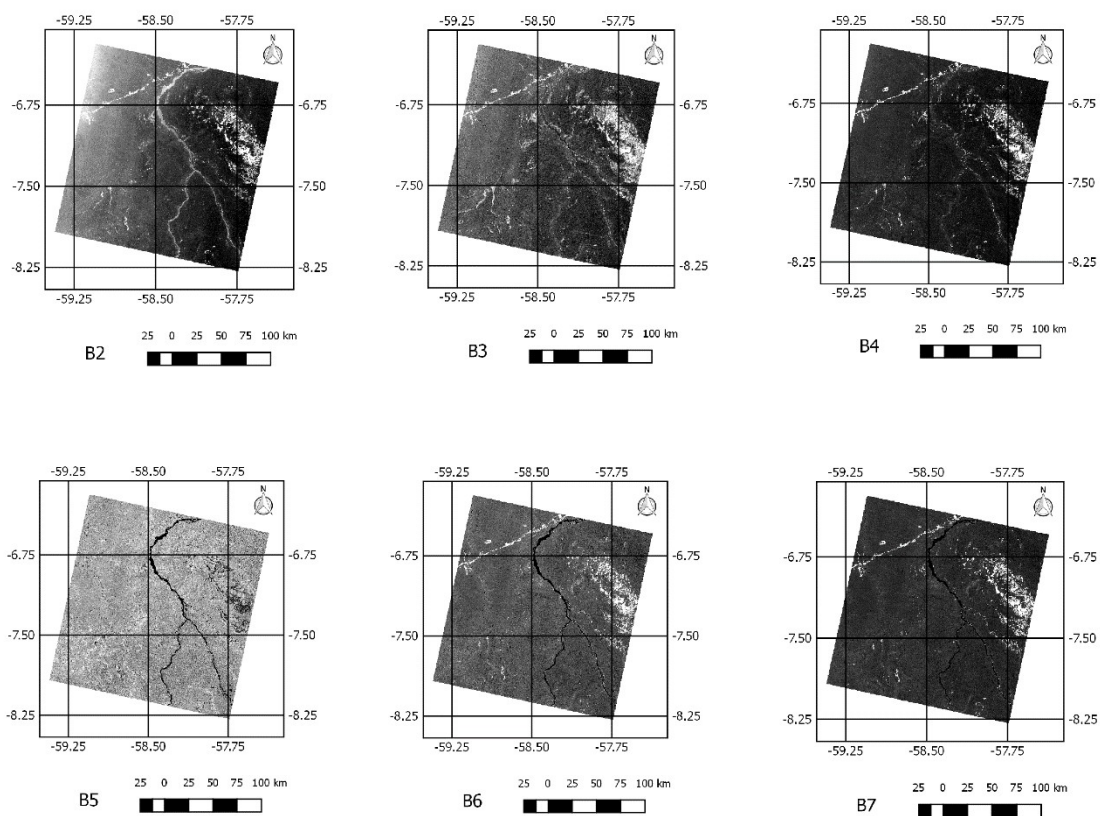
IMAGENS LANDSAT-8/OLI DAS REGIÕES DE APUÍ, HUMAITÁ E LÁBREA
ANALISADAS NESTE TRABALHO DE DISSERTAÇÃO

Figura 28 – Imagem 229/065 LANDSAT-8/OLI

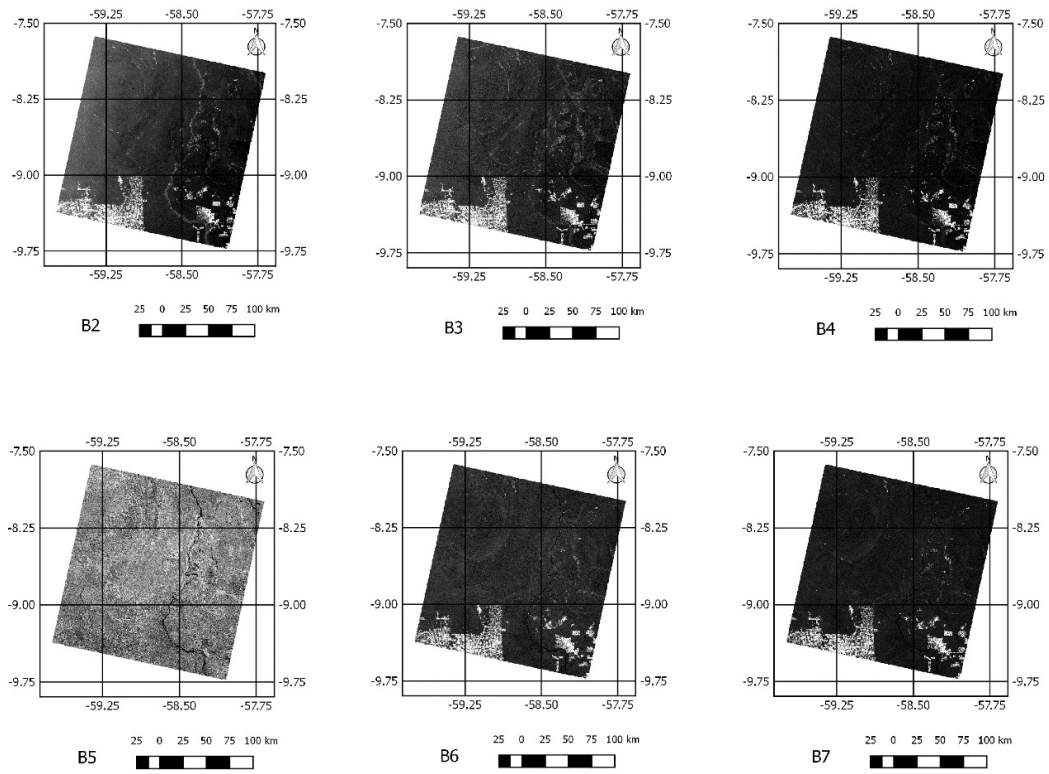


Figura 29 – Imagem 229/066 LANDSAT-8/OLI

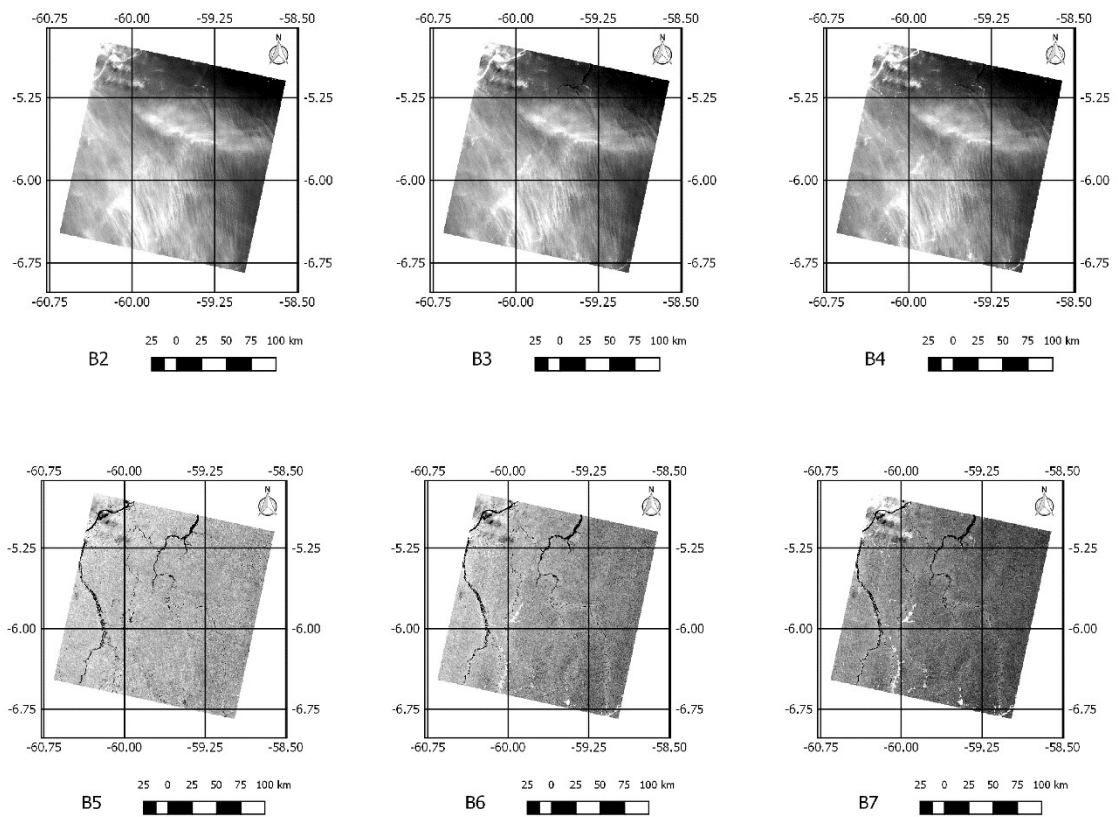


Figura 30 – Imagem 230/064 LANDSAT-8/OLI

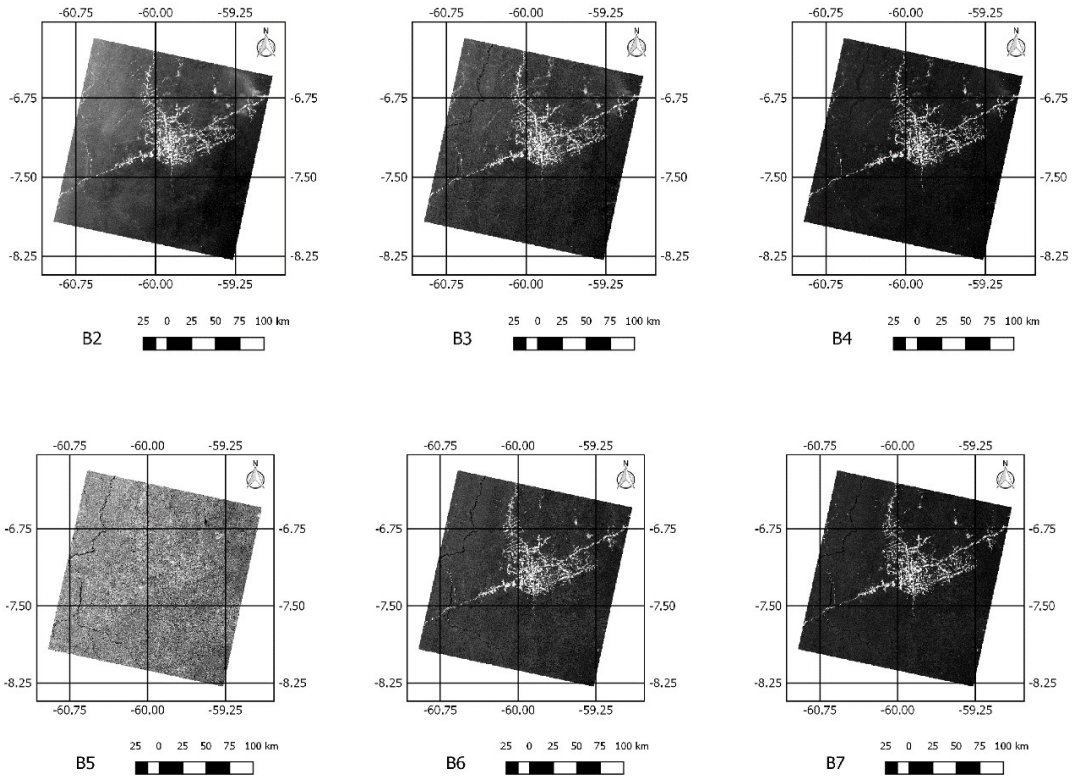


Figura 31 – Imagem 230/065 LANDSAT-8/OLI

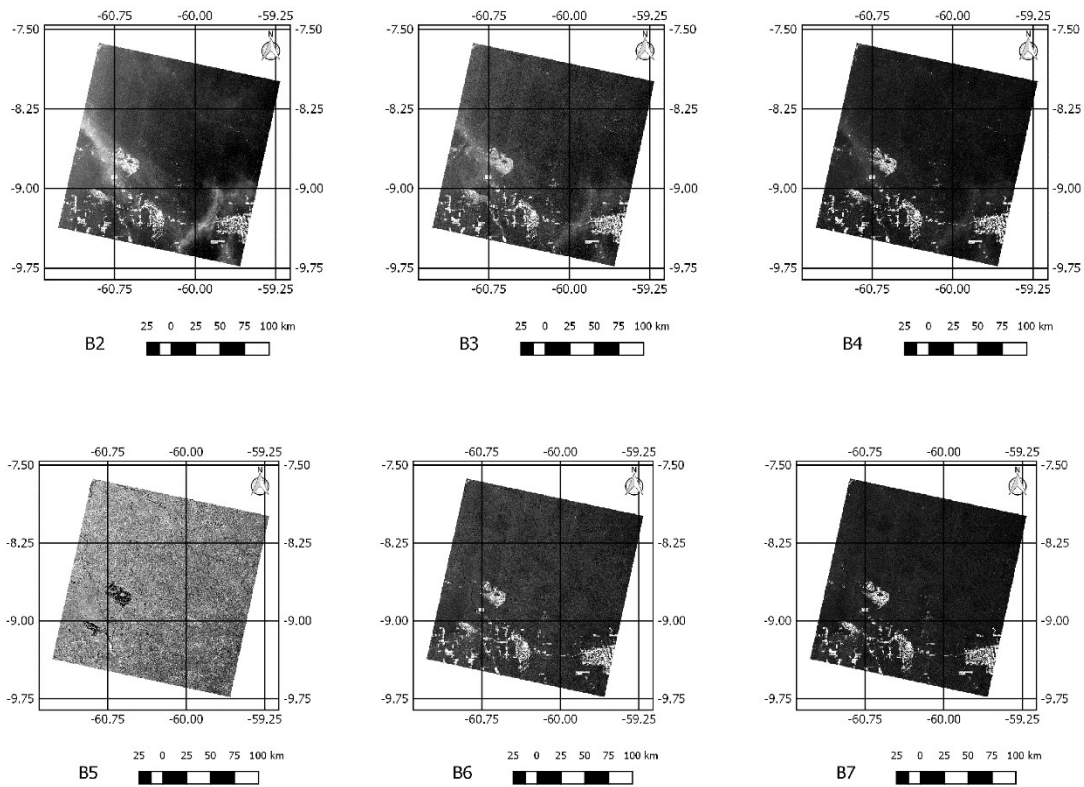


Figura 32 – Imagem 230/066 LANDSAT-8/OLI

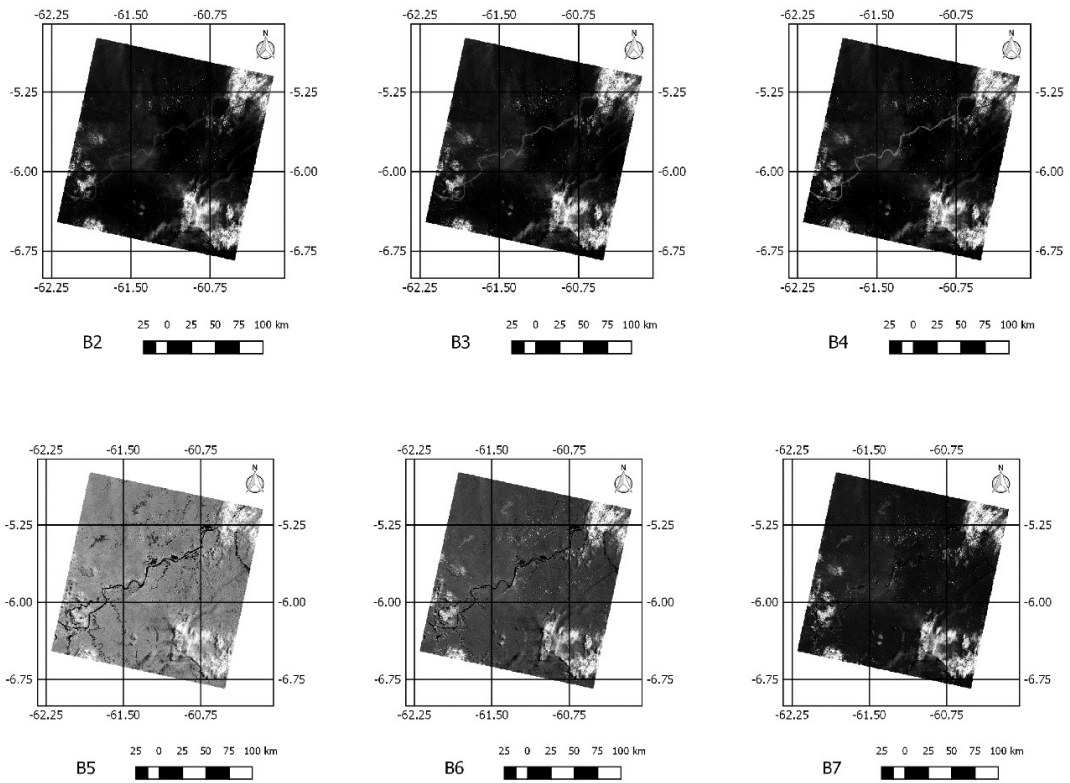


Figura 33 – Imagem 231/064 LANDSAT-8/OLI

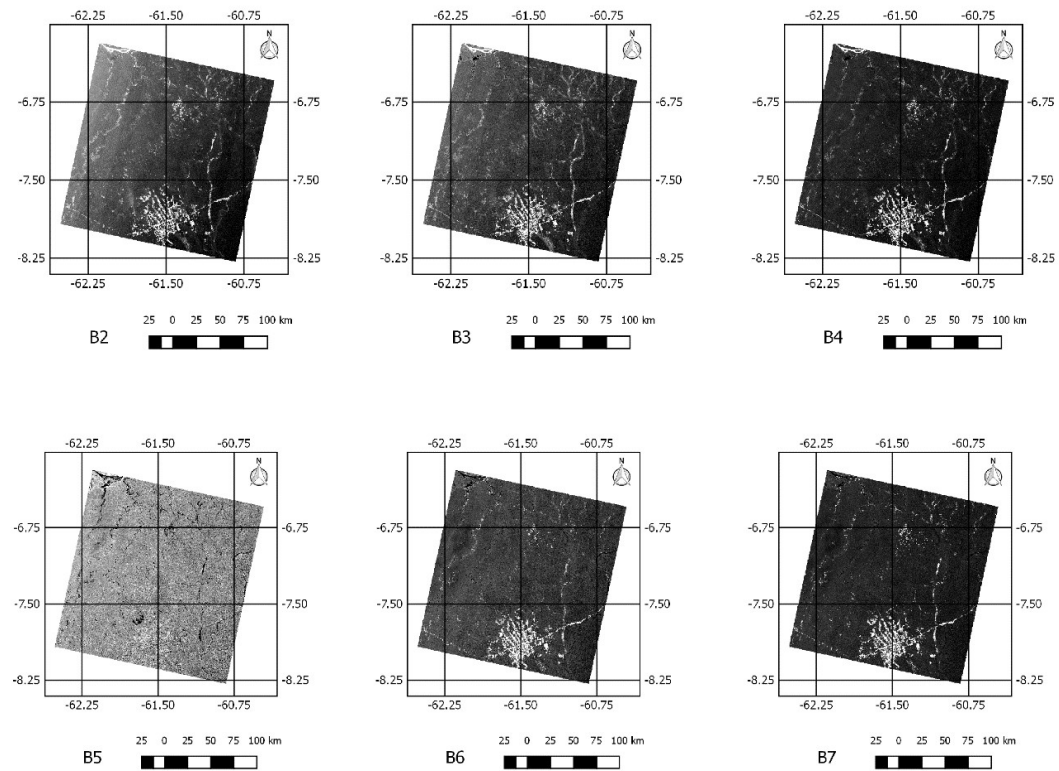


Figura 34 – Imagem 231/065 LANDSAT-8/OLI

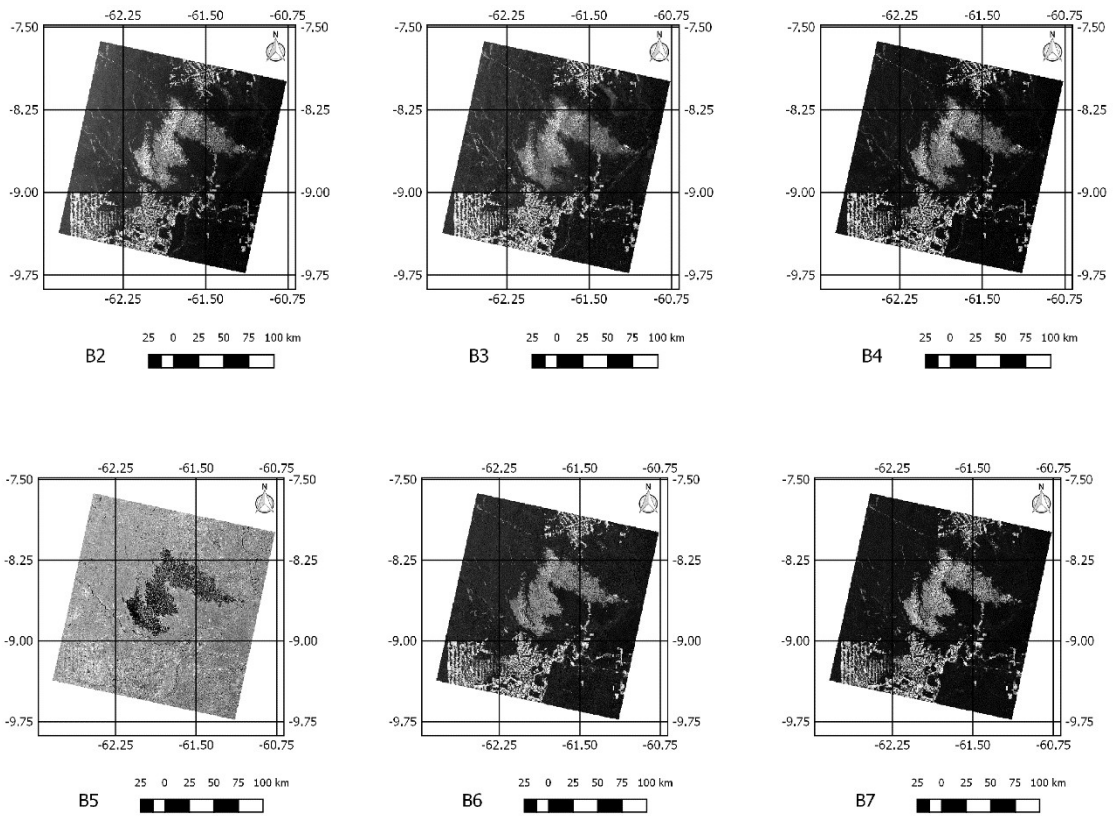


Figura 35 – Imagem 231/066 LANDSAT-8/OLI

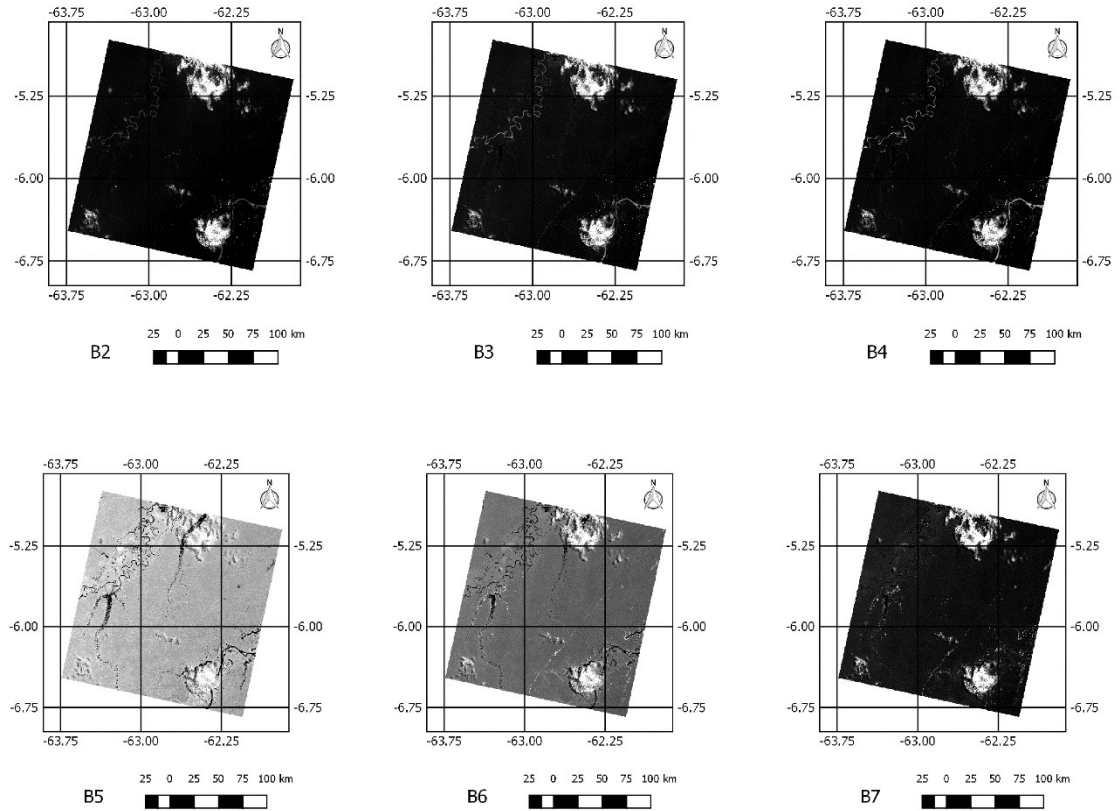


Figura 36 – Imagem 232/064 LANDSAT-8/OLI

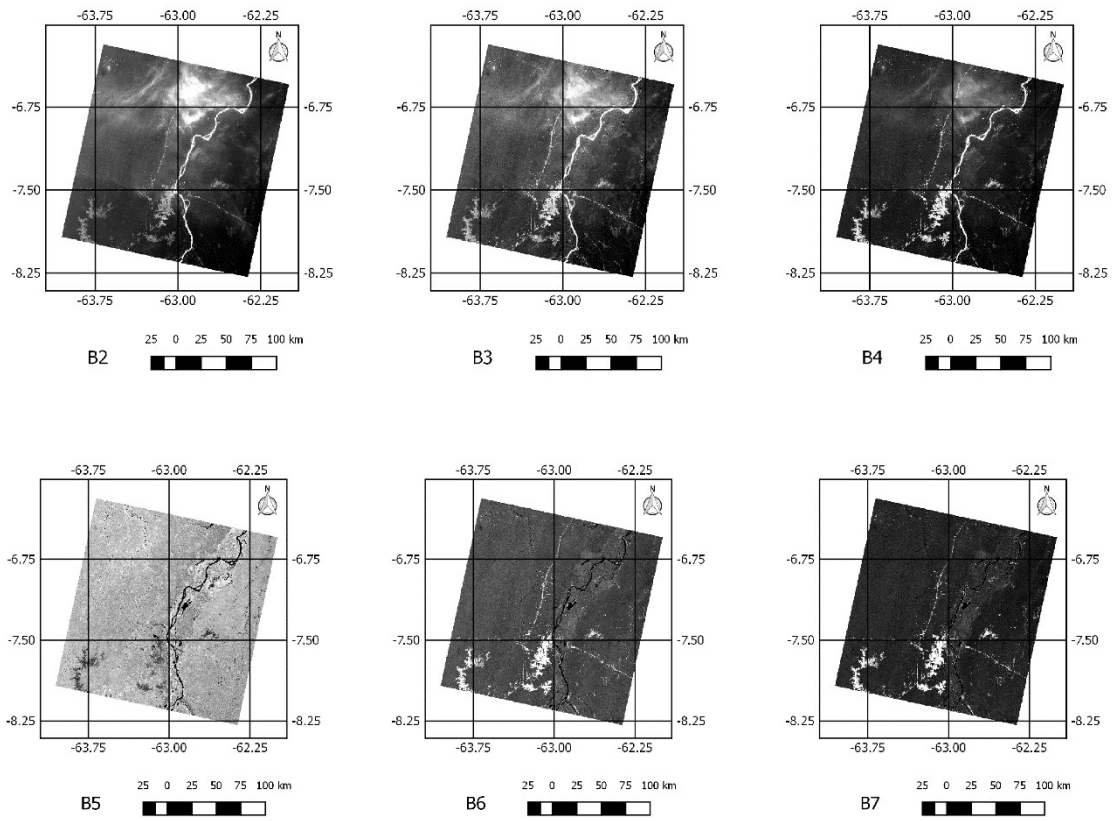


Figura 37 – Imagem 232/065 LANDSAT-8/OLI

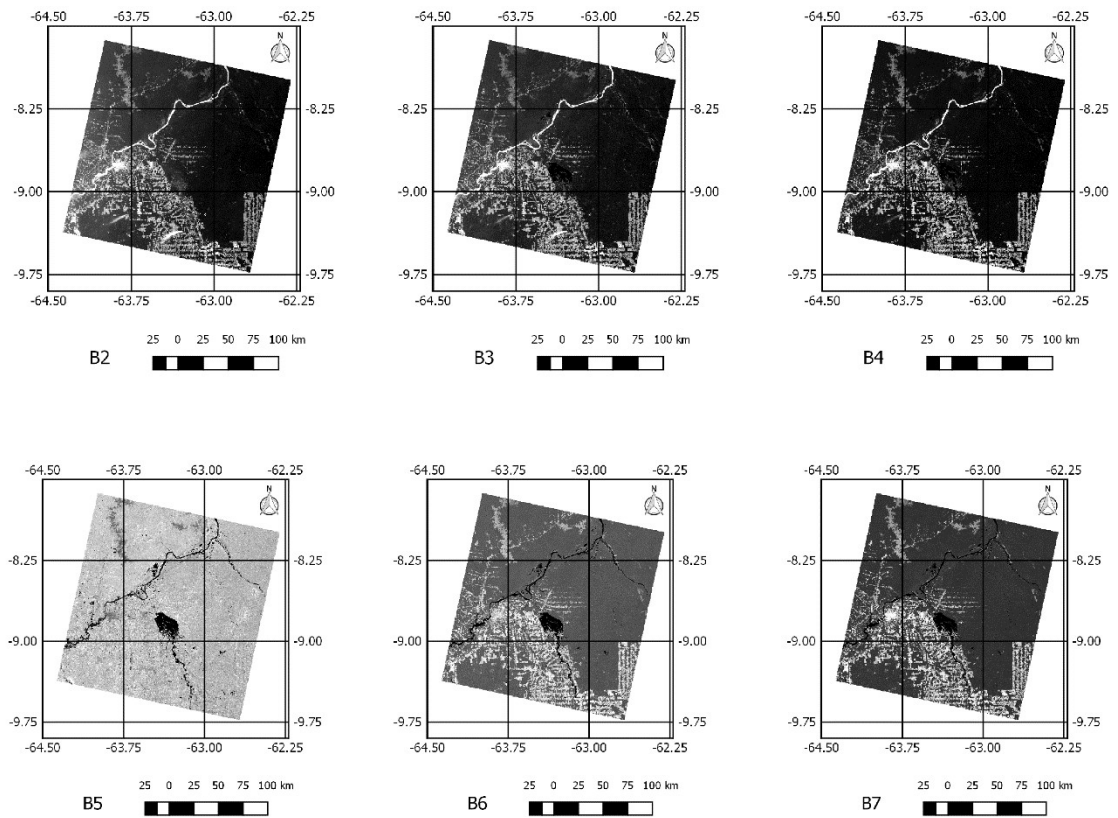


Figura 38 – Imagem 232/066 LANDSAT-8/OLI

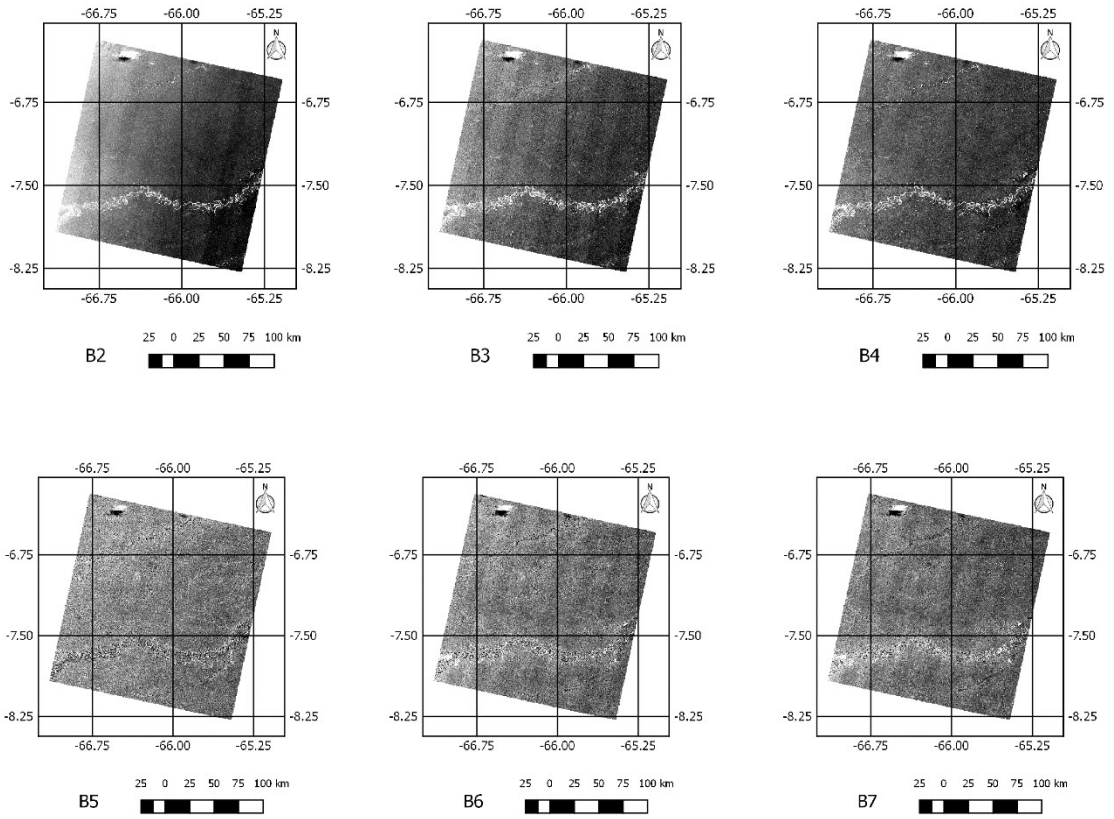


Figura 39 – Imagem 001/065 LANDSAT-8/OLI

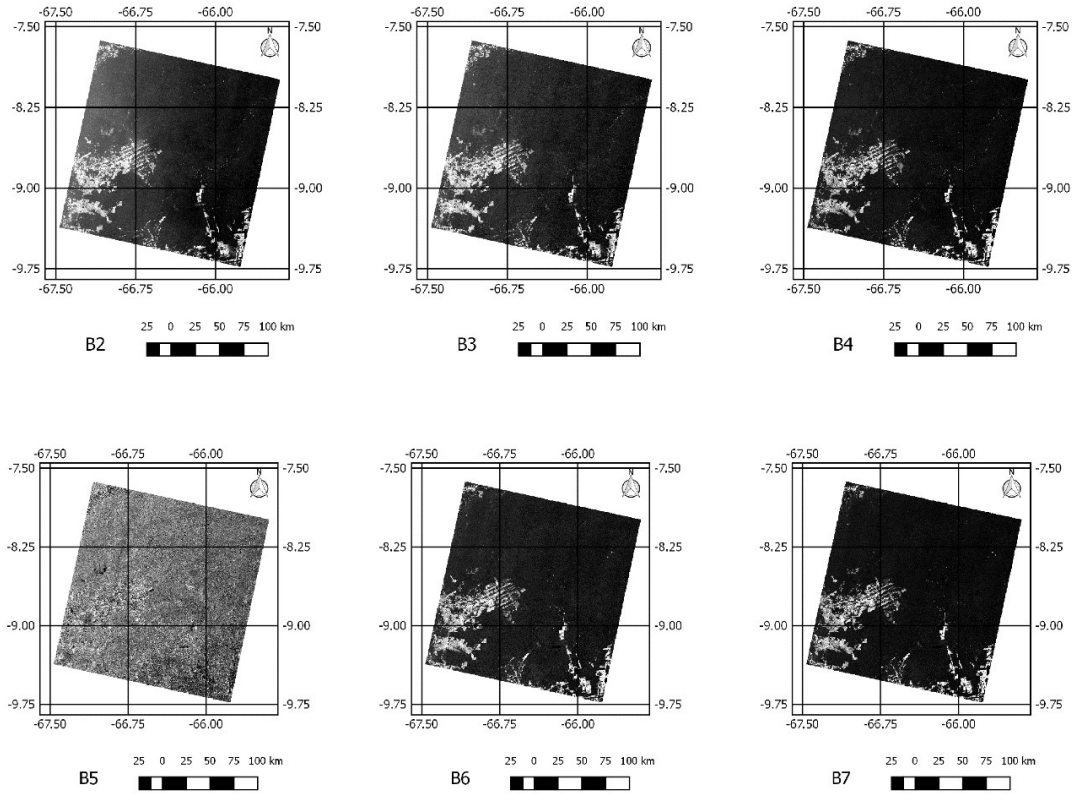


Figura 40 – Imagem 001/066 LANDSAT-8/OLI

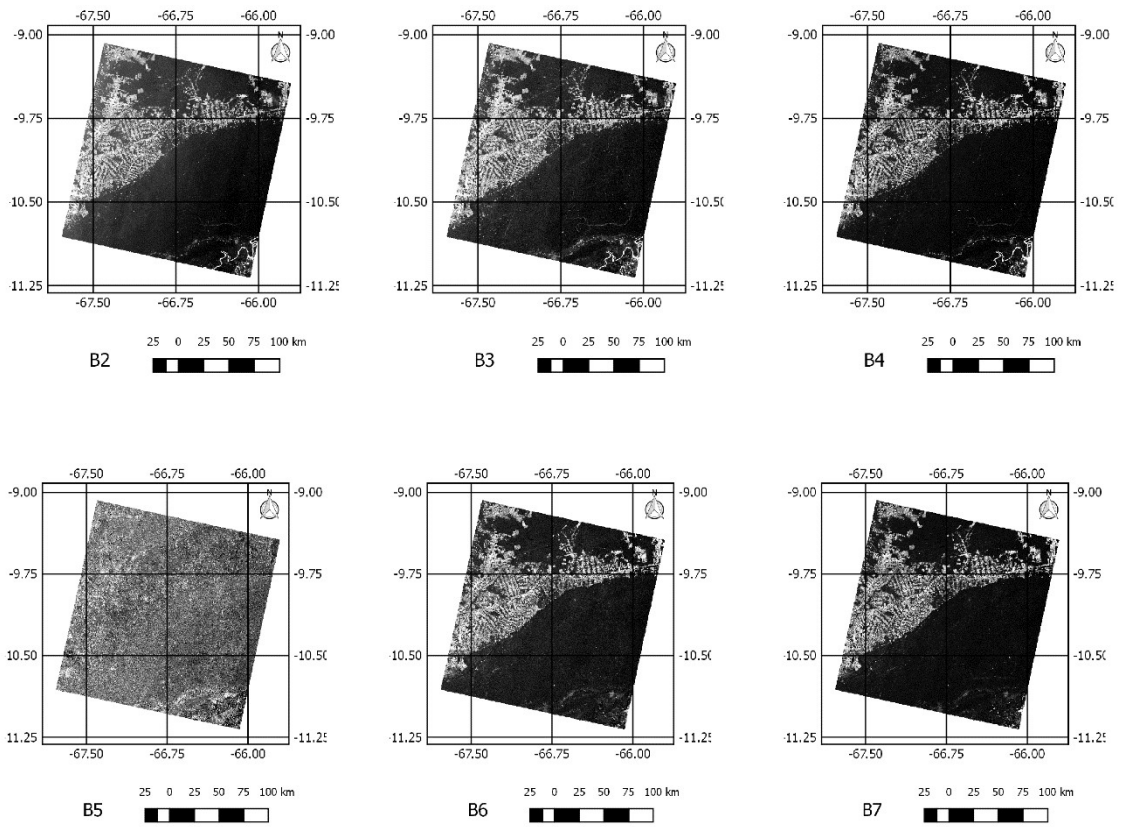


Figura 41 – Imagem 001/067 LANDSAT-8/OLI

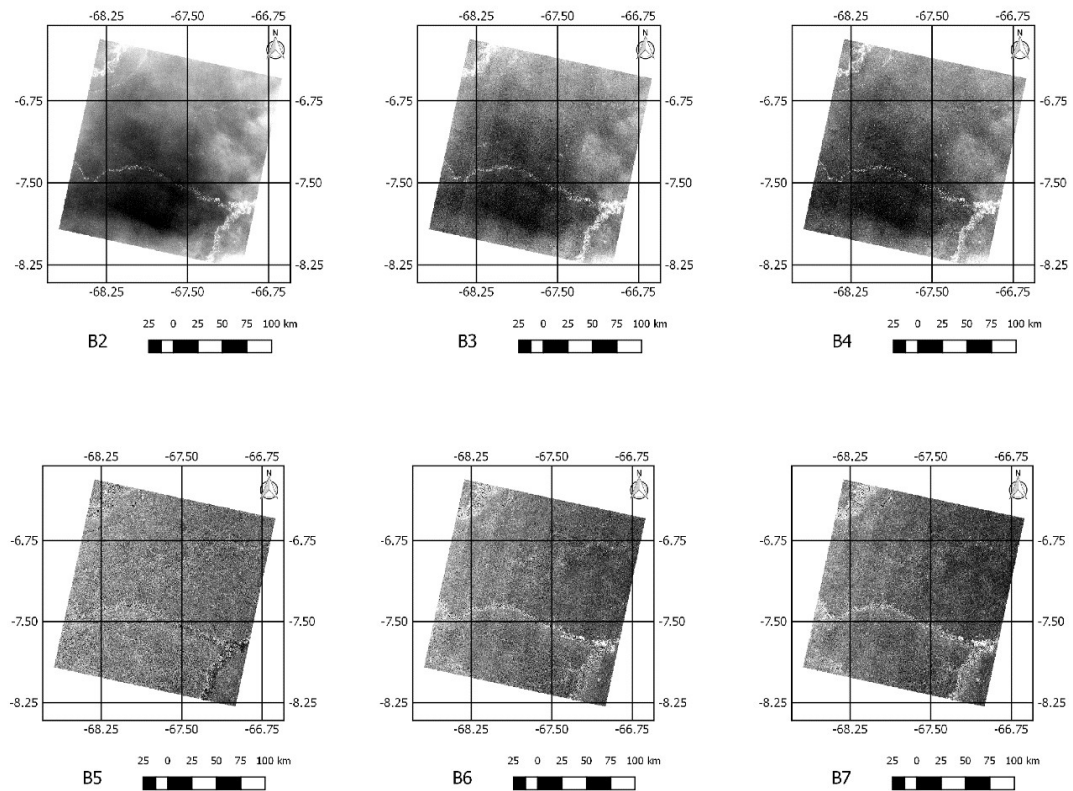


Figura 42 – Imagem 002/065 LANDSAT-8/OLI

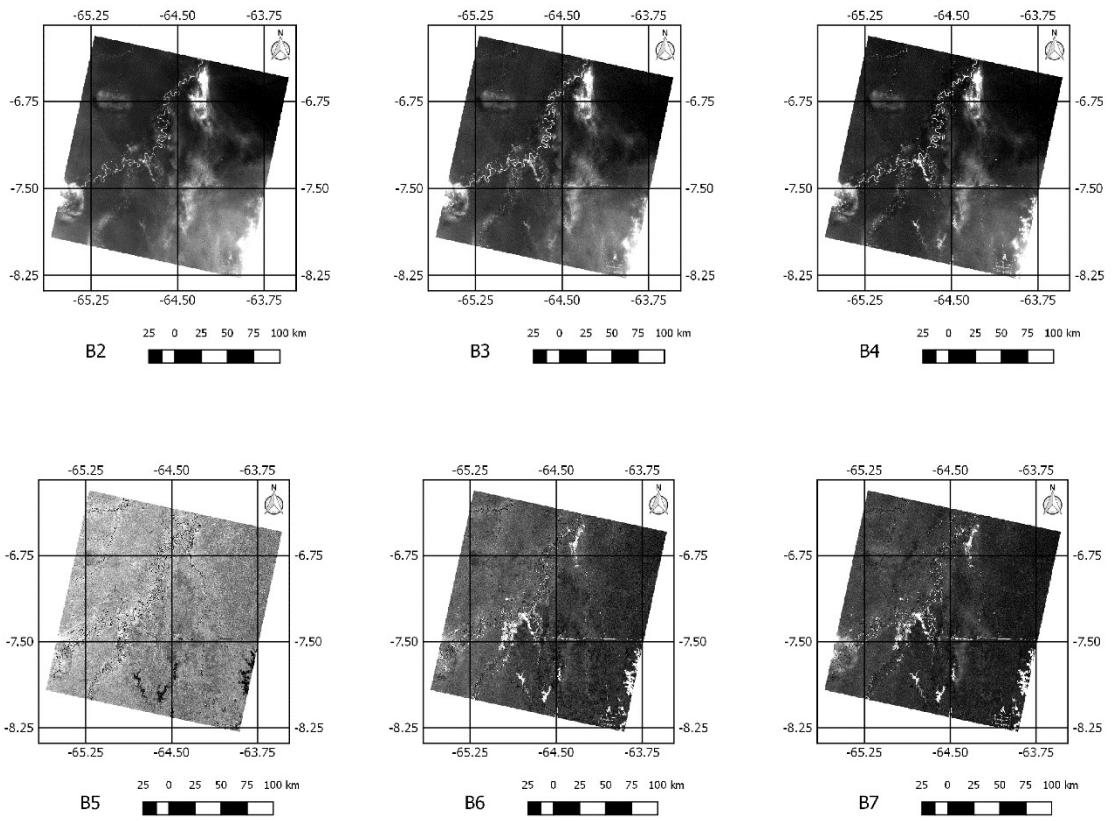


Figura 43 – Imagem 233/065 LANDSAT-8/OLI

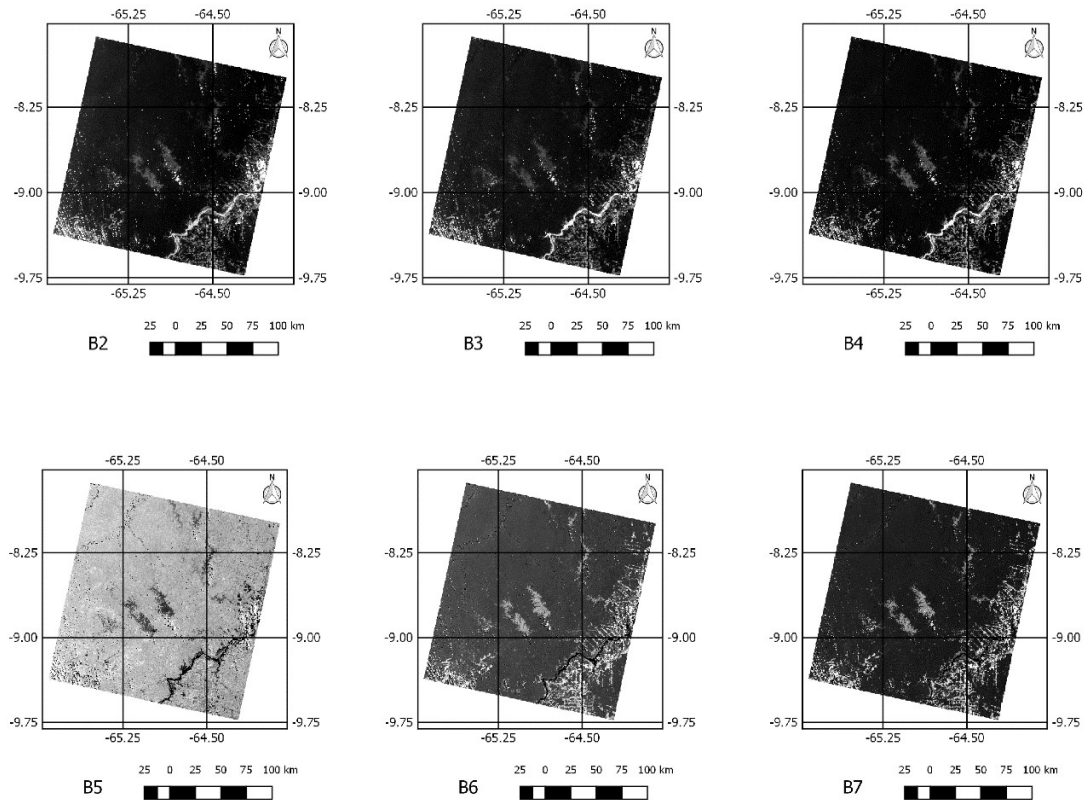


Figura 44 – Imagem 233/066 LANDSAT-8/OLI

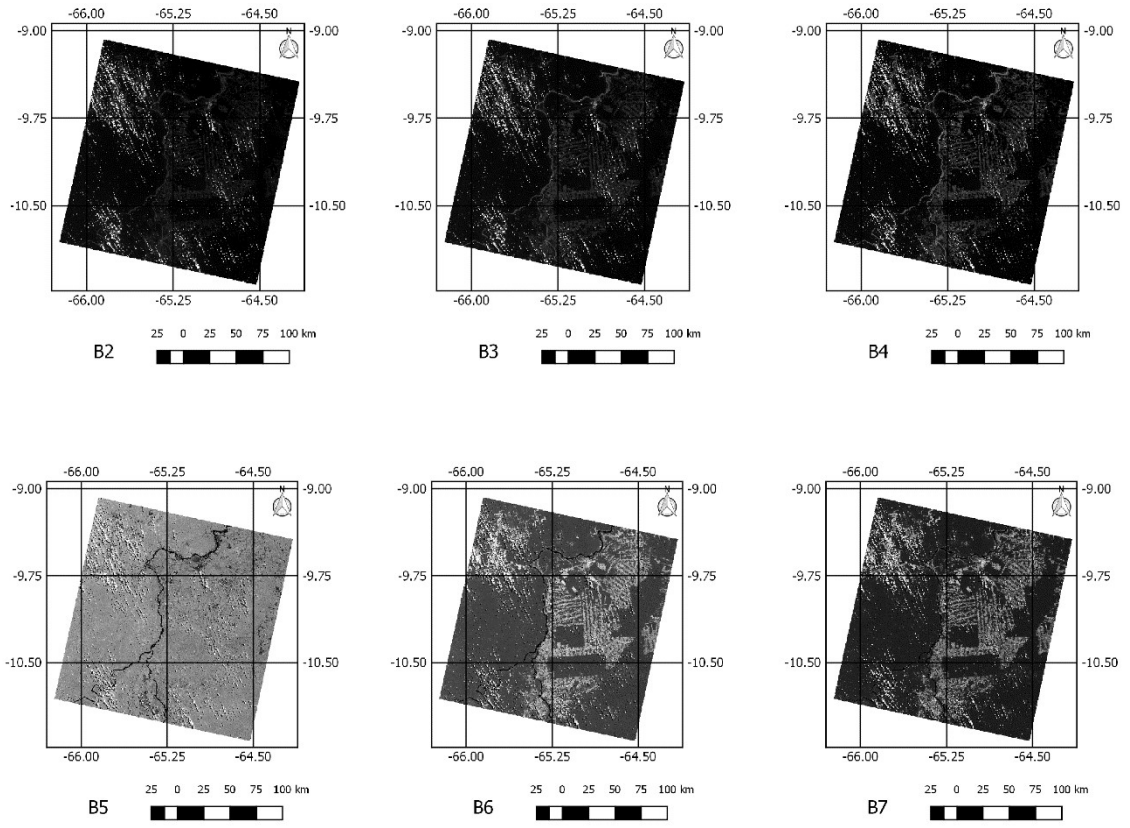


Figura 45 – Imagem 233/066 LANDSAT-8/OLI

APÊNDICE 2

ARTIGO ACEITO NO XIII ENCONTRO ANUAL DE COMPUTAÇÃO (ENACOMP)
2017

Detecção de áreas desmatadas utilizando redes neurais artificiais associados a métodos para melhoria da generalização

Joel P. Oliveira¹, Marly G. F. Costa², Cícero F. F. Costa Filho³

¹Centro Gestor e Operacional do Sistema de Proteção da Amazônia – Censipam
Avenida do Turismo, nº 1350 – Tarumã – Manaus – AM, Brasil

²Universidade Federal do Amazonas – UFAM
Av. General Rodrigo Octávio, nº 6200 – Coroado I – Manaus – AM, Brasil

{joelparente,marlygfcosta,cffcfilho}@gmail.com

Abstract. *This study proposes a methodology to study deforesting areas using Landsat optical images obtained through remote sensing. The methodology proposes the association of methods for improving neural networks generalization, as early stopping and regularization, with techniques for characteristic extractions, as scalar selection technique, principal component analysis and linear discriminant analysis. The performance of the methodology is evaluated using measures as accuracy, sensitivity, specificity and area under ROC curve. The results shown that the best results are obtained with the early stop generalization method associated with the linear discriminant analysis technique. The study is carried out in three regions of the Amazon State, located near the municipalities of Apuí, Humaitá and Lábrea. An area under ROC curve and an accuracy as good as 0.997 and 0.97, respectively, are obtained.*

Resumo. *Este estudo propõe uma metodologia para o estudo de áreas de desmatamento utilizando imagens ópticas Landsat obtidas por sensoriamento remoto. É proposta a associação de métodos para melhorar a generalização das redes neurais, como a parada antecipada e regularização, com técnicas de extração de características, como seleção escalar, análise de componentes principais e análise de discriminante linear. O desempenho da metodologia é avaliado utilizando medidas como precisão, sensibilidade, especificidade e área sob curva ROC. Os melhores resultados foram obtidos com o método de generalização parada antecipada associado à técnica de análise discriminante linear. O estudo foi realizado em três regiões do Estado do Amazonas, localizadas nos municípios de Apuí, Humaitá e Lábrea. Foi obtida uma área sob a curva ROC e uma precisão de 0,997 e 0,97 respectivamente.*

1. Introdução

Sensoriamento remoto é obtenção de dados sobre um objeto sem estar em contato físico direto com o mesmo [Jensen, 2009]. De forma mais específica, [Novo, 2008] define que sensoriamento remoto é a utilização de diversas tecnologias, como sensores, equipamentos instalados a bordo de aeronaves, espaçonaves, ou outras plataformas, com o objetivo de estudar os fenômenos que ocorrem na superfície da Terra.

Para realizar o monitoramento do desmatamento na Amazônia, as imagens geradas a partir de sistemas de sensoriamento remoto possuem grande importância, pois segundo [Noma *et al.*, 2013], imagens de satélite permitem realizar o monitoramento do desmatamento em larga escala e compreender a dinâmica da paisagem. Neste sentido, desde 1988, o Instituto Nacional de Pesquisas Espaciais (INPE) é o principal responsável pelo monitoramento da Floresta Amazônica. Neste aspecto, [Noma *et al.*, 2013] destacam dois projetos: Programa de Monitoramento do Desflorestamento na Amazônia Legal (PRODES) e TerraClass. O PRODES fornece dados por meio de mapas anuais de desmatamento na região Amazônica. Já o projeto TerraClass utiliza os dados gerados pelo PRODES para realizar uma classificação de uso e cobertura de solo.

Na literatura, diversos trabalhos propõem metodologias de classificação automática de imagens para monitoramento do solo. [Pimentel, 2014], utilizou redes neurais artificiais para classificar padrões temporais de uso do solo e cobertura da terra, utilizando séries temporais de índices de vegetação. [Reis, 2014] avaliou o potencial de imagens SAR, ópticas e a combinação das duas para detectar mudanças em uma área da floresta nacional do Tapajós. [Zanotta, 2014] propôs uma nova metodologia para detecção semi-supervisionada de desflorestamento na Amazônia brasileira utilizando imagens Landsat. Os trabalhos apresentados não utilizaram estratégias contra o sobre ajustamento (*overfitting*), fenômeno que ocorre quando uma rede neural é excessivamente treinada, trazendo como consequência o sobre ajustamento dos parâmetros na tarefa de reconhecimento dos padrões apresentados durante a fase de treinamento.

Nos últimos anos, o desempenho das redes neurais sofreu uma melhoria substancial, com o desenvolvimento de novas técnicas de generalização, como parada antecipada e regularização [Hagan *et al.*, 2014], ensejando novos critérios de parada para o treinamento das redes neurais, que não o Erro Quadrático Médio (*Mean Square Error – MSE*), tradicionalmente utilizado. Além disso, técnicas robustas de extração de características foram desenvolvidas, como Análise de Componentes Principais (*Principal Component Analysis – PCA*), Análise Discriminante Linear (*Linear Discriminant Analysis – LDA*) e Seleção Escalar de Características (SEC) [Theodoridis e Koutroumbas, 2008], que permitiram uma melhoria substancial na compactação e representatividade da informação. Nesse trabalho, pretende-se avaliar o desempenho dos métodos de generalização das redes neurais, quando associados aos métodos de extração de características anteriormente citados, para o reconhecimento de áreas desmatadas na região sul do estado do Amazonas. O texto está organizado da seguinte maneira. A Seção 2 apresenta o conjunto de dados e a metodologia. Na Seção 3, os resultados são exibidos e interpretados. Finalmente, algumas conclusões são exibidas na Seção 4.

2. Materiais e Métodos

2.1. Materiais

As imagens utilizadas nesse estudo foram imagens ópticas Landsat8 retiradas do *site*: <http://earthexplorer.usgs.gov>. As bases de imagens utilizadas foram dos municípios de Apuí, Humaitá, Lábrea localizados no sul do estado do Amazonas, conforme mostrado na Figura 1. Os dados de entrada do classificador foram extraídos do projeto PRODES, do ano de 2015, disponibilizado por [Inpe, 2017]. Foram extraídas amostras de feições correspondentes a áreas desmatadas e áreas não desmatadas (floresta e outros tipos de vegetação denominados de não-floresta segundo [Inpe, 2017]). Para cada uma das três regiões foram selecionadas 2000 amostras, 1000 amostras pertencentes a áreas não desmatadas e 1000 amostras pertencentes a áreas desmatadas. A metodologia utilizada neste estudo foi comparada com a abordagem metodológica adotada nos trabalhos de [Reis, 2014], [Pimentel, 2014] e [Zannota, 2014]. Esses

trabalhos foram escolhidos para comparação, pois os mesmos também utilizaram imagens ópticas da região Amazônica.

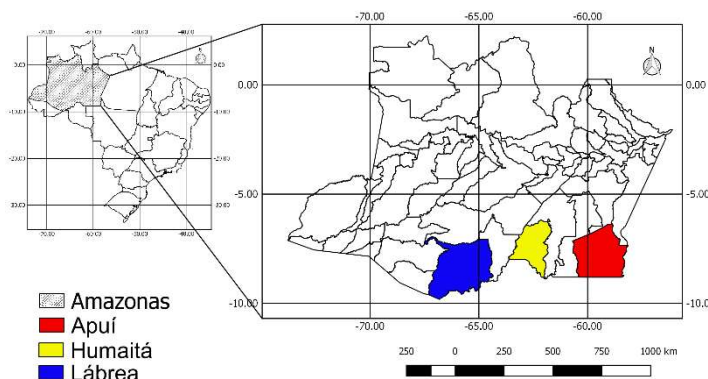


Figura 1. Mapa do Amazonas delimitando as regiões utilizadas nesse estudo

2.2. Métodos

2.2.1. Sistema de Reconhecimento

As etapas do método de reconhecimento de padrões utilizadas foram: extração de características, seleção de características e classificação.

O conjunto de características utilizados neste trabalho compreende características de textura. Com base na revisão da literatura, as características de textura mais utilizadas e que empregamos neste trabalho foram: média, variância, desvio padrão, homogeneidade, energia, correlação, contraste e entropia. Segundo [Usgs, 2017], para detecção da vegetação, as bandas de imagem sugeridas são: 2, 3, 4, 5, 6, 7. A essas bandas citadas anteriormente, acrescentamos mais duas bandas calculadas a partir dessas primeiras, conforme sugerido por [Huete *et al.*, 2002]: NDVI e EVI. As características foram numeradas de 1 a 64. A característica 1 corresponde a média da banda 2, a característica 2 corresponde a variância a banda 2, e assim sucessivamente, até a característica 64, que corresponde a entropia da banda EVI.

Para o treinamento das redes neurais foram selecionados conjuntos de 10, 20, 30, 40, 50 e 64 características. A seleção das mesmas deu-se através da utilização das técnicas PCA, LDA e SEC. Na técnica SEC a medida de separação de classes utilizada foi a Razão Discriminante de Fisher (*Fisher Discriminant Ratio – FDR*) [Theodoridis e Koutroumbas, 2008].

O classificador empregado foi uma rede neural, com arquitetura X-11-1, em que X pode assumir os valores de 64, 50, 40, 30, 20 ou 10, em função do número de variáveis de entrada. Os critérios de parada empregados para o treinamento da rede neural foram MSE, Parada Antecipada e Regularização. O limite máximo de iterações fixado para um treinamento foi de 1000. Os valores mínimos para os critérios MSE e Regularização que determinaram o final do treinamento foi de 10^{-6} . O treinamento da rede neural utilizando Parada Antecipada termina quando, no conjunto de validação, o MSE aumentar por mais de 6 épocas consecutivas. Para acelerar o treinamento da rede neural foi utilizado o método de Levenberg-Marquardt [Hagan *et al.*, 2014].

2.2.2. Metodologia de Treinamento e Teste

Para o treinamento da rede neural com o critério MSE e Regularização, a metodologia de treinamento e teste utilizada foi a de validação cruzada com 5 pastas [Blum *et al.*, 1999]. Nesse trabalho, utilizou-se um conjunto de 2000 amostras para treinamento e teste (1000 amostras de regiões de desmatamento e 1000 amostras de regiões de áreas não desmatadas). Esse total de

amostras foi dividido em 5 grupos de 400 amostras cada (200 de áreas desmatadas e 200 de áreas não desmatadas). A rede foi treinada com os dados de 4 grupos e testada com o quinto grupo. Como tem 5 grupos, esse processo foi repetido 5 vezes e uma média do desempenho foi calculada. Para o critério de Parada Antecipada, o conjunto de dados foi dividido em três partes iguais, treinamento, validação e teste. Três classificadores neurais distintos foram treinados e testados. Um deles foi treinado com amostras da região de Apuí. O segundo, com amostras da região de Humaitá. O terceiro, com amostras da região de Lábrea.

3. Resultados

A Tabela 1 mostra as características selecionadas para as regiões de Apuí, Humaitá e Lábrea utilizando a técnica SEC. Para a região de Apuí, a característica variância da Banda B4 foi a que apresentou um maior valor para a razão FDR, que foi de 1,32. Para a região de Humaitá, a característica variância da Banda B2 foi a que apresentou um maior valor para a razão FDR, que foi de 2,25. Para a região de Lábrea, a característica média da Banda B2 foi a que apresentou um maior valor para a razão FDR, que foi de 0,64.

Nas tabelas 2, 3 e 4 são apresentados dados de desempenho dos classificadores. Esses dados são resultados médios para acurácia, Área sob a curva de ROC (*Area Under ROC Curve – AUC*) [Hanley e Mcneil, 1982], sensibilidade e especificidade, para os métodos de treinamento MSE, Parada Antecipada e Regularização, combinados com as técnicas de extração de características SEC, PCA e LDA, para as regiões de Apuí, Humaitá e Lábrea, respectivamente.

3.1. Análise dos Resultados

A partir da Tabela 1 observa-se que as características 63 (Contraste, Banda EVI) e 62 (Correlação, Banda EVI) sempre estão presentes nos conjuntos de características selecionados. Isso pode ser justificado pelo fato da banda EVI ser formada a partir do cálculo do índice de vegetação EVI. Índices de vegetação foram desenvolvidos para explorar as propriedades espectrais de vegetação e, portanto, ressaltam o comportamento espectral da vegetação em relação ao solo e a outros alvos da superfície terrestre.

Os principais parâmetros utilizados para avaliação de um método de classificação são a AUC e a acurácia do método. Observando as Tabelas 2, 3 e 4, com respeito a AUC, derivam-se algumas conclusões importantes.

Tabela 1. Conjunto de Características selecionados pela técnica seleção escalar de características para as regiões de Apuí, Humaitá e Lábrea

Número de Características	Conjunto de Características		
	Apuí	Humaitá	Lábrea
10	A1= {18,19,2,63,5,62,26,61,3,53}	H1= {26,27,63,53,62,55,25,5,31,54}	H1= {1,5,17,62,9,63,61,54,47,41}
20	A2=A1+{55,30,13,42,31,6,50,21,58,15}	H2=H1+{13,7,35,15,21,34,6,23,29,19}	H2=H1+{13,49,30,57,21,39,33,14,29,2}
30	A3=A2+{14,29,27,7,54,33,45,43,23,37}	H3=H2+{47,14,39,45,33,30,37,11,22,18}	H3=H2+{6,15,37,7,22,45,31,38,53,55}
40	A4=A3+{22,51,59,39,46,57,47,49,38,41}	H4=H3+{46,61,38,40,56,48,64,51,52,59}	H4=H3+{60,46,44,2,36,25,28,20,52,12}
50	A5=A4+{17,34,60,35,52,9,44,36,20,10}	H5=H4+{16,24,43,32,8,60,44,36,10,20}	H5=H4+{26,18,4,27,40,56,3,48,64,16}

Tabela 2. Resultados médios para acurácia, AUC, sensibilidade e especificidade, para os métodos de treinamento MSE, parada antecipada e regularização para a região de Apuí

Região	Métodos de parada do treinamento	N. de Caract.	Técnica de Seleção de Características											
			SEC				PCA				LDA			
			Acur.	AUC	Sens	Esp	Acur.	AUC	Sens	Esp	Acur.	AUC	Sens	Esp
Apuí	MSE	10	99,80	0,998 ± 0,0022	99,80	99,80	95,90	0,959 ± 0,0102	95,60	96,20	99,95	0,999 ± 0,0011	99,90	100,00
		20	99,85	0,998 ± 0,0019	99,70	99,90	94,95	0,950 ± 0,0114	94,30	95,60	99,85	0,998 ± 0,0019	99,90	99,80
		30	99,80	0,998 ± 0,0022	99,70	99,90	93,15	0,931 ± 0,0132	92,40	93,90	99,95	0,998 ± 0,0019	99,90	100,00
		40	99,85	0,998 ± 0,0019	99,80	99,90	93,70	0,937 ± 0,0127	94,30	93,10	99,95	0,999 ± 0,0011	99,90	100,00
		50	99,85	0,998 ± 0,0019	99,70	100,00	92,60	0,926 ± 0,0138	91,20	94,00	99,95	0,999 ± 0,0011	99,90	100,00
		64	99,85	0,998 ± 0,0019	99,80	99,90	99,85	0,998 ± 0,0019	99,80	99,90	99,85	0,998 ± 0,0011	99,80	99,90
	Parada Antecipada	10	99,90	0,999 ± 0,0016	99,85	99,95	97,33	0,973 ± 0,0082	97,35	97,30	99,93	0,999 ± 0,0014	99,85	99,99
		20	99,93	0,999 ± 0,0014	99,85	100,00	98,63	0,986 ± 0,0059	98,70	98,55	99,90	0,999 ± 0,0016	99,85	100,00
		30	99,93	0,999 ± 0,0014	99,85	100,00	95,78	0,958 ± 0,0104	95,70	95,85	99,93	0,999 ± 0,0014	99,85	99,95
		40	99,93	0,999 ± 0,0014	99,85	100,00	93,78	0,938 ± 0,0126	93,60	93,95	99,93	0,999 ± 0,0014	99,85	100,00
		50	99,90	0,999 ± 0,0016	99,80	100,00	95,48	0,955 ± 0,0107	95,80	95,15	99,93	0,999 ± 0,0014	99,85	100,00
		64	99,93	0,999 ± 0,0014	99,85	100,00	99,93	0,999 ± 0,0014	99,85	100,00	99,93	0,999 ± 0,0014	99,85	100,00
	Regularização	10	99,80	0,998 ± 0,0022	99,70	99,90	96,40	0,964 ± 0,0096	96,10	96,70	99,95	0,999 ± 0,0011	99,90	100,00
		20	99,90	0,999 ± 0,0016	99,80	100,00	84,25	0,843 ± 0,0199	81,30	87,20	99,95	0,999 ± 0,0011	99,90	100,00
		30	99,85	0,988 ± 0,0019	99,80	99,90	85,20	0,852 ± 0,0170	87,30	83,10	99,95	0,999 ± 0,0011	99,90	100,00
		40	99,85	0,988 ± 0,0019	99,80	99,90	86,10	0,861 ± 0,0162	81,30	90,90	99,95	0,999 ± 0,0011	99,90	100,00
		50	99,95	0,999 ± 0,0011	99,90	100,00	87,20	0,872 ± 0,0155	82,30	92,10	99,95	0,999 ± 0,0011	99,90	100,00
		64	99,85	0,999 ± 0,0011	99,90	100,00	99,95	0,995 ± 0,0014	99,90	99,90	99,95	0,999 ± 0,0011	99,90	100,00

Tabela 3. Resultados médios para acurácia, AUC, sensibilidade e especificidade, para os métodos de treinamento MSE, parada antecipada e regularização para a região de Humaitá

Região	Método de parada do treinamento	N. de Caract.	Técnica de Seleção de Características											
			SEC				PCA				LDA			
			Acur.	AUC	Sens	Esp	Acur.	AUC	Sens	Esp	Acur.	AUC	Sens	Esp
Humaitá	MSE	10	94,25	0,943 ± 0,0121	94,00	94,50	97,05	0,971 ± 0,0087	97,50	96,60	99,55	0,996 ± 0,0034	99,20	99,90
		20	95,05	0,950 ± 0,0112	94,20	95,90	97,75	0,977 ± 0,0076	97,90	97,60	99,35	0,994 ± 0,0040	99,00	99,70
		30	98,85	0,988 ± 0,0054	98,60	99,10	98,00	0,980 ± 0,0071	97,80	98,20	99,60	0,996 ± 0,0032	99,30	99,90
		40	98,70	0,987 ± 0,0057	98,10	99,30	96,85	0,969 ± 0,0089	96,80	96,90	99,60	0,996 ± 0,0032	99,30	99,90
		50	98,80	0,988 ± 0,0055	98,40	99,20	95,25	0,953 ± 0,0110	95,40	95,10	99,50	0,995 ± 0,0035	99,30	99,70
		64	99,25	0,993 ± 0,0043	98,60	99,90	99,25	0,993 ± 0,0043	98,60	99,90	99,25	0,993 ± 0,0043	98,60	99,90
	Parada Antecipada	10	94,10	0,941 ± 0,0122	94,30	93,95	99,43	0,994 ± 0,0038	99,20	99,65	99,65	0,996 ± 0,0030	99,45	99,85
		20	97,00	0,970 ± 0,0087	96,25	97,75	98,98	0,990 ± 0,0051	98,95	99,00	99,68	0,997 ± 0,0029	99,45	99,90
		30	99,45	0,995 ± 0,0037	99,20	99,70	98,88	0,989 ± 0,0053	98,70	99,05	99,73	0,997 ± 0,0026	99,50	99,95
		40	99,53	0,995 ± 0,0035	99,30	99,75	99,08	0,991 ± 0,0048	99,15	99,00	99,73	0,997 ± 0,0026	99,50	99,95
		50	99,33	0,993 ± 0,0041	99,05	99,60	98,48	0,985 ± 0,0062	98,45	98,50	99,68	0,997 ± 0,0029	99,45	99,90
		64	99,60	0,996 ± 0,0032	99,40	99,80	99,60	0,986 ± 0,0032	98,40	99,80	99,60	0,996 ± 0,0032	99,40	99,80
	Regularização	10	93,10	0,931 ± 0,0133	92,00	94,20	98,25	0,982 ± 0,0067	98,40	98,10	99,50	0,995 ± 0,0035	99,30	99,70
		20	95,50	0,955 ± 0,0107	94,80	96,20	98,20	0,982 ± 0,0068	97,60	98,80	99,50	0,995 ± 0,0035	99,10	99,90
		30	99,15	0,991 ± 0,0046	98,70	99,60	97,20	0,977 ± 0,0076	97,70	97,70	99,60	0,996 ± 0,0032	99,30	99,90
		40	98,95	0,990 ± 0,0051	98,40	99,50	96,60	0,966 ± 0,0093	96,80	96,80	99,65	0,996 ± 0,0030	99,30	100,00
		50	98,75	0,988 ± 0,0056	98,10	99,40	96,25	0,963 ± 0,0098	96,10	96,40	99,65	0,996 ± 0,0030	99,30	100,00
		64	99,45	0,995 ± 0,0037	99,00	99,90	99,45	0,995 ± 0,0037	99,00	99,00	99,45	0,995 ± 0,0037	99,00	99,90

Tabela 4. Resultados médios para acurácia, AUC, sensibilidade e especificidade, para os métodos de treinamento MSE, parada antecipada e regularização para a região de Lábrea

Região	Métodos de parada do treinamento	N. de Caract.	Técnica de Seleção de Características											
			SEC				PCA				LDA			
			Acur.	AUC	Sens	Esp	Acur.	AUC	Sens	Esp	Acur.	AUC	Sens	Esp
Lábrea	MSE	10	92,10	0,921 ± 0,0142	92,30	91,90	98,50	0,985 ± 0,0062	98,40	98,60	99,50	0,995 ± 0,0035	99,50	99,50
		20	93,80	0,938 ± 0,0126	94,20	93,40	97,20	0,972 ± 0,0084	97,00	97,40	99,60	0,996 ± 0,0032	99,70	99,50
		30	96,40	0,964 ± 0,0096	95,70	97,10	95,85	0,958 ± 0,0103	96,20	95,50	99,50	0,996 ± 0,0034	99,50	99,60
		40	97,75	0,977 ± 0,0076	97,20	98,30	94,75	0,948 ± 0,0116	94,90	94,60	99,65	0,996 ± 0,0030	99,60	99,70
		50	99,50	0,995 ± 0,0035	99,30	99,70	93,90	0,939 ± 0,0125	93,70	94,10	99,70	0,997 ± 0,0027	99,60	99,80
		64	99,60	0,996 ± 0,0032	99,30	99,90	99,60	0,996 ± 0,0032	99,30	99,90	99,60	0,996 ± 0,0032	99,30	99,90
	Parada Antecipada	10	95,13	0,951 ± 0,0112	95,10	95,15	98,88	0,989 ± 0,0053	98,80	98,95	99,70	0,997 ± 0,0027	99,60	99,80
		20	95,85	0,959 ± 0,0103	96,00	95,70	98,75	0,987 ± 0,0056	98,90	98,60	99,75	0,998 ± 0,0025	99,65	99,85
		30	97,90	0,979 ± 0,0073	97,55	98,25	97,65	0,976 ± 0,0077	97,25	98,05	99,75	0,997 ± 0,0025	99,65	99,85
		40	98,53	0,985 ± 0,0061	98,35	98,70	98,03	0,980 ± 0,0071	97,95	98,10	99,73	0,997 ± 0,0026	99,70	99,75
		50	99,68	0,997 ± 0,0029	99,60	99,75	95,55	0,956 ± 0,0107	94,95	96,15	99,80	0,998 ± 0,0022	99,80	99,80
		64	99,68	0,997 ± 0,0029	99,60	99,75	99,68	0,997 ± 0,0029	99,60	99,75	99,68	0,997 ± 0,0029	99,60	99,75
	Regularização	10	93,50	0,935 ± 0,0129	92,50	94,50	94,50	0,985 ± 0,0062	95,00	94,00	99,50	0,995 ± 0,0035	99,60	99,40
		20	93,45	0,935 ± 0,0129	94,00	92,90	95,50	0,955 ± 0,0104	95,50	95,50	99,85	0,998 ± 0,0019	99,80	88,90
		30	95,65	0,956 ± 0,0105	95,80	95,50	96,70	0,967 ± 0,0083	97,00	96,40	99,75	0,998 ± 0,0025	99,70	99,80
		40	97,60	0,976 ± 0,0078	97,30	97,90	96,80	0,968 ± 0,0081	96,20	97,40	99,70	0,997 ± 0,0027	99,60	99,80
		50	99,65	0,996 ± 0,0030	99,40	99,90	97,50	0,975 ± 0,0076	97,10	97,90	99,70	0,997 ± 0,0027	99,80	99,60
		64	99,55	0,996 ± 0,0034	99,60	99,50	99,55	0,996 ± 0,0034	99,60	99,50	99,55	0,996 ± 0,0034	99,60	99,50

Em primeiro lugar, observa-se que a técnica LDA, para as três regiões, apresenta o melhor desempenho em relação a esses dois parâmetros. Por exemplo, para a região de Lábrea, para o método de regularização e para 30 variáveis, a área sob a curva de ROC e o desvio padrão obtidos para as técnicas SEC, PCA e LDA foram, respectivamente, $0,956 \pm 0,0105$, $0,967 \pm 0,0083$ e $0,998 \pm 0,0025$. Para esse exemplo, avaliando a significância da diferença entre as duas curvas ROC [Hanley e McNeil, 1982] das técnicas LDA e SEC, encontramos que $P = 0,000052 < 0,00001$. Então, a hipótese nula deve ser rejeitada (isto é, o desempenho da técnica LDA é superior ao da técnica SEC, ao nível de significância de 0,001%). Para esse exemplo, avaliando a significância da diferença entre as duas curvas ROC das técnicas LDA e PCA, encontramos que $P = 0,000506 < 0,0001$. Então, a hipótese nula deve ser rejeitada (isto é, o desempenho da técnica LDA é superior ao da técnica PCA, ao nível de significância de 0,01%).

Em segundo lugar, com respeito aos métodos de parada da rede neural, observa-se, nas três regiões, uma leve superioridade no desempenho da técnica de parada antecipada frente as outras duas, em relação aos dois parâmetros citados. Por exemplo, para a região de Humaitá, para a técnica PCA e para 50 variáveis a área sob a curva de ROC e o desvio padrão obtidos para os métodos MSE, parada antecipada e regularização foram, respectivamente, $0,953 \pm 0,0110$, $0,985 \pm 0,0062$ e $0,963 \pm 0,0062$. Para esse exemplo, avaliando a significância da diferença entre as duas curvas ROC dos métodos parada antecipada e MSE, encontramos que $P = 0,0054 < 0,01$. Então, a hipótese nula deve ser rejeitada (isto é, o desempenho do método parada antecipada é superior ao do método SEC, ao nível de significância de 1%). Para esse exemplo, avaliando a significância da diferença entre as duas curvas ROC dos métodos parada antecipada e regularização, encontramos que $P = 0,0278 > 0,01$. Então, a hipótese nula não deve ser rejeitada (isto é, o desempenho do método parada antecipada não é superior ao do método regularização, ao nível de significância de 1%).

Em terceiro lugar, observa-se, com respeito ao número de características, que para os métodos SEC e PCA, via de regra o desempenho com 64 características é melhor do que o desempenho com um número menor de características, em relação as variáveis anteriormente citadas. Por exemplo, para a região de Apuí, para a técnica PCA e parada o método de parada antecipada, a área sob a curva de ROC e o desvio padrão obtidos para 10 e 64 variáveis foram, respectivamente, $0,973 \pm 0,0082$ e $0,999 \pm 0,0014$. Para esse exemplo, avaliando a significância da diferença entre as duas curvas ROC dos métodos parada antecipada e MSE, encontramos

que $P = 0.00102 < 0,01$. Então, a hipótese nula deve ser rejeitada (isto é, o desempenho com 64 variáveis é superior ao desempenho com 10 variáveis, ao nível de significância de 1%).

O método proposto nesse trabalho foi comparado com: [Reis, 2014], [Pimentel, 2014] e [Zanotta, 2014]. Na mesma região utilizada por [Reis, 2014], o melhor desempenho foi obtido utilizando-se o método regularização associado à técnica LDA e 64 características, com uma acurácia média de 96,97%. Esse resultado foi superior àquele obtido pelo referido autor, uma acurácia de 84,2%. O autor trabalhou com oito características e utilizou o classificador de Mínima Distância Estocástica. Na região utilizada por [Pimentel, 2014], o melhor desempenho foi obtido utilizando-se o método de regularização associado à técnica LDA e com 64 características, com uma acurácia média de 96,94%. Esse resultado foi ligeiramente melhor que aquele obtido pelo referido autor, uma acurácia de 96,05%. O autor trabalhou com 1300 amostras e utilizou redes neurais como classificador. Na região utilizada por [Zanotta, 2014], o melhor desempenho foi obtido utilizando-se regularização associado à técnica LDA e com 64 características, com uma acurácia média de 96,95%. Esse resultado foi bem superior àquele obtido pelo referido autor, uma acurácia de 82%. O autor trabalhou com apenas 3 características e com o classificador baseado na análise do vetor de mudanças.

4. Conclusões

Os resultados mostraram a importância de se avaliar o resultado da combinação de novos métodos de generalização com técnicas de seleção de características. A técnica de seleção LDA apresentou um melhor desempenho em relação às técnicas SEC e PCA. O método Parada Antecipada apresentou um desempenho levemente superior em relação aos métodos de Regularização e MSE. Em relação à técnica LDA, observamos que o número de características utilizadas não tem influência sobre o desempenho do classificador. Como principal conclusão desse trabalho, ressaltamos que, em função das altas taxas de acurácia e de AUC obtidas, recomenda-se fortemente a utilização de técnicas de extração de características, juntamente com métodos de generalização de redes neurais, para classificação de áreas desmatadas, obtidas a partir de imagens ópticas. Em trabalhos futuros, propõe-se expandir a área geográfica de estudos, selecionando mais regiões do país, e utilizando-se classes de uso do solo mais variadas. Pretende-se também utilizar redes convolutivas para classificação das áreas desmatadas.

Referências

- Blum, A., Kalai, A., Langford, J. (1999) “Beating the hold-out: Bounds for k-fold and progressive crossvalidation”. In Proceedings of the twelfth annual conference on Computational learning theory, ACM, pages 203–208
- Hagan, M. T., Demuth, H. B., Beale, M. H. (1995) “Neural Network Design”. [S.l.]: PWS Pub., apud GIL, A. M. (2014) “Reconhecimento de dígitos manuscritos: busca de um classificador com máxima taxa de acerto”. In Dissertação de Mestrado do Curso de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Amazonas, Manaus.
- Hanley, J. A and McNeil, B. J. (1982) “The meaning and Use of the Area under a Receiver Operating Characteristic (ROC) Curve”. *Radiology*; 143(1):29-36.
- Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X., and Ferreira, L.G. (2002) “Overview of the radiometric and biophysical performance of the MODIS vegetation indices”, *Remote Sensing of Environment*, Volume 83, Issues 1–2, November, Pages 195-213.
- Inpe – Instituto Nacional de Pesquisas Especiais. (2017) “Projeto PRODES”, <http://www.obt.inpe.br/prodes/index.php>, Janeiro.

- Jensen, J. R. (2009) “Sensoriamento Remoto do Ambiente. Uma perspectiva em recursos terrestres”. Tradução de José Carlos N. Epiphanyo, Antonio R. Formaggio, Athos R. Santos, Bernado F. T. Rudorff, Cláudia M. Almeida, Lênio S. Galvão. São José dos Campos, SP. Editora Parêntese.
- Noma, A., Körting, T. S., Fonseca, L. M. G. (2013) “Uma Comparação entre Classificadores usando Regiões e Perfis EVI para Agricultura”. In Anais XVI Simpósio Brasileiro de Sensoriamento Remoto - SBSR, Foz do Iguaçu, PR, Brasil.
- Novo, E. M. L. M. (2008) “Sensoriamento Remoto: Princípios e Aplicações”. 3ª edição. São Paulo: Editora Edgard Blücher Ltda. 363p.
- Pimentel, T. R. G. (2014) “Classificação de padrões temporais de uso do solo e cobertura da terra em séries temporais de índice de vegetação utilizando um sistema neuro-difuso”. São José dos Campos. 2014. Dissertação de Mestrado do Curso de Pós-Graduação em Computação Aplicada do INPE, São José dos Campos.
- Reis, M. (2014) “Detecção de mudanças de uso e cobertura da terra utilizando dados óticos e de micro-ondas em uma região da Amazônia Brasileira”. São José dos Campos. Dissertação de Mestrado do Curso de Pós-Graduação em Sensoriamento Remoto do INPE, São José dos Campos.
- Symeonakis, E. (2016) “Modelling land cover change in a mediterranean environment using random forests and a multi-layer neural network model”. In Geoscience and Remote Sensing Symposium (IGARSS), IEEE International.
- Theodoridis, S.; Koutroumbas, K. (2008) “Pattern Recognition”. San Diego, California: Academic Press. 984 pages.
- Usgs - United States Geological Survey (2017) “What are the best spectral bands to use for my study? ”, <https://landsat.usgs.gov/what-are-best-spectral-bands-use-my-study>, Maio.
- Zanotta, D. C. (2014) “Uma metodologia adaptativa baseada em objetos para detecção semi-automática de desmatamento na Amazônia por imagens de sensoriamento remoto”. Tese de Doutorado do Curso de Pós-Graduação em Sensoriamento Remoto do INPE, São José dos Campos.