

**UNIVERSIDADE ESTADUAL DE PONTA GROSSA
SETOR DE CIÊNCIAS EXATAS E DA TERRA
PROGRAMA DE PÓS-GRADUAÇÃO EM GEOGRAFIA
MESTRADO GESTÃO DO TERRITÓRIO**

DINAMERES APARECIDA ANTUNES

**ANÁLISE ORIENTADA A OBJETO GEOGRÁFICO NA CARACTERIZAÇÃO
DO USO E OCUPAÇÃO DA TERRA EM SEGMENTOS DO RIO PITANGUI,
PARANÁ: AVALIAÇÕES PRELIMINARES**

**PONTA GROSSA
2015**

DINAMERES APARECIDA ANTUNES

**ANÁLISE ORIENTADA A OBJETO GEOGRÁFICO NA CARACTERIZAÇÃO
DO USO E OCUPAÇÃO DA TERRA EM SEGMENTOS DO RIO PITANGUI,
PARANÁ: AVALIAÇÕES PRELIMINARES**

Dissertação apresentada ao Programa de Pós-Graduação em Geografia - Mestrado em Gestão do Território da Universidade Estadual de Ponta Grossa, como requisito para obtenção do Título de Mestre.

Orientação: Prof^a Dra. Selma Regina Aranha
Ribeiro

**PONTA GROSSA
2015**

Ficha Catalográfica
Elaborada pelo Setor de Tratamento da Informação BICEN/UEPG

A627 Antunes, Dinameres Aparecida
Análise orientada a objeto geográfico na caracterização do uso e ocupação da terra em segmentos do rio Pitangui, Paraná: avaliações preliminares/ Dinameres Aparecida Antunes. Ponta Grossa, 2015. 140f.

Dissertação (Mestrado em Gestão do Território - Área de Concentração: Gestão do Território: Sociedade e Natureza), Universidade Estadual de Ponta Grossa. Orientadora: Prof^a Dr^a Selma Regina Aranha Ribeiro.

1.GEOBIA. 2.Classificação supervisionada. 3.Análise multivariada. I.Ribeiro, Selma Regina Aranha. II. Universidade Estadual de Ponta Grossa. Mestrado em Gestão do Território. III. T.


CDD: 528.8


TERMO DE APROVAÇÃO

DINAMERES APARECIDA ANTUNES

“ANÁLISE ORIENTADA A OBJETO GEOGRÁFICO NA CARACTERIZAÇÃO DO USO E OCUPAÇÃO DA TERRA EM SEGMENTOS DO RIO PITANGUI, PARANÁ”

Dissertação aprovada como requisito parcial para obtenção do grau de Mestre no Programa de Pós-Graduação em Geografia – Mestrado em Gestão do Território, Setor de Ciências Exatas e Naturais da Universidade Estadual de Ponta Grossa, pela seguinte banca examinadora:

Orientadora: 
Prof^a. Dr^a. Selma Regina Aranha Ribeiro
UEPG


Prof^a. Dr^a. Naissa Batista da Luz
FAO/EMBRAPA


Prof^a. Dr^a. Marcia Regina Calegari
UNIOESTE

Ponta Grossa, 04 de março de 2015.

Dedicatória

Dedico esse trabalho a minha amada mãe:

Amarilda de Fatima Antunes.

AGRADECIMENTOS

À Universidade Estadual de Ponta Grossa (UEPG) e ao Programa de Pós-Graduação em Geografia, pela oportunidade em realizar o curso.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelo subsídio financeiro que possibilitou o desenvolvimento da pesquisa.

À minha orientadora Prof^ª Dra. Selma Regina Aranha Ribeiro, pela orientação no desenvolvimento da dissertação.

Aos Professores do Programa de Pós-Graduação em Geografia pelos ensinamentos e estímulos.

À banca de defesa de qualificação pela contribuição interdisciplinar de cada um.

Aos amigos Kelly e Willian pela ajuda no escritório e em campo.

À minha família querida, principalmente a minha irmã Thamiris e minha mãe que tiveram paciência nos meus vários momentos ausente.

À todos que de alguma forma ajudaram na construção desse trabalho.

RESUMO

O objetivo desse trabalho foi mediante a GEOBIA (Geographic Object Based Image Analysis) caracterizar o uso e ocupação da terra em duas diferentes áreas no entorno do rio Pitangui, localizado entre os municípios de Castro, Carambeí e Ponta Grossa, Paraná. Devido à maior demanda de imagens de sensoriamento remoto com alta resolução espacial, há a necessidade de novas metodologias no processamento de imagem digital, pode-se citar a GEOBIA, que é análise orientada a objeto geográfico, que dentre as suas vantagens está a geração de polígonos para cada objeto geográfico criado e banco de dados relacional com diversos descritores, além de possibilitar o uso de informação espectral, espacial e de textura. A Análise de Componentes Principais (ACP) e Análise de Agrupamentos (AA) foram utilizadas com o objetivo de reduzir a dimensionalidade do banco de dados relacional das amostras de treinamento da classificação supervisionada e dessa forma, selecionar os descritores que proporcionassem melhores resultados para a classificação. Processou-se classificações supervisionadas orientadas a objetos utilizando descritores selecionados a partir da ACP, selecionados pela ACP mais AA, e todos os descritores gerados a partir da GEOBIA. Observou-se que os melhores resultados para ambas as áreas de estudo foram com os descritores selecionados pela ACP mais AA, que obtiveram na área de estudo 1 95,68% de precisão geral na matriz de confusão, e 0,95 no índice *kappa*, e para área de estudo 2 93,85% na matriz e 0,93 no índice *kappa*, valores considerados excelentes pela literatura, demonstrando que a seleção de descritores para a classificação orientada a objetos é pertinente, pois há descritores com informações redundantes que podem prejudicar o resultado final da classificação.

Palavras-chave: GEOBIA, classificação supervisionada, análise multivariada.

ABSTRACT

Our major goal is through GEOBIA (Geographic Object Based Image Analysis) obtain the characterization of use and land occupation in two different zones nearby Pitangui's river, situated among the cities of Castro, Carambeí and Ponta Grossa, Paraná State. Nowadays that is a growing demand of remote sensing images with high spatial resolution. As a result, there is a need to create new methodologies for digital image processing. Here we can mention GEOBIA, which has some advantages, such as a polygon generation for each geographical object created, relational data base with several descriptors, as well as, the possibility to use spectral, spatial and texture information. Yet we have been used the Principal Components Analysis (PCA) and the Cluster Analysis (CA) in order to reduce the size of our training samples from the relational data base, which are trained by means of supervised classification. With this, we aim to select descriptors that could bring the best results for our classification. We have obtained different geographic object oriented supervised classifications, whereas the descriptors were selected from PCA, also from PCA together with CA, as well as, all descriptors have been selected from GEOBIA. We have concluded that the best results for the two work zones were obtain from the descriptors selected from PCA together with CA. In work zone 1 we have obtained 95,68% of general precision in the confusion matrix and 0,95 in the kappa index. While, in work zone 2 we have obtained 93,85% in the confusion matrix and 0,93 in the kappa index. These numbers are consider outstanding in the literature. With this, we show that descriptors selection for geographic object oriented classification is an important approach, since there are descriptors with redundant information that could mislead the classification result.

Keywords: GEOBIA, supervised classification, multivariate analysis.

LISTA DE FIGURAS

Figura 1 - Crescimento de regiões	26
Figura 2 - Divisão e fusão	27
Figura 3 – Árvore Quadtree e seus descendentes	27
Figura 4- Localização das áreas de estudo.....	39
Figura 5- Área de estudo 1	40
Figura 6- Área de estudo 2	41
Figura 7- Compartimentos geomorfológicos da bacia hidrográfica do Pitangui	42
Figura 8 - Geologia da bacia hidrográfica do Pitangui.....	43
Figura 9- Procedimentos metodológicos.....	47
Figura 10 - Etapas de orientação a objetos A- Imagem original; B - Escala e fusão 50 e 30; C - Imagem segmentada sobreposta por polígonos gerados; D - Parte do banco de dados relacional;.....	49
Figura 11- Etapas para a classificação	54
Figura 12 - Ilustração de áreas visitadas em campo	55
Figura 13– Bandas da Imagem RapidEye para área de estudo 1	56
Figura 14 – Composições coloridas da Imagem RapidEye para área de estudo 1.....	57
Figura 15 – Chaves de interpretação para área de estudo 1.....	59
Figura 16- Áreas visitadas em campo (área de estudo 2)	60
Figura 17 – Bandas da imagem <i>RapidEye</i> para a área de estudo 2	61
Figura 18- Composições coloridas da imagem <i>RapidEye</i> para a área de estudo 2.....	63
Figura 19 – Chaves de interpretação para área de estudo 2.....	65
Figura 20- Pontos de localização das amostras de treinamento na área de estudo 1 em A e na área de estudo 2 em B.	66
Figura 21 - Procedimentos para Análise de Componentes Principais	67
Figura 22 - Procedimentos para Análise de Agrupamentos	70
Figura 23 - Dendrograma.....	71
Figura 24 – Testes de segmentação com diferentes valores de escala e fusão (área de estudo 1).....	74
Figura 25 – Área de estudo 1: A) Imagem segmentada, B) vetores, e C) Imagem segmentada sobreposta por vetores.....	74

Figura 26 – Testes de segmentação com diferentes valores de escala e fusão (área de estudo 1).....	75
Figura 27 – Área de estudo 2: A) Imagem segmentada, B) vetores, e C) Imagem segmentada sobreposta por vetores.....	76
Figura 28 - Gráfico de variância explicada pela ACP (área de estudo 1)	80
Figura 29 – Descritor espacial de direção principal - área de estudo 1	84
Figura 30- Descritor espacial de convexidade (área de estudo 1).....	84
Figura 31 – Descritor espacial de alongamento – Área de estudo 1	85
Figura 32- Descritor espacial de área - área de estudo 1.....	85
Figura 33 – Descritor de eixo maior - área de estudo 1.....	86
Figura 34 – Descritor de eixo menor - área de estudo 1	86
Figura 35 – Descritor de retangularidade - área de estudo 1	87
Figura 36- Variância explicada (área de estudo 2).....	90
Figura 37 – Descritores espaciais: 37A - Descritor de convexidade; 37B – Descritor de circuncidade; 37C – Descritor de direção principal	93
Figura 38 - Descritor espacial de convexidade.....	94
Figura 39 - Dendrograma do descritor espectral “média dos números digitais na banda 1”	98
Figura 40 –Dendrograma dos descritores espectrais “valor mínimo de pixel nas bandas 1, 2, 3, 4, 5”	99
Figura 41 – Dendrograma dos descritores “valor máximo de pixel nas bandas 1, 2, 3, 4, 5”	100
Figura 42 – Dendrograma dos descritores de textura “intervalo de textura nas bandas 1,2 e 3”	101
Figura 43– Dendrograma do descritores de “variância de textura na bandas 2”	102
Figura 44 – Dendrograma formado pelos descritores “entropia de textura nas bandas 1,3 e 5”	103
Figura 45 – Dendrograma do descritor espacial “fator de forma”	104
Figura 46 – Dendrograma do descritor espectral “Média dos números digitais das bandas 2 e 5”.....	107
Figura 47 – Dendrograma dos descritores espectrais “valor mínimo de pixel nas bandas 4 e 5”.....	108

Figura 48 – Dendrograma formado pelos descritores “valor máximo de pixel nas bandas 1 e 4”	110
Figura 49 – Dendrograma formado pelos descritor espacial “comprimento”	111
Figura 50– Classificações geradas para área de estudo 1: A- Classificação com os descritores selecionados a partir da ACP; B - Classificação com descritores selecionados a partir da ACP e AA; C - Classificação com todos os descritores;	115
Figura 51 – Classificações da área de estudo 2: A - Classificação com descritores selecionados a partir da ACP; B - Classificação com descritores selecionados a partir da ACP e AA; C - Classificação com os 54 descritores gerados	122

LISTA DE TABELAS

Tabela 1 - Autovalor e variância explicada e acumulada das componentes principais da área de estudo 1	78
Tabela 2 - Descritores de maior autovetor absoluto correspondentes as componentes principais geradas.....	80
Tabela 3 - Autovalor e variância explicada e acumulada das componentes principais da área de estudo 2.....	88
Tabela 4 – Descritores de maior autovetor absoluto correspondentes as componentes principais geradas.....	90
Tabela 5 – Matriz de confusão da classificação com os descritores selecionados a partir da ACP (área de estudo 1)	116
Tabela 6 - Áreas (m ²) da classificação com os descritores selecionados a partir da ACP (área de estudo 1)	117
Tabela 7– Matriz de confusão da classificação com os descritores selecionados a partir da ACP + AA (área de estudo 1).....	117
Tabela 8 – Matriz de confusão da classificação com os 54 gerados (área de estudo 1)	118
Tabela 9- Áreas (m ²) das classificações com os descritores selecionados a partir da ACP+AA e todos os 54 descritores obtidos mediante a GEOBIA (área de estudo 1)	118
Tabela 10 – Matriz de confusão da classificação realizada com os descritores selecionados mediante a ACP para a área de estudo 2.....	123
Tabela 11 - Áreas da classificação com os descritores selecionados mediante ACP (área de estudo 2)	123
Tabela 12 – Matriz de confusão da classificação realizada com descritores selecionados mediante ACP + AA para a área de estudo 2.....	125
Tabela 13 - Áreas (m ²) da classificação com os descritores selecionados mediante ACP e AA (área de estudo 2).....	125
Tabela 14 – Matriz de confusão para classificação com todos os descritores para a área de estudo 2.....	126
Tabela 15 - Áreas (m ²) da classificação com os 54 descritores gerados (área de estudo 2)	126

LISTA DE QUADROS

Quadro 1 – Especificações técnicas dos satélites do sistema RapidEye	46
Quadro 2 - Bandas espectrais das imagens RapidEye	46
Quadro 3 – Descritores de textura selecionados.....	82
Quadro 4 – Descritores de textura selecionados.....	92
Quadro 5 – Descritores selecionados mediante Análise de Agrupamentos para a área de estudo 1	95
Quadro 6 – Descritores selecionados mediante Análise de Agrupamentos para a área de estudo 1	105
Quadro 7 – Testes e descritores para a classificação supervisionada orientada a objetos.....	113
Quadro 8- Descritores utilizados para cada teste de classificação (ACP, ACP+AA, Todos) para a área de estudo 2.....	120
Quadro 9 – Acurácia dos testes de classificação para a área de estudo 2.....	121

Sumário

1	Introdução.....	14
2	Revisão Teórica.....	16
2.1	O Geoprocessamento e as Geotecnologias na Geografia.....	16
2.2	Análise Orientada a Objeto Geográfico (GEOBIA – <i>Geographic Object Based Image Analysis</i>).....	20
2.2.1	Segmentação.....	24
2.3	Classificação Digital Supervisionada.....	28
2.4	Análise Multivariada.....	30
2.4.1	Análise de Componentes Principais.....	32
2.4.2	Análise de Agrupamentos.....	35
3	Materiais e métodos.....	38
3.1	Localização e caracterização da área de estudo.....	38
3.1.1	Características da bacia hidrográfica do Pitangui.....	41
3.2	Materiais.....	45
3.2.1	Imagem RapidEye.....	45
3.2.2	<i>Softwares</i>	46
3.3	Procedimentos metodológicos.....	46
3.3.1	Análise Orientada a Objeto Geográfico (GEOBIA – <i>Geographic Object Based Image Analysis</i>).....	48
3.3.2	Classificação supervisionada orientada a objeto geográfico.....	54
3.3.3	Análise de Componentes Principais.....	67
3.3.4	Análise de Agrupamentos.....	69
3.3.5	Validação.....	71
4	Resultados e discussão.....	73
4.1	Análise Orientada a Objeto Geográfico (GEOBIA – <i>Geographic Object Based Image Analysis</i>).....	73
4.3	Seleção de descritores para a classificação.....	77
4.3.1	Seleção de descritores mediante Análise de Componentes Principais (ACP) ..	77
4.3.2	Seleção de descritores mediante a Análise de Agrupamentos (AA).....	95
4.4	Classificação e validação do uso e ocupação da terra.....	112
4.4.1	Classificação e validação do uso e ocupação da terra para a área de estudo1.	112
4.4.2	Classificação e validação do uso e ocupação da terra para a área de estudo2.	119
5	Considerações finais:.....	128
6	Referências:.....	130
7	Anexo.....	139

1 Introdução

A ciência geográfica necessita em seus estudos de metodologias e técnicas, nesse sentido destaca-se o Geoprocessamento e as Geotecnologias, com ênfase nos Sistemas de Informação Geográfica, Cartografia Digital, Processamento de Imagens Digitais e Sensoriamento Remoto que possibilitam gerar informações auxiliando a interpretação do analista.

As Geotecnologias como o Sensoriamento Remoto (SR), Processamento Digital de Imagens (PDI), e Sistemas de Informação Geográfica (SIG) proporcionam estudos mediante técnicas matemáticas e computacionais, que permitem o levantamento de informações qualitativas e quantitativas. E são pertinentes nas várias correntes teóricas da Geografia.

Tornando possível integração maior entre Sensoriamento Remoto e Sistemas de Informação Geográfica e atender a demanda de altas resoluções de imagens orbitais surge a GEOBIA (*Geographic Object Based Image Analysis*). Considera-se esta como uma nova metodologia no processamento das imagens de Sensoriamento Remoto, pois utiliza-se de informações espectrais, espaciais e de textura, o que fornece suporte ao trabalho do intérprete na geração de novas informações geográficas em diversas áreas como é o caso do objeto de estudo desse trabalho: uso e ocupação da terra.

Uma das características dessa nova metodologia (GEOBIA ou Análise Orientada a Objeto Geográfico) é a geração de objetos geográficos que podem ser mais próximos à realidade que *pixels* agregados, e banco de dados relacional para cada objeto geográfico gerado contendo descritores espectrais, espaciais e de textura.

Os objetos geográficos gerados e o banco de dados relacional são úteis para a extração de informações e por conter diversos descritores representando, por exemplo, tonalidades, forma, textura. Para compreender e tratar este rol de descritores o uso de Análises Multivariadas se faz pertinente.

Nesse trabalho foram utilizadas as Análises Multivariadas: Análise de Componentes Principais e Análise de Agrupamentos. A primeira com o objetivo maior de reduzir a dimensionalidade dos dados e posteriormente a seleção de descritores e a segunda para complementar a seleção de descritores para a classificação supervisionada orientada a objetos geográficos da imagem *RapidEye*.

O objetivo desse trabalho foi classificar o uso e ocupação da terra de duas diferentes áreas da bacia hidrográfica do rio Pitangui, utilizando a abordagem de

orientação a objetos geográficos (GEOBIA), e buscando classificação ótima mediante a seleção de descritores para a classificação, e para isso utilizou-se das técnicas de Análise Multivariada.

A área de estudo desse trabalho foi escolhida devido a sua importância para a região em que se localiza.

O rio Pitangui localiza-se entre os municípios paranaenses de Castro, Carambeí e Ponta Grossa. Se trata de um rio antecedente que atravessa diferentes compartimentos geológicos e geomorfológicos, suas águas são responsáveis pelo abastecimento da cidade de Ponta Grossa, Paraná. Há diversos usos e ocupações da terra na sua bacia hidrográfica, dentre eles podemos citar a área urbana, pois abrange a parte norte da área urbana de Ponta Grossa e todo o município de Carambeí; também há diversos cultivos em sua área de extensão, e grandes áreas de reflorestamentos. Além disso há elementos de grande importância paisagística para a região como *canyons* e a represa de Alagados.

Escolheu-se para realizar o trabalho duas pequenas áreas no entorno do rio Pitangui, essas áreas possuem diversos usos significativos da área rural da bacia como é o caso dos cultivos e reflorestamentos.

Pretende-se com o esse trabalho contribuir no melhor entendimento dos usos que se faz nessas duas áreas do rio Pitangui, além de buscar por uma metodologia de classificação de imagens de alta resolução, mediante a abordagem de classificação supervisionada orientada a objetos e utilizando a seleção de descritores por meio de Análises Multivariadas.

2 Revisão Teórica

Na revisão teórica serão explanados alguns tópicos que nortearam o desenvolvimento do trabalho: O Geoprocessamento e as Geotecnologias; Análise Orientada a Objeto Geográfico; Classificação digital supervisionada e Análise Multivariada.

O Geoprocessamento e as Geotecnologias abrangem diversas técnicas e tecnologias as quais são de uso interdisciplinar. Além disso, cada vez mais essas técnicas são usadas na ciência geográfica contribuindo para diversos estudos.

A Análise Orientada a Objeto Geográfico e classificação digital supervisionada são técnicas de processamento de imagens digitais de sensoriamento remoto, utilizadas nesse trabalho para caracterizar o uso e a ocupação da terra e são descritas nas seções 3.2 e 3.3.

As técnicas de Análise Multivariada foram utilizadas nesse trabalho porque possibilitam examinar grandes conjuntos de dados, utilizou-se a Análise de Componentes Principais e Análise de Agrupamentos no banco de dados relacional da GEOBIA, as quais são explicadas em detalhe na seção 3.4.

2.1 O Geoprocessamento e as Geotecnologias na Geografia

A necessidade de conhecer, detalhar e intervir no espaço geográfico buscando compreendê-lo e culminar em articulações entre a sociedade e o ambiente condiciona a busca de novas técnicas e instrumentos em diversas áreas do conhecimento. Neste contexto, um dos grandes impulsionadores são as tecnologias.

As diversas tecnologias têm desenvolvimento constante, gerando grande quantidade de dados a respeito do espaço físico territorial. Esta disponibilidade de dados necessita, muitas vezes, de transformações e processamentos para produzir informações e gerar conhecimento a respeito do espaço geográfico e fornecer suporte técnico científico em planejamentos e/ou gerenciamentos no que tange à tomada de decisões.

Nesse contexto, há destaque para o Geoprocessamento e as Geotecnologias, técnicas consagradas nos estudos geográficos atuais.

De acordo com Câmara e Davis (2001, p. 1) o Geoprocessamento:

[...] denota a disciplina do conhecimento que utiliza técnicas matemáticas e computacionais para o tratamento da informação geográfica e que vem influenciando de maneira crescente as áreas de Cartografia, Análise de Recursos Naturais, Transportes, Comunicações, Energia e Planejamento Urbano e Regional.

Câmara & Monteiro (2004, p. 1) mencionam que o Geoprocessamento é “tecnologia interdisciplinar que permite a convergência de diferentes disciplinas científicas para o estudo de fenômenos ambientais e urbanos”.

Segundo Vetorazzi (1996, p. 45) “o Geoprocessamento relaciona-se as mais variadas técnicas empregadas na coleta, armazenamento, processamento, análise e representação de dados com expressão espacial, isto é, possíveis de serem referenciados geograficamente (georreferenciados)”. Cita ainda que essas técnicas podem ser a Topografia convencional, com o emprego de instrumentos simples como trena e bússola, até a utilização de satélites de posicionamento e imageamento.

O Geoprocessamento tem uso multidisciplinar, ou seja, vários ramos do conhecimento se utilizam de suas técnicas. Na Geografia, tem sido muito utilizado e tem uma grande relação conceitual. Os autores Câmara, Davis e Monteiro (2001) analisam o conceito de espaço geográfico em diferentes correntes geográficas buscando estabelecer como o Geoprocessamento e suas tecnologias permitem a expressão destes conceitos.

Com relação à Geografia Idiográfica analisou-se a proposta de Hartshorne e o seu conceito de “unidade de área” equivalente também aos conceitos de unidade de paisagem (Tricart, 1977) e land-unit (Zonneveld, 1989), todos baseados no conceito-chave de delimitação de unidades homogêneas. No Geoprocessamento, a representação computacional de “unidade área” corresponderia ao polígono fechado, que delimita cada região de estudo e um conjunto de atributos, tipicamente armazenados num banco de dados relacional. (CÂMARA, DAVIS E MONTEIRO, 2001).

A corrente chamada “Geografia Teórica e Quantitativa” ou “Geografia Neopositivista” teve origem após o fim da Segunda Guerra Mundial, quando ocorreram modificações teóricas e metodológicas na ciência geográfica. É caracterizada pelo uso de técnicas matemático-estatísticas e tem como base o neopositivismo que almejava a caracterização geral e definitiva da ciência, dos métodos apropriados e de seus critérios de avaliação ao invés da subjetividade e da imprecisão dos conceitos. (VITTE, 2007).

De acordo com Vitte (2007) foi com a Geografia Teórica e Quantitativa que se iniciou o uso da modelagem e suas variáveis, e para o seu tratamento o uso do

computador teve papel fundamental favorecendo o armazenamento e correlação de dados, facilitando a manipulação de volumes maiores e condições de atualizações.

Sendo assim, pode-se dizer que há forte relação entre os conceitos do Geoprocessamento e da concepção da Geografia Quantitativa, uma vez utilizam técnicas matemáticas e estatísticas e de ferramentas computacionais.

Na Geografia Crítica, o conceito-chave é o espaço como “sistema de objetos e sistemas de ações” e a oposição entre “espaço de fluxos” e “espaço de lugares”. Essa corrente da ciência geográfica critica a Geografia Quantitativa por explicar com ênfase no uso de grandezas mensuráveis os processos socioeconômicos (CÂMARA, DAVIS E MONTEIRO, 2001).

Os autores supracitados fazem uma especulação a respeito das representações computacionais na perspectiva da Geografia Crítica:

Podemos apenas especular sobre as representações computacionais que serão utilizadas neste contexto, que possivelmente incluam técnicas de Representação de Conhecimento. Nossa hipótese de trabalho é que os GIS do futuro contemplarão representações não-cartográficas do espaço, com uma ênfase no estabelecimento de relações entre os diferentes atores sociais que atuam no espaço. (CÂMARA, DAVIS E MONTEIRO, 2001, p. 12-13).

O Geoprocessamento é de relevância nos estudos geográficos, pois possibilita que os conceitos geográficos sejam representados computacionalmente de modo a dar subsídio a análises, e facilitar a interpretações do geógrafo.

Com o avanço das tecnologias nas ciências em geral, as técnicas de Geoprocessamento também se desenvolveram, com isso surgiu o conceito de “Geotecnologias”.

De acordo com Rosa (2005, p. 81), as Geotecnologias “são o conjunto de tecnologias para coleta, processamento, análise e oferta de informações com referência geográfica”, e o Geoprocessamento o conjunto de técnicas que representam qualquer tipo de processamento georreferenciado.

Estão inclusos nas Geotecnologias, segundo Rosa (2005), a Cartografia Digital, o Sensoriamento Remoto (SR), o Processamento de Imagens Digitais (PDI), os Sistemas de Informação Geográfica (SIG's), a Topografia Digital ou Automatizada, e o Sistema Global de Navegação por Satélite (GNSS – *Global Navigation Satellite System*).

O Sensoriamento Remoto (SR) constitui-se como ciência e arte de adquirir informações sobre objetos e fenômenos sem contato com esses, tipicamente utiliza regiões do espectro eletromagnético e instrumentos localizados em plataformas móveis. Quanto à forma de operação dos sensores, o SR pode ser classificado em passivo e ativo, sendo que o primeiro deles utiliza apenas a energia natural refletida ou emitida a partir da superfície terrestre, e o ativo utiliza energia artificial, produzida por radares instalados nos próprios sensores (FIGUEIREDO, 2005).

O SR possibilita visão sinóptica e panorâmica de dada região e condições homogêneas para observação de fenômenos. É uma forma eficaz de monitoramento em escala global e local, pois possibilita rapidez e periodicidade no fornecimento de imagens e informações (JENSEN, 2009).

De acordo com Gelli (2001), tradicionalmente a extração de informações das imagens de SR pelas equipes técnicas de geociências foi feita por métodos analógicos. No entanto, as evoluções no campo da informática permitem crescente utilização do processamento digital. Os dados adquiridos por SR são transformados em informações e imagens mais simples, por meio de técnicas de Processamento de Imagens Digitais (PDI) e de interpretação e reconhecimento de padrões, utilizando *hardwares* e *softwares*.

O PDI envolve processos cujas entradas e saídas são imagens e processos de extração de atributos de imagens até o reconhecimento de objetos individuais. (GONZALEZ; WOODS, 2000).

Ainda em se tratando de Geotecnologias na Geografia os Sistemas de Informação Geográfica (SIG's) se destacam nos processamentos e possuem função de:

[...] integrar, numa única base de dados, informações espaciais provenientes de dados cartográficos, dados de censo e de cadastro urbano e rural, imagens de satélites, redes, dados e modelos numéricos de terrenos; combinar as várias informações, mediante algoritmos de manipulação, para gerar mapeamentos derivados; consultar, recuperar, visualizar e imprimir o conteúdo da base de dados geocodificados (CÂMARA; ORTIZ, 199?, p. 1).

Os Sistemas de Informação Geográfica e o Sensoriamento Remoto (SR) são adequados para estudos do meio ambiente, pois proporcionam a manipulação, organização e transformação dos dados em informações, facilitando a interpretação e ações.

Os SIG's como ressalta Rego (2007, p. 84):

[...]descreve as formas e a estrutura da paisagem em um ou vários momentos, mas a análise é do analista geográfico, bem como a interpretação e a derivação das funções e dos processos subjacentes, permitindo o entendimento do conjunto expresso pelo espaço geográfico.

O Sensoriamento Remoto fornece dados de entrada que limitam a escala de trabalho e o SIG é interface de ligação e manipulação dos dados de entrada e/ou saída, ambos estão contidos no conjunto de técnicas do Geoprocessamento e Geotecnologias. Possibilitam ampliar o conhecimento geográfico tornando-o menos subjetivo, mais quantitativo e passível de análises numéricas e estatísticas, proporcionando técnicas que fornecem aporte nos estudos e compreensão do espaço geográfico.

Lang e Blaschke (2009) designam SIG e SR como as mais importantes técnicas holísticas para a análise, planejamento e gestão da paisagem.

As Geotecnologias como o Sensoriamento Remoto (SR), Processamento de Imagens Digitais (PDI), e Sistemas de Informação Geográfica (SIG) possibilitam aporte para vários estudos mediante técnicas matemáticas e computacionais, que permitem o levantamento de informações qualitativas e quantitativas que são cruciais para o entendimento e planejamento.

2.2 Análise Orientada a Objeto Geográfico (GEOBIA – *Geographic Object Based Image Analysis*)

Devido a maior acessibilidade e crescente volume de imagens de satélite de alta resolução surge a necessidade de novos procedimentos para o Processamento de Imagens Digitais (PDI), diante disso abordagens baseadas em objetos ganham destaque.

A GEOBIA – Análise Orientada a Objeto Geográfico é metodologia no processamento das imagens de Sensoriamento Remoto, em que o nível de abstração deixa de ser o *pixel* para ser o objeto geográfico.

Objetos geográficos são constituídos por agregados de *pixels* próximos espacialmente e com características similares. De acordo com Hay e Castilla (2008), o substantivo geográfico implica necessariamente na representação de uma porção da superfície terrestre. Um objeto geográfico pode ser definido dentro da GEOBIA como uma delimitada região geográfica que pode ser identificada em um período de tempo como o referente de um termo geográfico. Os termos geográficos nesse sentido são as

típicas legendas contidas nos mapas, como, por exemplo, as classes de uso e ocupação da terra: corpos hídricos, floresta, área urbana etc.

A Análise de Imagem Baseada em Objeto (OBIA – *Object-Based Image Analysis*) é usada em muitas áreas diferentes (Biomedicina, Astronomia, Microscopia, Visão Computacional etc) e agora é usada no Sensoriamento Remoto mediante PDI, neste, optando-se pelo termo GEOBIA com o pseudo-prefixo Geo como qualificador, para enfatizar o componente geográfico e o foco principal em Sensoriamento Remoto e Sistemas de Informação Geográfica (HAY & CASTILLA, 2008).

No caso da GEOBIA, o objetivo primordial é:

“o desenvolvimento da teoria, métodos e ferramentas suficientes para replicar e/ou exceder a interpretação humana de imagens de Sensoriamento Remoto por meio de caminho auto e semi-automatizados, resultando em um processo mais preciso, reproduzível, com menos subjetividade, tempo e custos. Não sendo então o da construção de novas ferramentas, mas sim a geração de informação geográfica permitindo aos usuários interpretar e responder a algumas perguntas como os usos e ocupações da terra, a gestão dos recursos naturais, entre outras” (HAY, G.J., CASTILLA, G., 2008, p. 6).

Os autores supracitados destacam a existência de diversas vantagens da GEOBIA: particionar a imagem em objetos semelhantes à forma como os seres humanos conceituam e organizam a paisagem para compreendê-la; objetos de imagem possuem características úteis (por exemplo, forma, textura, contexto e relações com outros objetos) que não podem ser extraídos de *pixels* individuais; maior facilidade na integração com SIG. Ainda há a possibilidade de que com esse procedimento possam ser obtidos delineamentos de objetos geográficos mais precisos.

Hay e Castilla (2008) citam vários fatores que contribuíram para o desenvolvimento da análise de imagens orientada a objetos geográficos (GEOBIA): a política de comercialização de imagens de alta resolução espacial; ferramentas de computação cada vez mais acessíveis; e o reconhecimento das limitações nas abordagens em nível de *pixels*, que não contemplam contexto, textura e forma.

Os autores Lang e Blaschke (2009) argumentam que o aumento da resolução espacial, junto ao aumento exponencial do nível de detalhe, demanda técnicas de interpretação e análise que permitam trabalhar adequadamente os detalhes das imagens orbitais, possibilitando analisar objetivamente os seus conteúdos essenciais.

Dessa forma, altas resoluções espaciais exigem novos processamentos que extrapolem a observação de *pixels* individuais. Por esta razão, no processamento digital de dados de SR está ocorrendo grande revolução metodológica:

Com a alta resolução onde um *pixel* não é mais automaticamente maior/igual aos objetos pesquisados, subitamente não são procurados somente desenvolvimentos contínuos de métodos de classificação, porém novas concepções metodológicas. O usuário é surpreendido não somente pela quantidade de dados, mas também pelo grau de detalhes das informações. Este último aspecto exige novos caminhos para a regionalização da informação, abstração conduzida pela teoria e o relacionamento de *pixels* individuais em construções lógicas. (BLASCHKE, GLÄSSER, e LANG, 2007, p. 17).

Diante desse contexto, Lang e Blaschke (2009) mencionam os procedimentos não baseados em *pixels* unitários, que são exclusivamente estatísticos e não utilizam a dimensão espacial da informação.

Para imagens de média e alta resolução, a orientação a objetos é sugerida na literatura como a mais indicada, pois considera não só a resposta espectral dos alvos, mas também suas variantes geométricas, topológicas e de textura (ANTUNES, 2003).

Para Venturieri e Santos (1998), a análise baseada em regiões é um processo de preparo para as imagens digitais a uma futura classificação temática considerando os segmentos, ou orientação a objetos, no qual os elementos analisados são regiões resultantes do agrupamento de *pixels*. A orientação de objetos divide uma imagem em partes, ou, como o nome supõe, em objetos que a constituem e o processo é finalizado quando os objetos de interesse na aplicação estiverem isolados.

De acordo com Hay e Castilla (2008), a GEOBIA é sub-área da Ciência de Informação Geográfica. Propõe-se ao desenvolvimento de métodos automatizados para particionar imagens de Sensoriamento Remoto em objetos de imagem significativos, avaliar as suas características mediante análises espaciais, espectrais e de textura de modo a gerar nova informação geográfica e topologia em formatos compatíveis aos SIG's. Para isso, utiliza os dados de SR e gera saídas prontas para os SIG's como polígonos e banco de dados relacional.

SIG e SR se complementam em estudos do espaço físico-territorial e do uso e da ocupação da terra e a GEOBIA contribui nisso, pois representa forte desenvolvimento adicional na área de análise automática de imagens. Constitui ligação entre os métodos do SR e do processamento de informação geográfica, entre o domínio *raster* (imagem)

do SR e o domínio vetor de SIG, pois já disponibiliza como saída para o SIG informações no formato de polígonos (ou seja, objetos da imagem) representando objetos geográficos (HAY; CASTILLA, 2008; LANG; BLASCHKE, 2009; BLASCHKE, GLÄSSER; LANG, 2007).

A interpretação da cena em que é possível a inclusão de informação espectral, espacial e de textura é importante, já que mais parâmetros podem ser derivados para descrever os elementos (SOLBERG et al. 1996).

Tso e Mather (2009) mencionam que a adição de informações auxiliares (não espectrais), tais como as relações espaciais, consistem em uma das maneiras mais poderosas de caracterizar as classes de interesse. Informações auxiliares podem ser extraídas da própria imagem, como a textura e o contexto espacial, e também se adicionando modelos digitais de terreno, mapas de solos, geologia, índices de vegetação, entre outros. A acurácia da classificação pode ser melhorada em relação a processos de classificação de informações espectrais exclusivamente.

Conforme Fan e Xia (2001), a dimensão espectral refere-se às diferentes bandas do espectro eletromagnético, o contexto espacial é definido pela correlação espacial entre os *pixels* adjacentes e as suas vizinhanças. Tanto a dimensão espectral quanto o contexto espacial são utilizados na metodologia de orientação a objetos e torna o procedimento vantajoso se comparado aos que somente utilizam dados espectrais dos *pixels*.

A GEOBIA possibilita a inclusão de dados adicionais aos espectrais, o que vem agregar mais atributos para melhor descrever cada região e assim gerar informações geográficas mais completas que somente a radiância refletida dos alvos.

De acordo com os trabalhos de Blaschke, Glässer e Lang (2007), Lang e Blaschke (2009), faltam ainda fundamentações metodológicas nos processamentos de imagens orientadas a objetos. O desafio metodológico consiste na extração de informações relevantes a partir de imensa quantidade de informações genéricas, que são os dados brutos. Portanto, são necessários procedimentos de redução de complexidade, ou seja, de regionalização.

Apesar da enorme capacidade das concepções automáticas, elas ainda não alcançam a capacidade de interpretação da detecção humana. Ainda que a máquina apresente nítidas capacidades na diferenciação de níveis digitais e no processamento de informações digitais, o intérprete humano apresenta maior superioridade para a detecção de estruturas e de informações relacionadas ao contexto. Adicionalmente, um

intérprete caracteriza-se ainda pelo conhecimento e experiência armazenados, que aumentam gradualmente a cada nova análise de imagem. Concepções de análise baseada em objetos ou contexto, devido a limitações inerentes dos algoritmos e do conjunto correspondente dos segmentos disponibilizados, ainda não estão em condições de considerar “limites concepcionais” (LANG E BLASHKE, 2009, p. 157).

Embora haja limitações à GEOBIA, ela se apresenta como uma nova metodologia na análise de imagens de sensoriamento remoto, possibilitando maior integração com os SIG's e o uso de dados auxiliares. Fornece suporte ao trabalho do intérprete na geração de novas informações geográficas em diversas áreas como é o caso de monitoramento e planejamento ambiental, e do levantamento do uso da terra.

Diversos trabalhos utilizam a análise orientada a objetos em diferentes estudos.

Os autores Abbas et al. (2010) realizaram um estudo com o objetivo de desenvolver um inventário de cobertura florestal úmida temperada do Parque Nacional Ayubia, no Paquistão, mediante imagens de satélite de alta resolução (0,6m). Utilizaram a análise orientada a objetos geográficos para poder trabalhar com multi-escalas. Os objetos da paisagem representando componentes homogêneos foram delimitados utilizando rotinas de segmentação, que permitem a inclusão de elementos espaciais. Uma árvore de classificação foi desenvolvida e utilizou-se o classificador do vizinho mais próximo.

Arroyo, Johansen e Phinn (2010) realizaram um estudo com a meta de aplicar GEOBIA para mapear automaticamente classes de cobertura terrestre área de estudo localizada a 50 quilômetros de Melbourne, Victoria, Austrália. Eles utilizaram dados de imagem em cinco subconjuntos que mostram diferentes ambientes e características. Quatro delas foram usadas para criar um método de classificação GEOBIA para mapear os tipos de cobertura da terra. Foram utilizadas segmentação e classificação para estabelecer uma base de conhecimento contextual. Assim, as características do contexto tornaram-se útil para classificação de tipos de cobertura da terra. A acurácia dos quatro mapas de cobertura do solo utilizados para desenvolver o esquema de classificação foi de 77,5%.

2.2.1 Segmentação

O processo de segmentação é a base da GEOBIA. A segmentação realiza o particionamento da imagem, tentando imitar o processo cognitivo humano em que

segmentar é natural e fácil, uma vez que os sistemas biológicos reconhecem e interpretam os objetos assim que os veem a partir de um conjunto de informações que são combinadas e processadas em paralelo no cérebro (CONCI, 2014).

Segundo os autores Robinson, Redding e Crisp (2002), a tarefa primordial da segmentação de imagens é extrair objetos de interesse. O resultado depende da escala e critérios de semelhança, criando objetos maiores ou menores, homogêneos ou não homogêneos, suavizados ou não. A segmentação deve parar quando os objetos ou regiões de interesse forem detectados.

Os algoritmos de segmentação para imagens geralmente estão baseados em uma de duas categorias básicas relacionadas às propriedades dos valores de intensidade: descontinuidade e similaridade.

Na descontinuidade, “o pressuposto é que as fronteiras das regiões são suficientemente diferentes entre si e em relação ao fundo da imagem para permitir a detecção de limite com base nas descontinuidades locais em intensidade” (GONZALEZ E WOODS, 2000, p. 455).

Nessa abordagem, a segmentação da imagem acontece com base em mudanças locais abruptas de intensidade, há três tipos de características de imagem que se procura detectar: os pontos isolados, as linhas e as bordas. Conforme Gonzalez e Woods (2000), a segmentação baseada nas bordas é a mais utilizada no enfoque de descontinuidade, e é utilizada nesse trabalho.

Conforme Seara e Elizandro (2005), a detecção de bordas é uma das técnicas básicas utilizadas pela visão humana no reconhecimento de objetos. É o processo de localização e realce dos *pixels* de borda, aumentando o contraste entre a borda e o fundo.

A maioria dos algoritmos de segmentação com a finalidade de detectar descontinuidade utiliza uma máscara sobre os *pixels* da imagem.

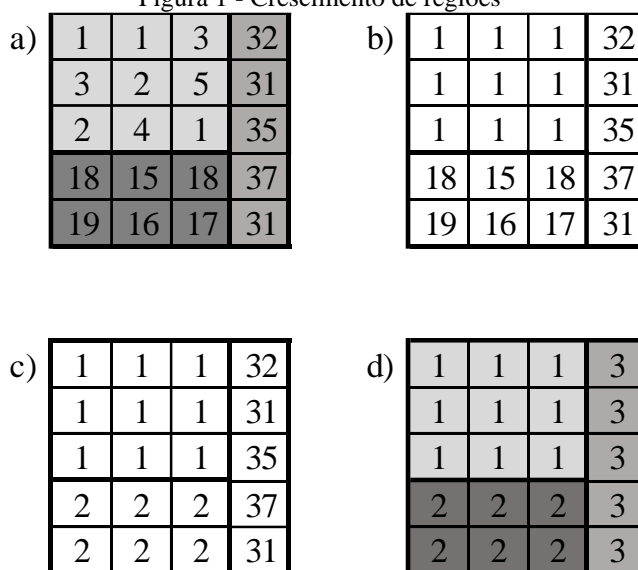
Cada pixel e seus *pixels* vizinhos tem um valor do nível de cinza multiplicado por uma constante. A soma destes valores representa a máscara de resposta daquele ponto. [...] Ampliando esta máscara atravessando a imagem linha por linha, um novo vetor é criado. É o mesmo que a imagem original, mas contém os valores da máscara de resposta ao invés do valor do pixel. Estes valores da máscara de resposta podem então ser comparados com o valor mínimo de *threshold*¹ para determinar quais *pixels* são mais prováveis de serem parte de uma borda. Este *threshold* pode ser ajustado para variar seletivamente de acordo com os *pixels* de borda, permitindo um usuário conduzir o algoritmo para um melhor desempenho para uma imagem específica (WANGENHEIM, 2005, P. 13-14).

As técnicas de similaridade são baseadas na divisão da imagem em regiões que sejam semelhantes de acordo com um conjunto de critérios pré-estabelecidos. A limiarização, o crescimento de regiões, a divisão e fusão de regiões são exemplos dos métodos desta categoria (Gonzalez e Woods, 2000).

Segundo Gonzalez e Woods (2000), o crescimento de regiões é uma técnica de agrupamento espacial no qual somente os *pixels* adjacentes, contíguos na imagem, podem ser agrupados.

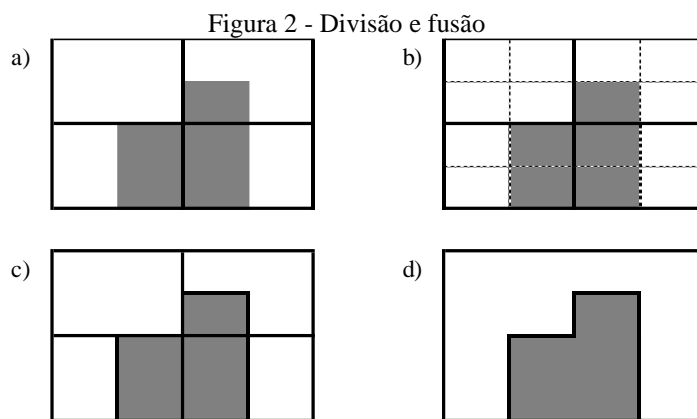
Esta técnica é ilustrada na Figura 1 e consiste em partir de um ou vários *pixels* considerados uniformes, analisar os *pixels* na vizinhança e adicionar os vizinhos que forem similares ao grupo inicial, de maneira que a região formada cresça num processo iterativo que termina quando as fronteiras naturais dos objetos presentes na cena forem atingidas. Ao final do processo, a imagem encontra-se dividida em várias sub-regiões. A região cresce por meio da agregação de *pixels* com características similares, como cor e brilho.

Figura 1 - Crescimento de regiões



Fonte: Adaptado de Gonzalez e Woods (2000).

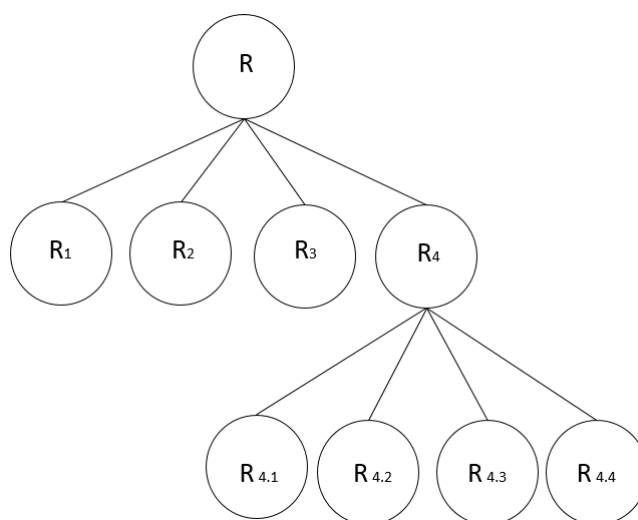
O método de divisão e fusão é ilustrado na Figura 2 e subdivide a imagem em quatro blocos, testando cada um destes blocos e verificando se os *pixels* pertencentes aos mesmos atendem a algum critério de homogeneidade. Os blocos que atenderem ao critério não serão mais divididos. O bloco que não atender será subdividido em blocos menores. O processo é repetido até um critério de parada ser satisfatório, em seguida, é realizada a junção dos blocos vizinhos que sejam homogêneos. (GONZALEZ; WOODS, 2000).



Fonte: Adaptado de Gonzalez e Woods (2000).

De acordo com Gonzalez e Woods (2000), esse método geralmente é associado a uma árvore em que cada nó possui exatamente quatro descendentes (*quadtree*), conforme explicita a Figura 3, o que possibilita decompor e agrupar partes de uma imagem.

Figura 3 – Árvore *Quadtree* e seus descendentes



Fonte: Gonzalez e Woods (2010, p. 505).

Programas computacionais oferecem módulos que realizam processamentos para a segmentação e orientação a objetos geográficos. Caso do Ecognition, da empresa Definiens, que utiliza o algoritmo FNEA (*Fractal Net Evolution Approach*); e do ENVI, da empresa Exelis, que é um *software* utilizado nesse trabalho e tem implementado os algoritmos EDGE e FLSA (*Full Lambda Schedule Algorithm*).

O algoritmo EDGE identifica características com limites distintos, calcula uma imagem gradiente usando o método de detecção de borda Sobel, onde os mais altos valores de pixel representam áreas com o maior contraste de pixel. Após isso, se aplica

o algoritmo *watershed* à imagem gradiente em vez da imagem original, aplicando “inundação” à imagem inicial com os menores valores de gradiente (a parte uniforme dos objetos) para os mais altos valores de gradiente (nas bordas) (EXELISVIS, 20--).

O algoritmo FLSA foi criado por Robinson, Redding e Crisp (2002), com base em Koepfler et. al. (1994), para ser utilizado com imagens SAR (*Syntetic Aperture Radar*). No entanto, sua aplicação é generalizada. Com o enfoque de detecção de bordas e posterior fusão de regiões, o FLSA divide a imagem em regiões para depois fundir regiões conexas, quando o valor do custo de fusão superar o limite estabelecido.

De acordo com Seara e Elizandro (2005), o processo de segmentação é essencialmente empírico e deve se ajustar a diferentes tipos de imagem, não havendo modelo formal para a segmentação, que se caracteriza como a etapa mais difícil e delicada do processo, pois todas as medidas serão realizadas sobre as regiões identificadas nela.

Não existe um método de segmentação comum para todas as imagens, tanto pela diversidade de aplicações do PDI quanto pela ligação desse método com os conceitos subjetivos do analista (ANTUNES, 2003).

A segmentação se realiza de acordo com o objetivo da análise e conhecimento do intérprete, a escolha da escala e critérios de semelhança estão relacionados com as resoluções espacial e espectral da imagem, além do “olhar” do técnico.

No próximo tópico, será discorrido sobre a classificação digital de imagens, sendo este método utilizado nesta pesquisa para culminar na caracterização do uso e ocupação da terra.

2.3 Classificação Digital Supervisionada

A classificação consiste no reconhecimento de padrões, sendo que padrão é um arranjo de descritores ou características. Segundo Gonzalez e Woods (2000), uma classe de padrões é uma família de padrões que compartilham propriedades comuns.

Ribeiro (2003) menciona que a classificação digital tem como objetivo ordenar, discriminar e categorizar os objetos ou *pixels* em classes, baseando-se em características espectrais dos objetos.

Com os métodos de classificação e processamento de imagens digitais é possível identificar as áreas críticas afetadas por ação natural ou antrópica, facilitando o

planejamento e ações políticas com diretrizes para a gestão destas áreas. (RIBEIRO, 2013).

Na classificação de imagem, utiliza-se a fotointerpretação ou análise visual de imagens. A fotointerpretação é de suma relevância para a classificação digital, a partir dela são definidos os padrões de cada classe, por meio das amostras de treinamento, juntamente com uma verdade de campo, mapa ou material bibliográfico da área de estudo.

A fotointerpretação baseia-se em elementos de interpretação de imagem que foram desenvolvidos mediante experiências empíricas. “Os elementos de interpretação de imagem incluem localização, tonalidade e cor, tamanho, forma, textura, padrão, sombra, altura e profundidade, volume, declividade, aspecto, sítio, situação e associação” (JENSEN, 2009, p. 133).

Segundo Moreira (2011), na classificação digital procura-se rotular cada *pixel* da imagem segundo a forma de uso e ocupação da terra, semelhante ao que se faz na abordagem visual. Na classificação orientada a objetos, ao invés de rotular o pixel, rotula-se o objeto geográfico ou a região individualizada por meio da segmentação de imagem. A rotulação dos valores dos níveis digitais é feita mediante algoritmos de classificação que realizam o reconhecimento dos padrões espectrais. A classificação pode ser supervisionada ou não-supervisionada.

Na classificação digital não supervisionada, o usuário não define as classes de interesse, somente fixa o número de classes que deseja obter e os algoritmos de processamento de imagens se encarregam de separar os *pixels* na imagem em grupos uniformes (CENTENO, 2004).

Vários trabalhos atualmente usam técnicas de classificação digital, como é o caso de Oliveira et. al. (2013) que, por meio da classificação supervisionada pelo algoritmo da máxima verossimilhança (MAXVER), mapeia fragmentos florestais, e Junges, Fontana e Pinto (2013) que utilizaram a classificação não supervisionada de índice de vegetação para identificação de áreas de cultivo.

De acordo com Centeno (2004), a classificação supervisionada combina a análise visual com a técnica digital, utiliza a capacidade interpretativa do técnico, este definirá o número de classes e selecionará amostras de treinamento. Estas amostras são conjuntos de *pixels* representativos e puros selecionados de cada classe na imagem, a partir dessas o algoritmo realiza a generalização do padrão para a cena toda, culminando na classificação da imagem.

A classificação digital supervisionada segue as seguintes fases: definição de classes por parte do técnico, coleta das amostras de treinamento, verificação da acurácia das amostras de treinamento, classificação pelo algoritmo estatístico, e verificação da qualidade da classificação.

Dessa forma podemos concluir que a classificação digital oferece vantagens como, por exemplo, o uso de todas as bandas espectrais, o que é impossível para os olhos do ser humano, há o benefício de repetir processos com facilidade, é menos subjetiva principalmente quanto ao limite de feição que a visual, e otimiza a realização do mapeamento e trabalhos de campo.

Com a GEOBIA, a classificação pode ser realizada a partir dos objetos geográficos e conter não só a informação espectral, mas também as espaciais e de textura.

Com a imagem classificada e após validação a campo, pode-se caracterizar o uso e cobertura da terra e suas classes e culminar em um mapa temático de uso e cobertura da terra.

No próximo tópico será abordada a Análise Multivariada e suas técnicas Análise de Componentes Principais e Análise de Agrupamentos que nesse trabalho foram utilizadas para selecionar descritores e melhorar a classificação digital supervisionada.

2.4 Análise Multivariada

Atualmente, com o desenvolvimento tecnológico, o progresso científico e a maior capacidade de obter informações de fenômenos e acontecimentos, existe maior volume de dados disponíveis para análise. Informações que antes eram descartadas ou não eram coletadas, hoje estão dispostas em banco de dados. No entanto, há dificuldades em trabalhar com tanta informação disponível para gerar conhecimento e de fato auxiliar na tomada de decisões.

A Estatística definida por Larson e Farber (2010) como a ciência que coleta, organiza, analisa e interpreta dados para a tomada de decisões possibilita a análise de grande volume de dados. A partir de métodos estatísticos é possível trabalhar com esses diversos dados. Para a análise de variáveis, os métodos estatísticos são classificados em dois grupos: a estatística univariada, que trata as variáveis de maneira isolada, e a estatística multivariada, que trata as variáveis de forma conjunta (VICINI, 2005).

Existem dados que podem ser compreendidos como estatísticas univariadas, mas há dados que necessitam de técnicas mais complexas como a estatística multivariada para serem entendidos.

As técnicas de análise multivariada são populares, pois elas permitem que organizações criem conhecimento, melhorando assim o processo de tomada de decisão. Possibilitam também o trabalho com banco de dados e a análise de inúmeras variáveis e/ou objetos. Na ciência geográfica, a análise multivariada permite examinar grandes conjuntos de informações e diversas variáveis no entendimento de elementos para a explicação de fenômenos geográficos, como é o caso do estudo de tendências e padrões espaciais, a formulação de modelos e a elaboração de previsões (ALENCAR, 2009).

A Análise Multivariada também está presente em nosso dia-a-dia nas diversas escolhas que tomamos, pois, de acordo com Hair et al. (2009), qualquer análise em que se considerem mais que duas variáveis pode ser considerada a princípio como multivariada.

A Análise Multivariada, utiliza-se de métodos estatísticos para descrever e analisar objetos com muitas variáveis simultaneamente e entender o relacionamento entre as diversas variáveis, que faz desse conjunto de técnicas uma metodologia de grande potencial de aplicação. É sugerida quando há a necessidade de redução ou simplificação, distribuição e para o agrupamento de dados, investigação sobre interdependência de variáveis, predição e também teste de hipóteses (JOHNSON; WICHERN, 1988).

Hair et al. (2009), citam que muitas técnicas multivariadas são extensões da análise univariada e bivariada. Algumas técnicas multivariadas fornecem meios de executar em uma única análise aquilo que antes exigia múltiplas análises univariadas para ser realizado.

Pode-se citar como técnicas de análise multivariada: Análise de Componentes Principais, Análise de Agrupamentos, Análise Discriminante, Análise de Correlação Canônica, Análise de Regressão Logística, entre outras.

As técnicas de análise multivariada são de uso interdisciplinar e têm sido utilizadas em várias áreas do conhecimento e em diversos trabalhos.

Alencar (2009) realizou um estudo da aplicação das análises de Componentes Principais e Agrupamentos e apresenta exemplo de aplicação em dados socioeconômicos de 23 municípios pertencentes à mesorregião do Vale do Mucuri, em Minas Gerais. O autor cita que a ACP possibilitou exame nos diversos dados e a AA

separou estes dados em grupos, e que os resultados obtidos primeiramente pela ACP e posteriormente com a AA se mostram complementares.

Silva et al. (2006) identificaram e mapearam zonas homogêneas do meio biofísico da bacia hidrográfica do Alto Taquari localizada nos estados de Mato Grosso e Mato Grosso do Sul. Utilizaram análise de agrupamentos e de correspondência simples e múltipla para seleção de atributos ambientais passíveis de serem medidos, como resultado mapearam e relacionaram estes em base de Sistemas de Informação Geográfica.

Rezende, Fernandes e Silva (2007) definiram potenciais de desenvolvimento dos municípios da região Sul de Minas Gerais, utilizando análise fatorial. A Análise Fatorial permitiu que se estabelecessem a hierarquia entre os municípios, a definição dos que têm maior e menor potencial de crescimento industrial, comercial, social e agropecuário.

Os trabalhos acima citados demonstram que é possível utilizar técnicas de Análise Multivariada nos estudos geográficos, para seleção de atributos e descritores, e redução na dimensionalidade de dados.

A análise de dados geográficos, bem como sua organização e classificação podem ser realizadas com o uso de técnicas da estatística multivariada. Nesse trabalho utilizou-se a Análise de Componentes Principais e Análise de Agrupamentos, estas serão descritas nos próximos subtópicos.

2.4.1 Análise de Componentes Principais

A Análise de Componentes Principais – ACP ou PCA (do inglês *Principal Component Analysis*) – foi introduzida por Pearson em 1901 e desenvolvida de forma independente por Hotelling em 1933.

De acordo com Vicini (2005), a ideia matemática do método é conhecida há muito tempo, apesar do cálculo das matrizes dos autovalores e autovetores não ter sido possível até o advento da evolução dos computadores. O seu desenvolvimento foi conduzido, em parte, pela necessidade de se analisar conjuntos de dados com muitas variáveis correlacionadas.

A Análise de Componentes Principais tem por objetivo concentrar a informação contida em variáveis originais em um conjunto menor de variáveis estatísticas, com perda mínima de informações. Pode ser utilizada para analisar inter-relações entre um

grande número de variáveis e explicá-las em termos de suas dimensões inerentes comuns (HAIR et al., 2009).

Marques (2014), cita que os principais objetivos da ACP são reduzir o número de variáveis e analisar quais as variáveis, ou quais os conjuntos de variáveis explicam a maior parte da variabilidade total, revelando que tipo de relacionamento existe entre eles.

Trata-se de um método fatorial que se traduz na redução do número de variáveis não por uma simples seleção, mas pela construção de novas variáveis sintéticas, obtidas pela combinação linear das variáveis iniciais, por meio de fatores (BOROUCHE, 1982).

Para investigar as relações entre um conjunto de variáveis correlacionadas, pode ser útil transformar o conjunto de variáveis originais em um novo conjunto de variáveis não correlacionadas, chamadas de componentes principais, tendo propriedades especiais em termos de variâncias. Os componentes principais são combinações lineares das variáveis originais e são derivadas em ordem decrescente de importância, ou seja, a 1ª componente principal é a combinação linear normalizada com variância máxima (MARQUES; MENDES MARQUES, 2005).

A reprodução da variabilidade total requer as “p” variáveis, porém, frequentemente a maior parte da variância pode ser explicada por um número pequeno $k < p$ de componentes principais. Neste caso, existe praticamente a mesma quantidade de informações nas “k”, componente principais, que nas “p”, variáveis originais. As “k”, componente principais, podem então substituir as “p”, variáveis originais (MARQUES, 2014).

De acordo com Richards (1986), a Análise de Componentes Principais pode ser resumida em três passos: (a) derivação da matriz de correlação ou de variância covariância, (b) cálculo dos autovetores e autovalores, e (c) transformação linear do conjunto de dados.

Esta transformação tem como base a rotação do espaço de atributos na direção dos autovetores com o intuito de tornar ortogonal o conjunto de dados, promovendo a união das informações de maior correlação nas primeiras componentes (MATHER, 1999). A ACP reduz a dimensionalidade dos dados, concentrando as informações significativas nas primeiras componentes geradas pela transformação linear (SCHOWENGERDT, 1997).

Para a determinação das componentes principais, é necessário calcular a matriz de variância-covariância (Σ), ou a matriz de correlação (R), encontrar os autovalores e

os autovetores e, por fim, escrever as combinações lineares, que serão as novas variáveis, denominadas de componentes principais. Cada componente principal é uma combinação linear de todas as variáveis originais, independentes entre si e estimadas com o propósito de reter, em ordem de estimação e em termos da variação total, a informação contida nos dados iniciais (REGAZZI, 2001).

A utilização das componentes principais para definir um espaço de fatores que englobe os dados, não modifica os dados em si, mas simplesmente encontra um sistema de coordenadas mais conveniente, capaz de remover ruídos dos dados sem distorcê-los e diminuir sua dimensionalidade sem comprometer seu conteúdo de informações. (NETO e MOITA, 1998).

Ainda de acordo com Marques (2014), a ACP é muito útil como método auxiliar em Regressão, Análise Fatorial e Análise de Agrupamentos. A técnica é utilizada em diversas áreas da ciência como a Biologia, Psicologia, Agronomia, Ciência da Computação, Sensoriamento Remoto e Geografia.

Petrisor et al. (2011) propõem uma abordagem que combina Análise de Componentes Principais e a modelagem de Sistemas de Informação Geográfica para produzir e mapear uma hierarquia de subunidades espaciais na Romênia. Nesta abordagem, os autores determinam as variáveis mais influentes e seus pesos com base na percentagem de variabilidade explicada. Estas variáveis são incorporadas em um modelo de SIG. Os resultados indicam a utilidade da abordagem não só como instrumento de pesquisa, mas também para fins de planejamento estratégico. A técnica permitiu identificar as variáveis com maior relevância para o nível de desenvolvimento nos setores de educação, economia, sociocultural e demografia.

Jesus e Epiphânio (2010) utilizaram a Análise de Componentes Principais para geração de componentes de brilho e verdor, aplicados a dados dos sensores MSS (Multispectral Scanner), Landsat TM (Thematic Mapper), Landsat e CCD (Câmera Imageadora de Alta Resolução). Mediante as componentes principais, geraram Análise por Vetores de Mudanças (AVM), fornecendo informações sobre a intensidade e o tipo de mudança ocorrida na evolução temporal da expansão agrícola do município de Chapadão do Céu, Goiás.

Abson et al. (2012) apresentam análise de vulnerabilidade sócio ecológica para a região da Comunidade de Desenvolvimento da África Austral (*Southern Africa Development Community* – SADC) usando tanto PCA e as técnicas de normalização para gerar índices de vulnerabilidade sócio ecológicos. Os mapas de vulnerabilidade

resultantes indicam a variabilidade espacial regional dos quatro componentes estatisticamente independentes. De acordo com os autores, a técnica apresenta instrumento de política potencialmente útil para a identificação de áreas de maior preocupação em termos de nível relativo e as causas e impactos da vulnerabilidade ecológica e social às mudanças ambientais em grandes escalas espaciais.

Outra técnica multivariada importante aqui usada é a Análise de Agrupamentos a qual será descrita na próxima subseção.

2.4.2 Análise de Agrupamentos

A Análise de Agrupamento (AA) ou *Cluster Analysis* é uma técnica estatística que visa agrupar variáveis com características comuns, sem perder informações de todo o conjunto em estudo. É amplamente utilizada nas diversas áreas do conhecimento, por se tratar de medida contínua que possibilita a interpretação individual de cada grupo e a relação que este grupo possui com os demais. Trata-se de uma técnica usada para classificar objetos ou casos, em grupos relativamente homogêneos que podem ser chamados também de conglomerados. (JOHNSON; WICHERN, 2002).

Os autores Hair et al. (2009) mencionam que essa técnica realiza a tarefa inata a todos os indivíduos, ou seja, o reconhecimento de padrões e agrupamento. O objetivo principal da análise de agrupamentos é definir a estrutura de dados colocando as observações mais parecidas em grupos.

De acordo com Gore Jr. (2000), a análise de agrupamentos foi introduzida por Tryon em 1939 e Castell em 1944 com procedimentos matemáticos para a organização de objetos com base na similaridade observada.

A finalidade principal desta técnica multivariada é agregar objetos com base nas características que eles possuem. É intitulada também de análise Q, construção de tipologia, análise de classificação e taxonomia numérica. É utilizada nas mais diversas áreas como psicologia, biologia, sociologia, economia, engenharia e administração. Os métodos têm uma dimensão em comum: classificação de acordo com relações entre os objetos sendo agrupados. Essa dimensão representa a essência de todas as abordagens de agrupamento – a classificação de dados, como sugerido pelos agregados naturais dos dados em si (HAIR et al., 2009).

De acordo com os autores supracitados a análise de agrupamentos é usada para descrição taxonômica (identificar grupos naturais dentro dos dados), simplificação de

dados (a habilidade de analisar grupos de observações semelhantes em vez de todas as observações individuais), e identificação de relações, já que a estrutura simplificada da análise de agrupamentos retrata relações não reveladas de outra forma.

Segundo Johnson e Wichern (1998), a técnica tem como objetivo descobrir os agrupamentos naturais das variáveis, dividindo os elementos da amostra em grupos, de forma que os elementos pertencentes a um mesmo grupo sejam similares entre si com respeito às variáveis, e os elementos de grupos diferentes sejam heterogêneos, em relação a estas mesmas características. Segundo os autores, essa técnica não requer suposição sobre o número de grupos ou sua estrutura, o agrupamento é feito a partir de similaridades ou distâncias.

A Análise de Agrupamentos objetiva gerar uma estrutura classificatória ou de reconhecimento da existência de grupos, que dividem o conjunto de observações em um número de grupos homogêneos, segundo algum critério de homogeneidade (REGAZZI, 2001).

A análise de agrupamentos é muito usada em conjunto com outras técnicas multivariadas como a análise fatorial e de componentes principais.

Citam-se alguns trabalhos que utilizam a Análise de Agrupamentos:

Adami et al. (2008), com o objetivo de avaliar dados temporais de imagens MODIS (*Moderate Resolution Imaging Spectroradiometer*) para estudo da dinâmica espaço-temporal de duas regiões do Bioma Pantanal, utilizou das análises de componentes principais, autocorrelação e agrupamentos. A partir da Análise de Agrupamentos foi possível a identificação de dois grupos direcionando o estudo em toda a área, e com a ACP definiu-se quatro compartimentos de diferentes vegetações.

Fritzons e Montavani (2010) estabeleceram algumas relações entre índices morfométricos e a litologia de bacias hidrográficas. Os índices morfométricos foram submetidos à análise de agrupamentos para agrupar as bacias hidrográficas com morfometria semelhantes. Concluíram a partir do resultado da análise de agrupamentos que os grupos formados fazem boa correlação com o substrato rochoso.

Boratto e Avellar (2013) utilizaram o índice de vegetação por diferença normalizada e as técnicas de estatística multivariadas de análise de componentes principais e análise de agrupamentos com o objetivo de espacializar, quantificar e classificar as áreas verdes do município de Sete Lagoas. Concluiu-se que a ACP apresenta correlação fraca em relação às variáveis sócio econômicas e ambientais, o que

mostrou que o processo de degradação das áreas verdes está relacionada à falta de planejamento.

Melo Júnior et al. (2006) utilizaram da análise de agrupamentos e também de componentes principais para determinar regiões homogêneas quanto à distribuição de frequências de chuvas no leste do estado de Minas Gerais. Como resultado foram agrupados cinco e três regiões para os critérios de classificação em meso e macroescala climática da área de estudo.

No próximo tópico serão abordados os materiais e métodos utilizados nesse trabalho.

3 Materiais e métodos

3.1 Localização e caracterização da área de estudo

Para esse trabalho, foram definidos dois segmentos da bacia do rio Pitangui, a qual se localiza entre os municípios de Castro, Carambeí, e Ponta Grossa, Paraná, conforme ilustra a Figura 4.

O rio Pitangui é entendido como um rio de contrastes, por percorrer compartimentos geomorfológicos e domínios geológicos imprimindo diferenças na rede hidrográfica e em suas paisagens (MELO, GUIMARÃES E SANTANA, 2010). A bacia do rio Pitangui possui uma área de 1.001,98 Km², sua rede de drenagem segue em direção ao interior do estado (oeste). O rio nasce no Primeiro Planalto Paranaense (PPP), atravessando a Escarpa Devoniana (ED) mediante vale encaixado e em seguida sendo barrado artificialmente, formando o reservatório de Alagados, utilizado para produção de energia hidrelétrica e abastecimento de água para o município de Ponta Grossa. Extravasando a represa, o rio Pitangui segue seu curso no Segundo Planalto Paranaense (SPP) em direção à foz, no rio Tibagi no município de Carambeí.

Os dois recortes localizam-se no Segundo Planalto Paranaense (Figuras 5 e 6), na área rural dos municípios de Ponta Grossa e Carambeí, Paraná.

Figura 4- Localização das áreas de estudo

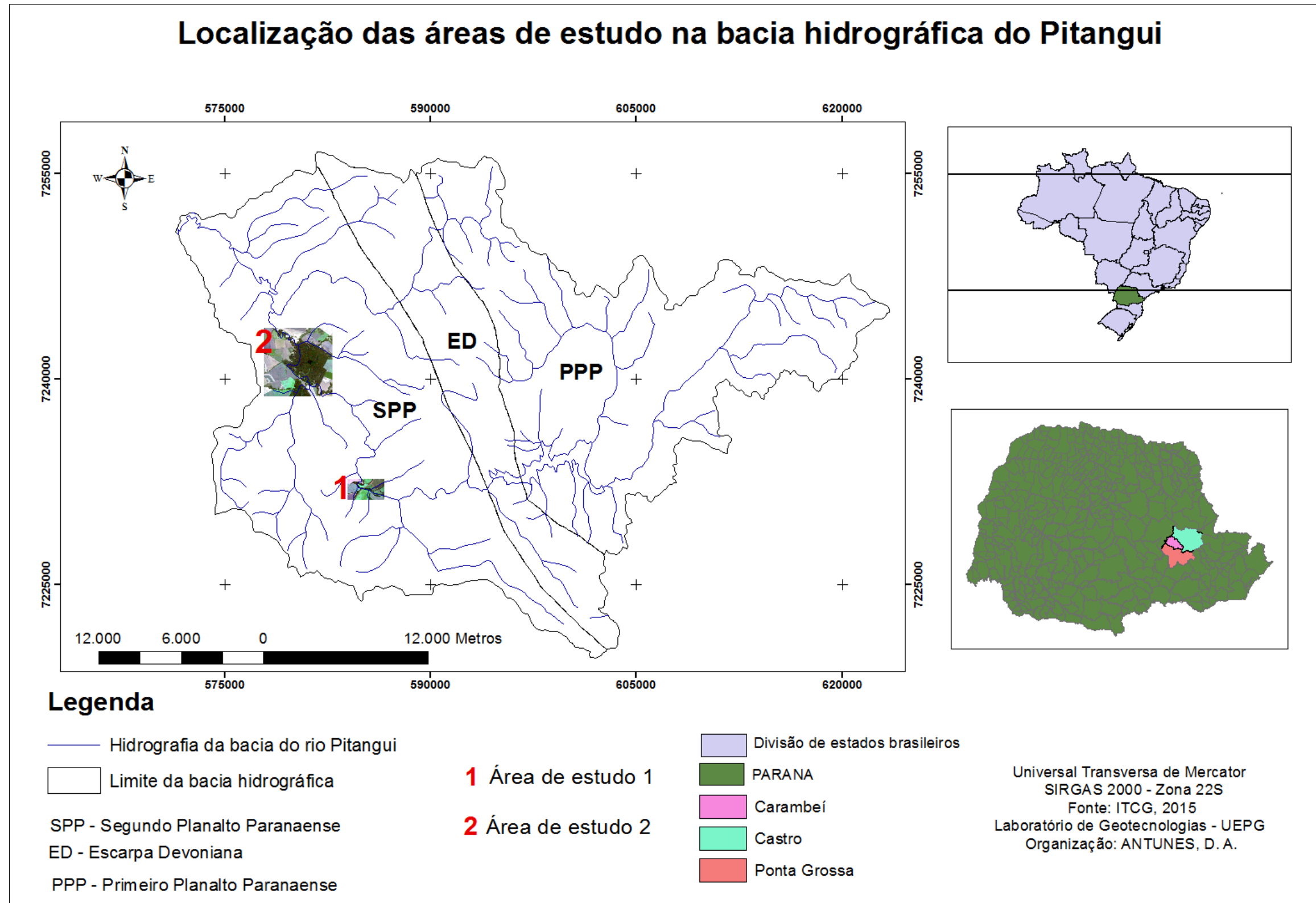
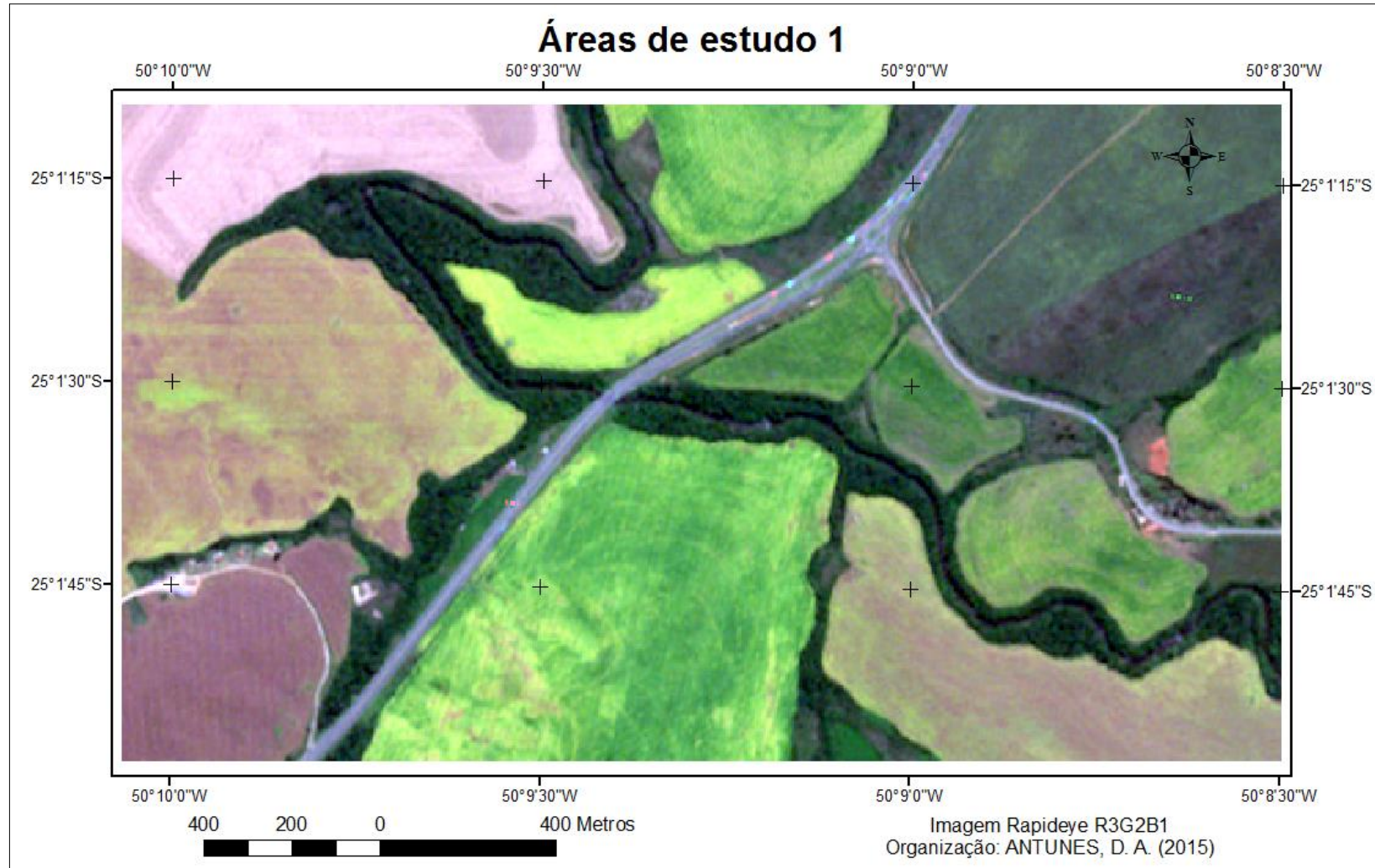
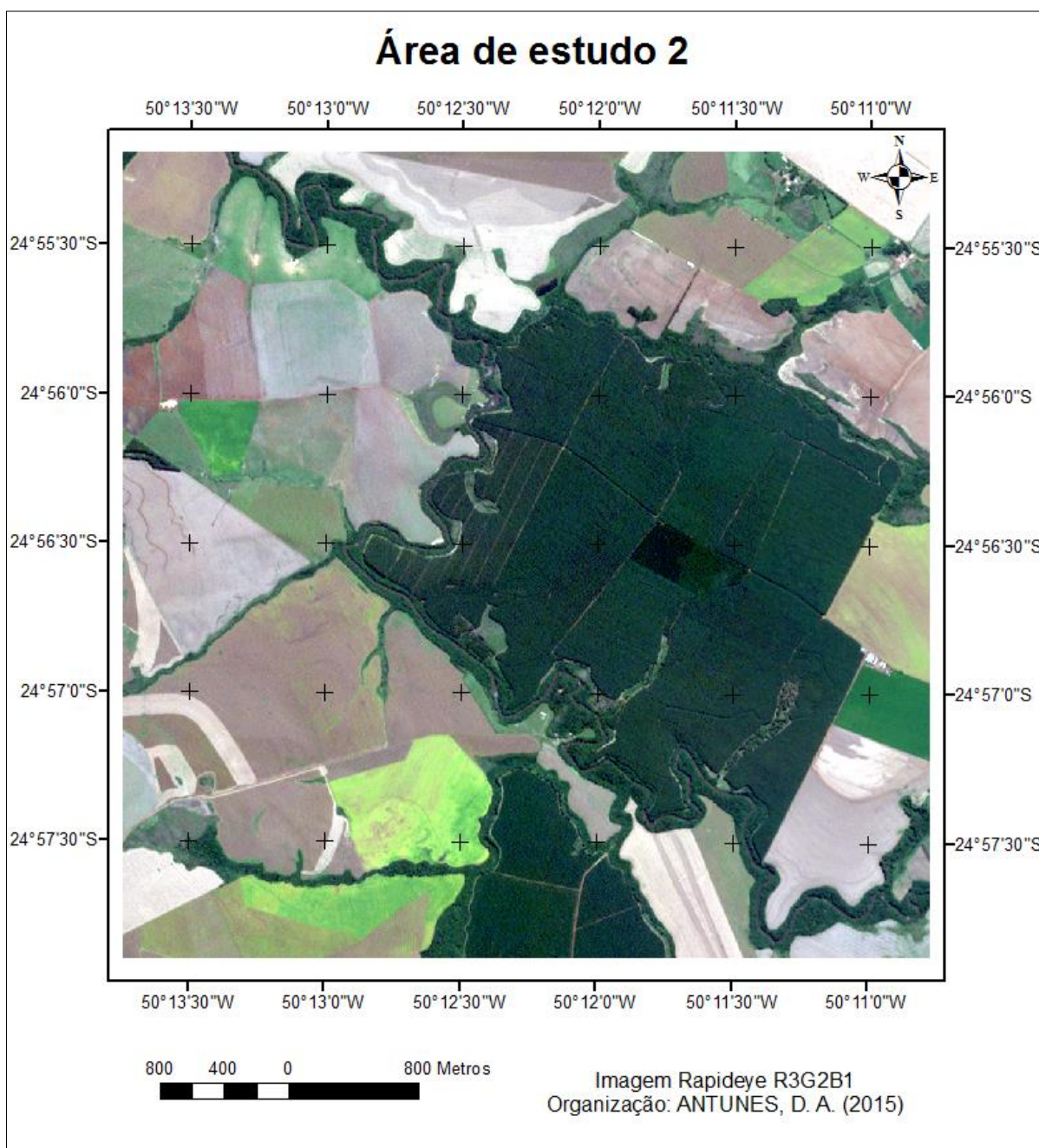


Figura 5- Área de estudo 1



Fonte: a autora.

Figura 6- Área de estudo 2



Fonte: a autora.

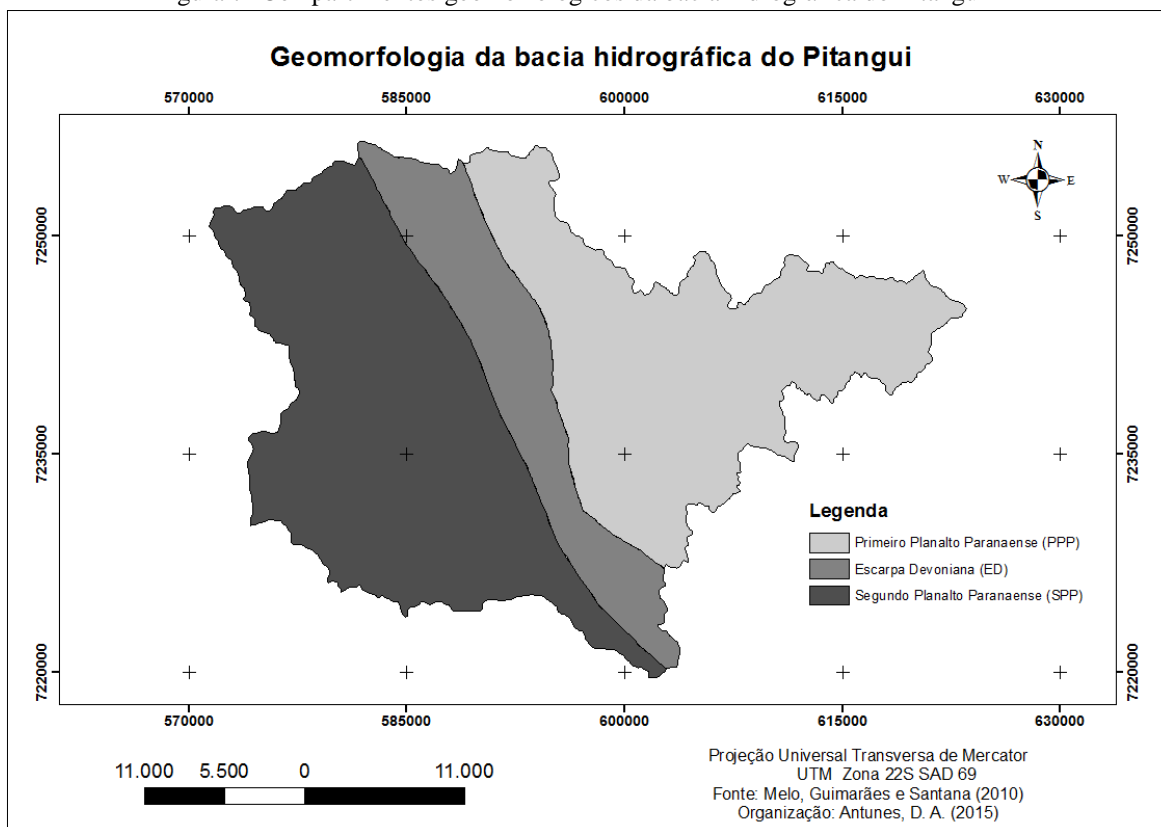
3.1.1 Características da bacia hidrográfica do Pitanguí

Descrevem-se aqui as características físicas da bacia do rio Pitanguí como a Geomorfologia, Geologia, clima, solos e hidrografia, as quais são importantes, pois influenciam os usos e ocupações da terra e a distribuição dos recursos naturais.

3.1.1.1 Geomorfologia

A bacia hidrográfica do Pitangui está situada sobre os compartimentos geomorfológicos: Primeiro Planalto Paranaense (PPP), Escarpa Devoniana (ED) e Segundo Planalto Paranaense (SPP), conforme ilustra a Figura 7.

Figura 7- Compartimentos geomorfológicos da bacia hidrográfica do Pitangui



Fonte: Adaptado de Melo, Guimarães e Santana (2010)

A unidade morfológica, denominada Primeiro Planalto Paranaense é relativamente uniforme, esculpida em rochas cristalinas, tais como xistos metamórficos e gnaisses, cortados por diques de pegmatitos e intrusões graníticas, com altitudes médias entre 850-950 metros, formando paisagens suavemente onduladas com planície e várzeas intercaladas constituídas por sedimentos colúvio-aluvionares recentes e paludais ao longo dos principais cursos de água. (MINEROPAR, 2006). A bacia do rio Pitangui possui aproximadamente um terço de seu território situado no PPP. (MELO, GUIMARÃES E SANTANA, 2010).

A Escarpa Devoniana faz a transição entre o PPP o SPP (MELO ET AL, 2007). Delimita a leste os Campos Gerais do Paraná e estende-se como uma faixa em forma de arco, com cerca de 260 quilômetros de extensão e desníveis altimétricos usualmente

entre 100 e 200 metros. A Escarpa Devoniana representa um degrau topográfico, com paredes abruptas e verticalizadas, que separa o Primeiro e o Segundo Planalto Paranaense. (MENEGUZZO; MELO 2012).

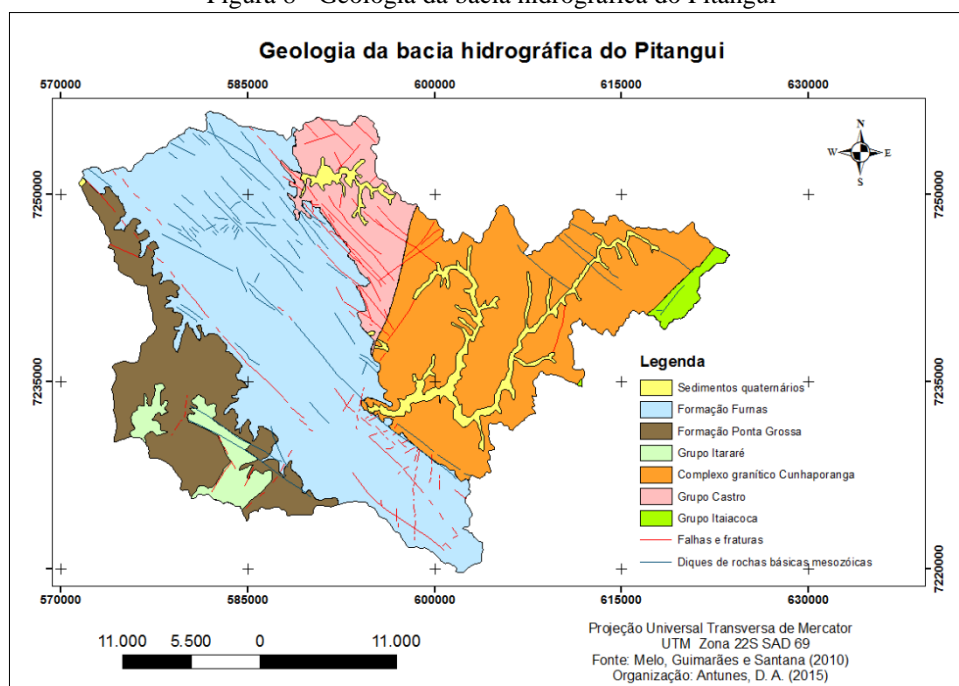
No SPP o relevo predominante é colinoso, com elevações amplas com topos de 900 e 1.000m de altitude. É também onde se encontra a área de foz do rio Pitangui, com 755m (GEAH, MELO, MORO, 2010).

3.1.1.2 Geologia

Os autores Melo, Guimarães e Santana (2010) argumentam que a geologia do rio Pitangui está impressa sobre dois compartimentos geológicos: a Bacia Sedimentar do Paraná e o seu embasamento.

De acordo com os mesmos, no curso superior do rio há ocorrência de rochas principalmente do Complexo Granítico Cunhaporanga e em menor proporção dos grupos Castro e Itaiacoca, que são unidades geológicas que constituem os representantes locais do domínio conhecido por embasamento. Já na porção média e no trecho de jusante encontram-se rochas da Bacia do Paraná, onde o leito do rio e praticamente toda a porção de sua margem direita têm como substrato a Formação Furnas, enquanto na margem esquerda aparecem as Formações Furnas, Ponta Grossa e o Grupo Itararé. Ainda encontram-se na bacia sedimentos quaternários, falhas e fraturas e diques de rochas básicas mesozóicas. A Figura 8 ilustra a geologia da bacia hidrográfica do Pitangui.

Figura 8 - Geologia da bacia hidrográfica do Pitangui



Fonte: Melo, Guimarães e Santana (2010)

3.1.1.3 Vegetação

Quanto à vegetação pode-se encontrar espécies florestais e campestres, ainda que extremamente impactados pela ação antrópica (CARMO et al, 2010). No Primeiro Planalto Paranaense predominam várzeas, campos higrófilos, campos secos e formações florestais.

Acompanhando a Escarpa Devoniana e nas cabeceiras do rio Pitangui, são encontradas expressivas áreas com remanescentes de Floresta Ombrófila Mista em estágios secundários de sucessão ecológica, enquanto mais próximas à escarpa são encontradas florestas em estado mais avançados de sucessão (GEAH; MELO; MORO, 2010).

No Segundo Planalto Paranaense existe a presença de formações campestres associadas às florestas de galeria e capões de mata, bem como de campos nativos onde são desenvolvidos cultivos intensivos (VELOSO; RANGEL FILHO; LIMA, 1991).

3.1.1.4 Clima

O clima da região do rio Pitangui é o Cfb da classificação de Koeppen (Maack, 2012), característico de zona temperada sempre úmida, possuindo temperatura média do mês mais frio de 18°C, do mês mais quente inferior a 22°C e sem a existência de uma estação seca definida. A precipitação média anual situa-se entre 1.400 e 1.800 mm, sendo que o período mais seco é de junho a agosto com médias de 250 a 350 mm e o mais chuvoso é de dezembro a fevereiro com médias de 500 a 600 mm (CRUZ, 2007).

3.1.1.5 Solos

Os solos predominantes na bacia hidrográfica do rio Pitangui e sua distribuição refletem-se em parte aos traços principais do relevo, das unidades rochosas presentes na área. São encontrados Cambissolos, Neossolos Litólicos, Organossolos, Gleissolos, e Latossolos (GEAH; MELO; MORO, 2010).

3.1.1.6 Hidrografia

A hidrografia da bacia do rio Pitangui é marcada fortemente por características assimétricas. Esta assimetria pode ser justificada pela complexidade morfológica da área: a margem esquerda situa-se em sua maior parte no SPP, enquanto boa parte da margem direita abrange o PPP, parte do SPP e o relevo de transição inclui a ED. O marcante controle por estruturas geológicas nas direções noroeste-sudeste e nordeste-sudeste; as diferenças nos tipos rochosos e relevo associado que aparecem no PPP e a diversidade da dinâmica fluvial de bacias vizinhas, são fatores que também contribuem para a assimetria da mesma (GEALH, MELO, MORO, 2010).

No Primeiro Planalto (os rios da parte montante dos rios Pitangui e Jotuba e outros da margem direita à jusante do rio São João) a drenagem apresenta-se com direção nordeste-sudoeste, controlada pela estrutura do embasamento da Bacia do Paraná. A margem esquerda do rio Pitangui no SPP é caracterizada por apresentar afluentes com padrão de drenagem dendrítico (forma ramificada). No relevo de transição (ED) os rios mostram-se controlados pelas estruturas geológicas relacionadas com o Arco de Ponta Grossa e apresentam-se alinhados na direção noroeste-sudeste (GEALH, MELO, MORO, 2010).

O rio Pitangui, assim como outros rios presentes na região, se caracteriza pela formação antecedente, isto é, anterior ao relevo atual. Por isso, muitos dos rios dos Campos Gerais apresentam notável controle estrutural imposto pelas estruturas rúpteis (falhas, fraturas, diques) de direção predominante noroeste-sudeste, paralelas ao eixo do Arco de Ponta Grossa, bem como, em outras direções associadas.

3.2 Materiais

3.2.1 Imagem RapidEye

As imagens RapidEye fazem parte do sistema de mesmo nome composto por cinco satélites de observação da Terra lançados em 29 de Agosto de 2008 pela empresa alemã RapidEye. As imagens são entregues ortorretificadas e com correção atmosférica. O Quadro 1 traz as especificações técnicas dos satélites do sistema RapidEye.

Quadro 1 – Especificações técnicas dos satélites do sistema RapidEye

Número de Satélites:	5
Altitude da Órbita:	630 km, órbita síncrona com o Sol
Hora de Passagem no Equador:	11:00 hs (aproximadamente)
Velocidade:	27.000 km/h
Largura da Imagem:	77 km
Tempo de Revisita:	Diariamente (off-nadir); 5,5 Dias (nadir)
Capacidade de Coleta:	4,5 milhões de Km ² /dia
Tipo do Sensor:	Multiespectral (pushbroomimager)
Bandas Espectrais:	5 (Red, Green, Blue, Red-Edge, Near IR)
Resolução Espacial (nadir):	6,5 m
Tamanho do Pixel (ortorretificado):	5 m
Armazenamento de Dados a Bordo:	1.500 km de dados de imagens por órbita
Resolução Radiométrica:	12 bits
Velocidade de Download (banda X):	80 Mbps

Fonte: Felix, Kazmierczak e Espindola (2009).

No Quadro 2 são apresentadas as bandas espectrais das imagens RapidEye.

Quadro 2 - Bandas espectrais das imagens RapidEye

Banda 1	440 – 510 μ	Azul
Banda 2	520 – 590 μ	Verde
Banda 3	630 – 685 μ	Vermelho
Banda 4	690 – 730 μ	Red-Edge
Banda 5	760 – 850 μ	Infravermelho próximo

Fonte: Felix, Kazmierczak e Espindola (2009).

A imagem RapidEye foi disponibilizada pelo Programa de Pós Graduação em Geografia da Universidade Estadual de Ponta Grossa, com data de agosto de 2012.

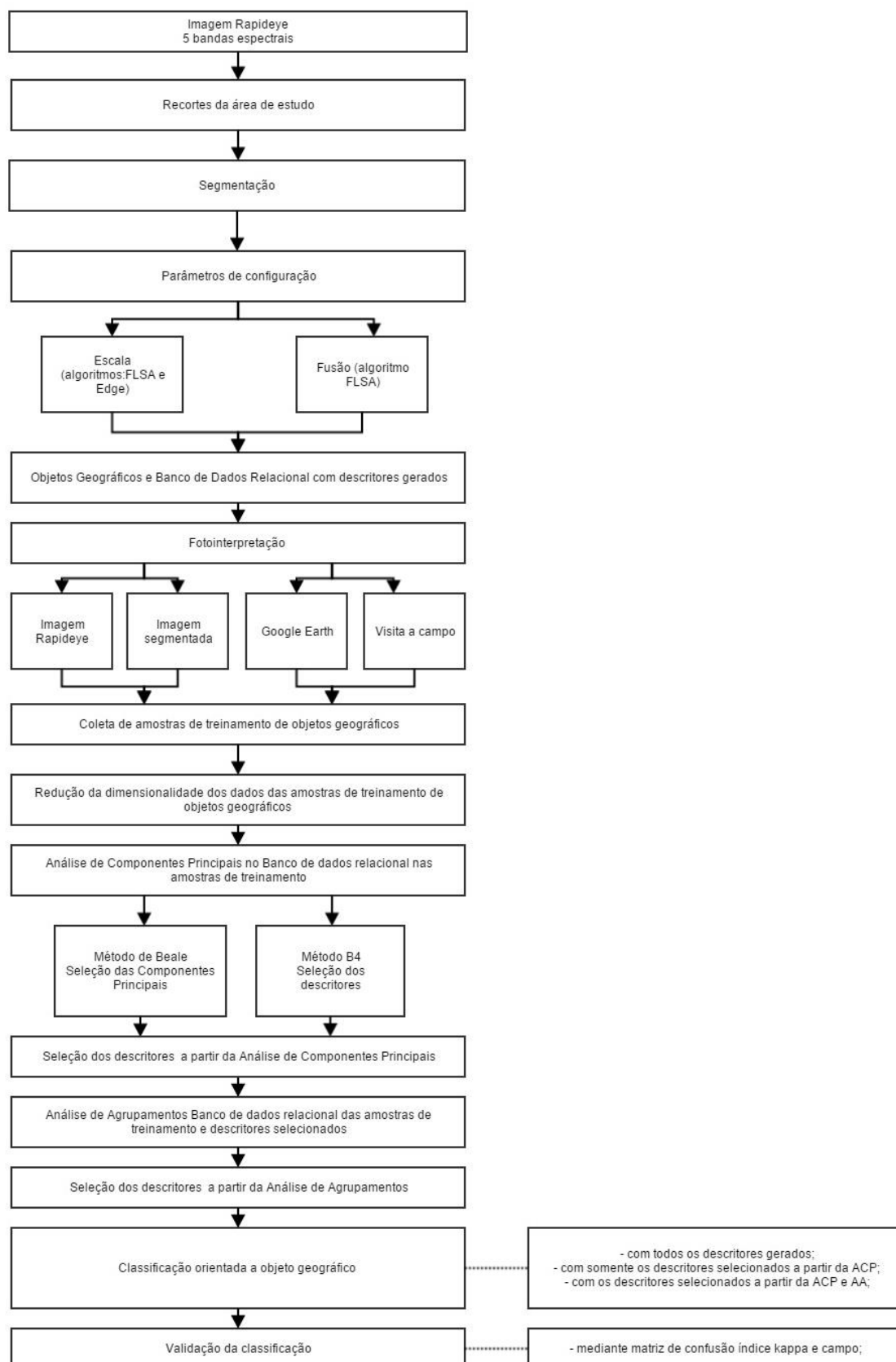
3.2.2 Softwares

- ENVI 5.1: *Software* de PDI desenvolvido pela empresa EXELIS, nesse trabalho possibilita a realização da análise e classificação orientada a objetos;
- R: *Software* estatístico livre utilizando-se para as técnicas de Análise de Componentes Principais (ACP);
- Past: *Software* estatístico livre utilizando-se para gerar os dendrogramas da Análise de Agrupamentos;
- ArcGis 10.2: *Software* de SIG desenvolvido pela empresa ESRI.

3.3 Procedimentos metodológicos

No diagrama de blocos a seguir são explicitados os procedimentos metodológicos realizados nesse trabalho.

Figura 9- Procedimentos metodológicos



Fonte: A autora.

3.3.1 Análise Orientada a Objeto Geográfico (GEOBIA – *Geographic Object Based Image Analysis*)

Como abordado no referencial teórico, a GEOBIA constitui-se em uma metodologia de processamento das imagens digitais de satélite. Nela são considerados atributos espectrais, espaciais e de textura, o que representa mudança da abordagem convencional *pixel a pixel* para a orientação a objeto geográfico, estes que são agrupamentos de *pixels* semelhantes.

Sabe-se que a GEOBIA segmenta a imagem e também gera topologia mediante o banco de dados relacional, o qual contém os descritores espectrais, espaciais e de textura de cada objeto geográfico gerado. O dado de entrada para a GEOBIA foi a imagem RapidEye com suas 5 bandas espectrais.

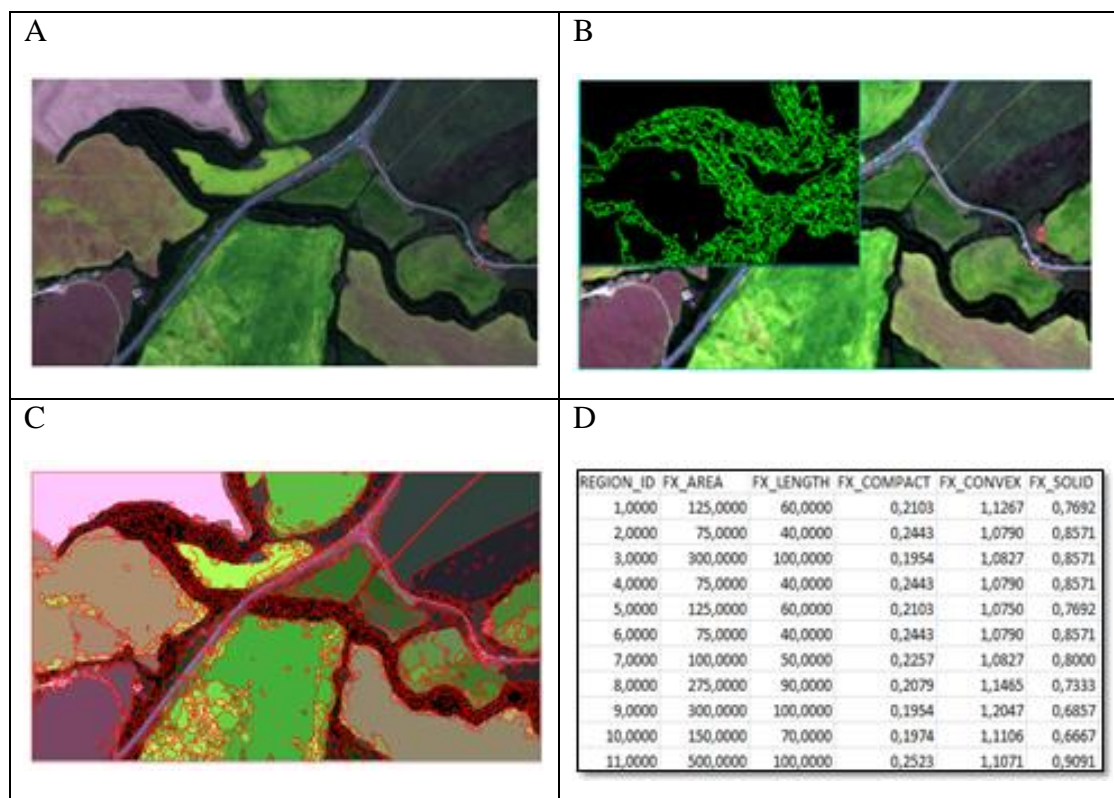
A segmentação foi realizada mediante os algoritmos *Edge* e FLSA, os quais possibilitam a configuração do parâmetro tamanho dos segmentos em função do interesse de estudo e resolução espacial. Dessa forma os *pixels* são agrupados criando os objetos geográficos.

A segmentação é definida no nível de escala que opera com o algoritmo *Edge*, quanto maior o valor para esse nível, maior são os objetos gerados, mas em menor número. O nível de fusão que opera com o algoritmo FLSA realiza um refinamento no procedimento, ao passo que agrega objetos menores e semelhantes dentro de segmentos maiores.

Para a seleção dos parâmetros da segmentação, foram realizados e analisados visualmente experimentos com diferentes valores para o nível de escala e fusão. Até chegar ao nível ideal com separação dos objetos geográficos, gerou-se segmentações com objetos geográficos muito grandes e em pequeno número, também “super-segmentação” com objetos muito pequenos e em grande número.

Ao fim da segmentação, são gerados polígonos de cada objeto e banco de dados relacional contendo descritores espectrais, espaciais e de textura para cada objeto, como ilustra a Figura 10.

Figura 10 - Etapas de orientação a objetos A- Imagem original; B - Escala e fusão 50 e 30; C - Imagem segmentada sobreposta por polígonos gerados; D - Parte do banco de dados relacional.



Fonte: A autora.

O banco de dados relacional contém atributos ou descritores extraídos da imagem, podendo ser do tipo espectral, espacial ou de textura.

Os atributos ou descritores espectrais permitem trabalhar com a heterogeneidade espectral dentro da mesma feição do terreno ou objeto. Já os atributos ou descritores de forma e de textura permitem melhor classificação automática de imagens, pois por meio deles é possível distinguir segmentos com características espectrais semelhantes.

A seguir, segue a explanação de cada um dos descritores gerados de acordo com EXELISVIS, 2014:

3.3.1.1 Descritores Espectrais

Os descritores espectrais são obtidos a partir de cada banda espectral da imagem e são quantificados pelo comportamento espectral dos objetos ou alvos que compõem a cena (Swain e Davis, 1978). Os valores são calculados para cada segmento (agrupamento de *pixels*), onde todos os *pixels* desse conjunto contribuem para o cálculo.

São descritores espectrais:

- a) **Média dos números digitais, por banda (AVGBAND):** valor correspondente ao valor médio dos números digitais dos *pixels* que compõem o objeto, em cada uma das bandas disponíveis.
- b) **Desvio padrão dos números digitais, por banda (STDBAND):** valor correspondente ao desvio padrão dos números digitais dos *pixels* do objeto, em cada uma das bandas disponíveis.
- c) **Valor Máximo de *pixel*, por banda (MAXBAND):** valor correspondente ao maior valor digital observado nos *pixels* dos objetos, em cada uma das bandas disponíveis.
- d) **Valor Mínimo de *pixel*, por banda (MINBAND):** valor correspondente ao menor número digital observado nos *pixels* do objeto, em cada uma das bandas disponíveis.

3.3.1.2 Descritores Espaciais

Os objetos geográficos oriundos da segmentação possuem características geométricas que podem ser traduzidas em forma de comprimentos, áreas ou relações matemáticas entre medidas do segmento. Buscam representar a forma do objeto (polígono) matematicamente, de modo que o computador possa interpretar.

São descritores espaciais:

- a. **Área (AREA):** valor da área do segmento ou do polígono, em unidades de mapa. Caso o segmento possua um furo, a área do furo é descontada.
- b. **Comprimento (LENGHT):** valor do comprimento do perímetro (limites) do segmento, somados aos perímetros de possíveis furos que estejam dentro deste segmento.
- c. **Compacidade (COMPACT):** medida de forma que indica o quão compacto é o segmento, sendo o círculo a forma mais compacta. A compacidade pode ser descrita pela seguinte equação:

$$\text{Compacidade} = \frac{\sqrt{4 \cdot \frac{\text{Área}}{\pi}}}{\text{Perímetro}}$$

(1)

- d. **Convexidade (CONVEXITY):** medida do quanto um polígono pode ser dito convexo ou côncavo. Para polígono convexo sem furos, o valor de convexidade é

1, já para polígonos côncavos é menor que 1. Ou seja, em um polígono convexo, unindo-se quaisquer dois de seus pontos, o segmento de reta obtido está sempre contido no polígono. Já em um Polígono Côncavo, existem pelo menos dois pontos que unidos formam um segmento de reta que não se encontra contido no polígono. Pode ser dada pela seguinte equação:

$$\text{Convexidade} = \frac{\text{Comprimento fecho convexo}}{\text{Comprimento}} \quad (2)$$

- e. **Solidez (SOLIDIT):** compara a área do polígono com a área do fecho convexo no entorno do polígono. O valor de solidez para um polígono convexo, sem buracos, é 1, e para um polígono côncavo é menor que 1. Segue a equação:

$$\text{Solidez} = \frac{\text{Área do polígono}}{\text{Área do fecho convexo}} \quad (3)$$

- f. **Circuncidade (ROUNDNESS):** compara a área do polígono com quadrado do diâmetro máximo do polígono. O diâmetro máximo é o comprimento do eixo maior de uma retângulo envolvente (*bounding box*) em torno do polígono. O valor de circuncidade para um círculo é 1. É dado pela equação:

$$\text{Redondeza} = \frac{4 \cdot \text{área}}{\pi \cdot (\text{eixo maior})^2} \quad (4)$$

- g. **Fator de Forma (FORMFACTOR):** compara a área de um polígono com o quadrado do seu perímetro. Para um círculo o fator de forma é 1. Segue a equação:

$$\text{Fator de forma} = \frac{4 \cdot \pi \cdot \text{área}}{(\text{perímetro})^2} \quad (5)$$

- h. Alongamento (ELONGATION):** indica a razão ou a proporção entre o eixo maior e menor do polígono e tais eixos são obtidos a partir de um retângulo envolvente em torno do polígono. Para um quadrado o valor é 1 e para um retângulo é maior que 1. Pode ser descrita pela equação:

$$\text{Alongamento} = \frac{\text{Eixo maior}}{\text{Eixo menor}} \quad (6)$$

- i. Retangularidade (RECT_FIT):** indica o quanto um polígono se aproxima da forma retangular, pela comparação da área do polígono com a área de um retângulo envolvente ao polígono. Para um retângulo o valor é 1, para uma forma não retangular o valor é menor 1. Segue a equação:

$$\text{Retangularidade} = \frac{\text{Área}}{\text{Eixo maior} \cdot \text{Eixo menor}} \quad (7)$$

- j. Direção Principal (MAINDIR):** indica a direção, em graus, do eixo maior do polígono em relação ao eixo x. Os valores variam de 0° a 180°, sendo que 90° significa direção Norte-Sul e 0° a 180° significa Leste-Oeste.
- k. Eixo maior (MAJAXISLEN):** o comprimento do eixo maior de um retângulo envolvente em torno do polígono. O valor é baseado no tamanho do *pixel* no terreno, no caso da imagem ser georreferenciada. Caso não seja georreferenciada, o valor é dado em *pixels*.
- l. Eixo menor (MINAXISLEN):** o comprimento do eixo menor de um retângulo envolvente em torno do polígono. O valor é baseado no tamanho do *pixel* no terreno, no caso da imagem ser georreferenciada. Caso não seja georreferenciada, o valor é dado em *pixels*.
- m. Número de furos (NUMHOLES):** representa o número de furos ou buracos inseridos no polígono.
- n. Relação área / furos (HOLESOLRAT):** é a razão entre a área do polígono e a área do contorno. A área do contorno é a área da parte não vazia do polígono,

quando este possui buracos. O valor será igual a 1 quando o polígono não possuir furos. Segue a equação:

$$\text{Relação área/furos} = \frac{\text{Área total do polígono}}{\text{Área do contorno}} \quad (8)$$

3.3.1.3 Descritores de Textura

A textura é utilizada como elemento auxiliar no processo de classificação, especialmente na distinção entre duas classes com comportamento espectral similar, podendo ser descrita, de forma qualitativa, como “uniforme”, “grosseira”, “regular”, “fina” ou “média” (Claro, 1995).

De acordo com EXELISVIS (2014), os atributos de textura são computados em processo de duas etapas, onde a primeira passagem se aplica um *kernel* quadrado de tamanho pré-definido para a banda de imagem de entrada. Os atributos são calculados para todos os *pixels* na janela do *kernel* e o resultado é referenciado ao *pixel* central do *kernel*. Em seguida, os resultados de atributos são a média em cada *pixel* no agrupamento de *pixel* para criar o valor do atributo para as bandas.

São descritores de textura:

- a. Intervalo da textura, por banda (TX_RAN):** corresponde ao intervalo de valores dos *pixels* compreendidos dentro da janela ou *kernel* de variância.
- b. Textura média, por banda (TX_AVG):** corresponde à média dos valores dos *pixels* compreendidos dentro da janela.
- c. Variância da textura, por banda (TX_VAR):** corresponde à variância dos valores dos *pixels* compreendidos dentro da janela, para a banda.
- d. Entropia da textura, por banda (TX_ENT):** corresponde ao valor médio da entropia dos valores dos *pixels* compreendidos dentro da janela. De acordo com Marco et al. (2010), a entropia é uma medida clássica adotada em teoria da informação que determina o grau de desorganização. Pode ser descrita pela seguinte fórmula:

$$H = - \sum_{i=0}^{G-1} p(i) \log_2 [p(i)] \quad (9)$$

Os descritores da GEOBIA foram selecionados e utilizados na classificação supervisionada orientada a objeto geográfico que será abordada a seguir.

3.3.2 Classificação supervisionada orientada a objeto geográfico

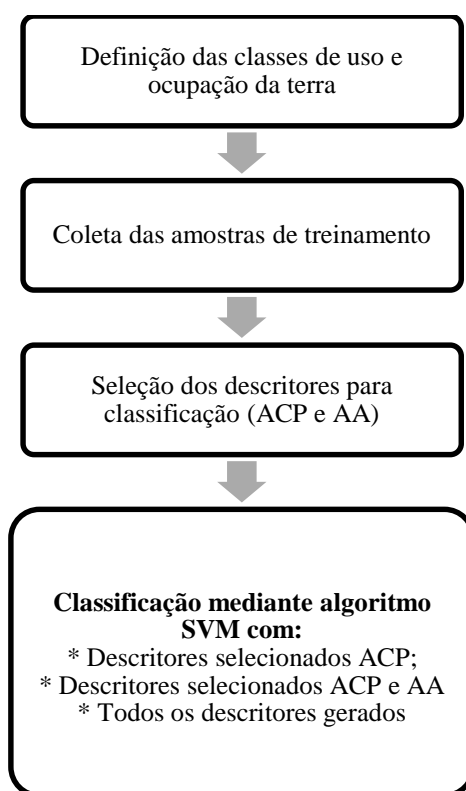
A classificação de imagens digitais pode ser realizada ao nível de abstração do *pixel* ou ainda orientada a objeto geográfico. Esta última é adotada nesse trabalho por ter a vantagem de utilizar dados espaciais e de textura além dos espectrais.

O enfoque nesse trabalho é na classificação supervisionada baseada em exemplos, ou seja, onde o técnico define as classes e apresenta amostras de treinamento ao algoritmo que irá classificar a imagem toda e assim gerar a imagem classificada.

Foram realizados três testes de classificação para cada área de estudo: com os descritores selecionados mediante ACP, com a seleção de descritores da ACP e AA, e todos os descritores gerados a partir da GEOBIA.

A Figura 11 ilustra as etapas seguidas para a realização da classificação:

Figura 11- Etapas para a classificação



Fonte: a autora.

Mediante os elementos de fotointerpretação (tonalidade e cor, forma, tamanho, padrão, textura, associação e sombra) da imagem *RapidEye*, *Google Earth* e verificação à campo, definiram-se as classes de uso e ocupação da terra para a classificação digital.

Para a fotointerpretação a imagem *RapidEye* foi utilizada em diferentes bandas e composições coloridas para as áreas de estudo 1 e 2 ilustradas nas Figuras de 12 a 14.

Definiram-se para a primeira área de estudo 11 classes de uso e ocupação da terra, sendo: Cultivo 1 (milho 1), cultivo 2 (aveia), cultivo 3 (soja), cultivo 4 (milho 2), cultivo 5 (trigo), cultivo 6, cultivo 7, solo exposto, corpos hídricos, cobertura florestal e estradas/construções. As duas classes de milho se devem a cultivos com ciclo fenológico diferentes. Na Figura 12 destaca-se mediante números algumas áreas visitadas em campo.

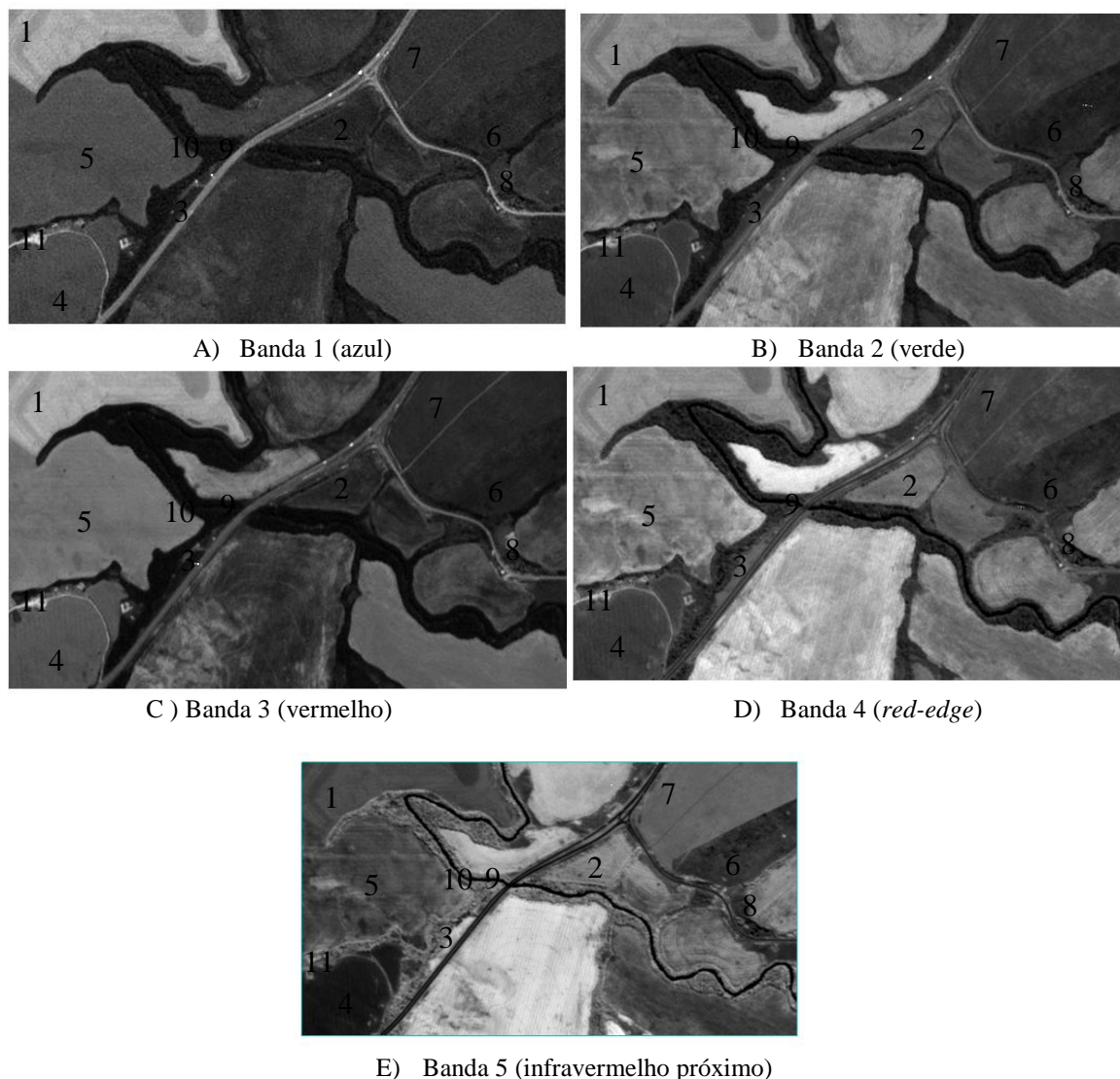
Figura 12 - Ilustração de áreas visitadas em campo



Fonte: a autora.

Legenda: 1 – Cultivo 1 (milho 1); 2 – Cultivo 2 (aveia); 3- Cultivo 3 (soja); 4 – Cultivo 4 (milho 2); 5 – Cultivo 5 (trigo); 6 – Cultivo 6; 7 – Cultivo 7; 8 – Solo exposto; 9 – Corpos hídricos; 10 – Cobertura florestal; 11 – Estradas e construções;

A Figura 13 exhibe as cinco bandas espectrais da imagem *RapidEye*.

Figura 13– Bandas da Imagem *RapidEye* para área de estudo 1

Fonte: a autora

Legenda: 1 – Cultivo 1 (milho 1); 2 – Cultivo 2 (aveia); 3- Cultivo 3 (soja); 4 – Cultivo 4 (milho 2); 5 – Cultivo 5 (trigo); 6 – Cultivo 6; 7 – Cultivo 7; 8 – Solo exposto; 9 – Corpos hídricos; 10 – Cobertura florestal; 11 – Estradas e construções;

A banda 1 (Figura 13A) corresponde ao comprimento de onda do azul. Nessa banda há baixa radiância refletida para cultivos em geral, no entanto, para os cultivos 1, 4, e 5 apresenta radiância mais alta que nos outros cultivos devido ao ciclo fenológico destas, que provavelmente estavam mais avançados. Podem-se notar bem as estradas/construções e o solo exposto nesse comprimento de onda.

A banda 2 (Figura 13B) é relativa ao comprimento de onda do verde. Nela, há diferentes tons de cinza para as vegetações, pode-se notar em tons mais claros os cultivos 1, 2, e 5.

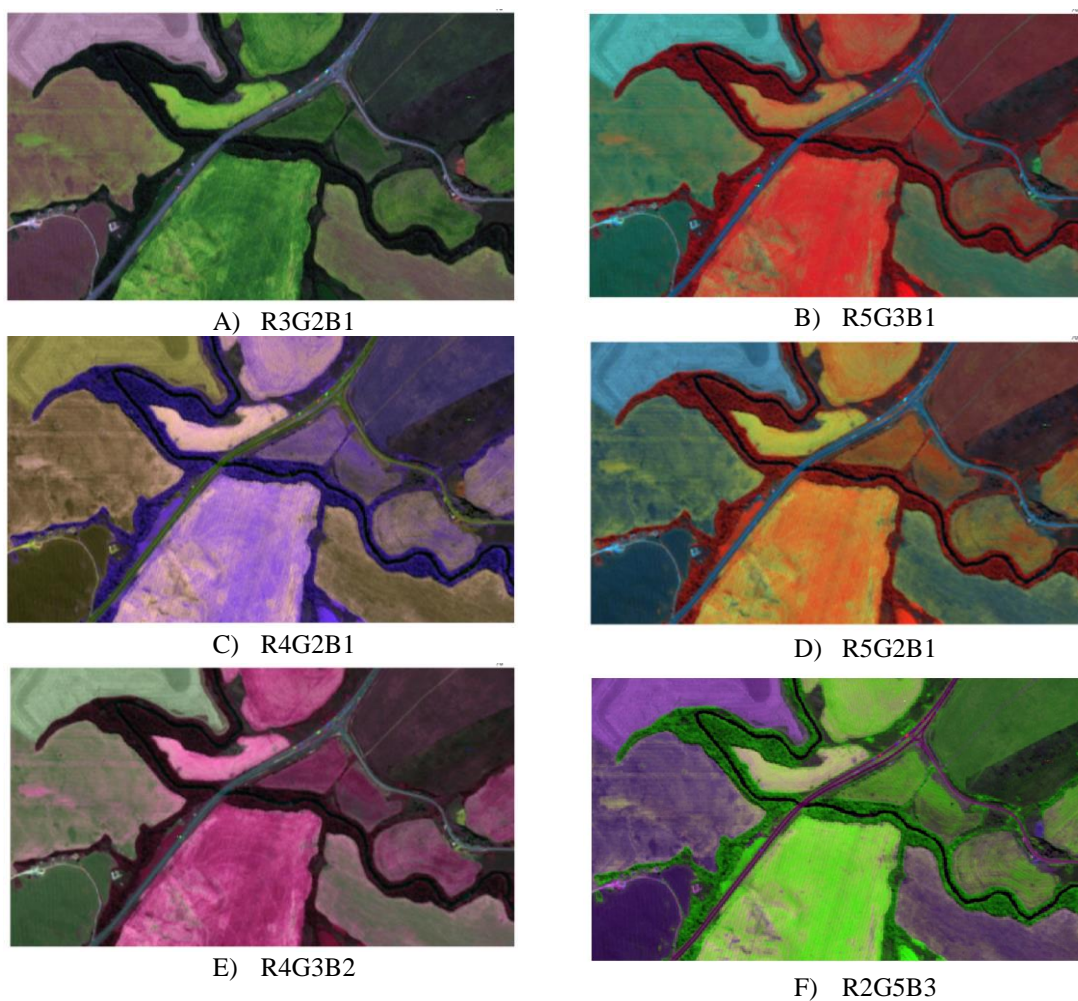
A banda 3 (Figura 13C) corresponde ao comprimento de onda do vermelho, nela também há diferentes nuances para a vegetação, mas com destaque para os cultivos 1, 5, e 4 que aparecem mais definidos que os outros.

A banda 4 (Figura 13D) corresponde ao comprimento de onda do *red-edge*, o limiar entre o vermelho e o infravermelho próximo. Nessa banda evidenciam-se os cultivos 1, 3, e 5.

A banda 5 (Figura 13E) corresponde ao infravermelho próximo. Nela há baixa reflectância para os corpos d'água que aparecem bem escuros. Devido à alta reflectância para a vegetação a cobertura florestal aparece mais clara e os cultivos 1, 2 e 5 também.

A Figura 14 traz algumas composições coloridas da imagem RapidEye da área de estudo 1.

Figura 14 – Composições coloridas da Imagem RapidEye para área de estudo 1



Fonte: a autora.

A Figura 14A é uma composição das bandas dos comprimentos de onda do visível (azul, verde e vermelho), ou seja, verdadeira cor. Geralmente essa composição

apresenta boa correlação com a experiência visual do técnico. Nas Figuras 14B, 14C, 14D, 14E e 14F são correspondentes a composições falsa-cor.

Na Figura 14B é possível notar bem as classes. A cobertura florestal aparece com tom avermelhado, os corpos hídricos num tom azul escuro, os cultivos 1 (milho 1), 4 (milho 2) e 5 (trigo) aparecem em diferentes tons de verde. O cultivo 3 (soja) aparece bem alaranjado. O cultivo 2 (aveia) aparece em tons de laranja e amarelo, em algumas regiões parecido com o cultivo 6. O cultivo 7 aparece num tom escuro e o solo exposto verde limão.

Nas figuras 14C e 14D os corpos hídricos e cobertura florestal são bem perceptíveis, também os cultivos. No entanto, o solo exposto é menos perceptível na figura 13C.







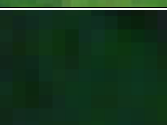


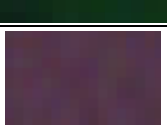


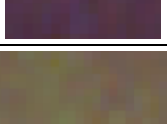
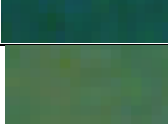

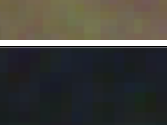


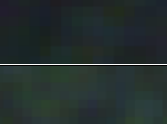
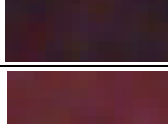

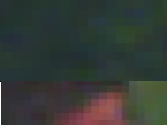











Na Figura 14E percebe-se bem os cultivos, pois aparecem em diferentes nuances. Os corpos hídricos e cobertura florestal não são bem separáveis nessa composição, pois aparecem bem escuros.

Na Figura 14F o cultivo 2 (aveia) fica bem evidente com nuances de verde claro, já os cultivos 1, 4, e 5 não são bem separáveis, apresentam tons de azul semelhantes, bem como o solo exposto. Nessa composição é bem perceptível também os corpos hídricos (cor escura) e a cobertura florestal que aparece bem verde.

As várias composições coloridas auxiliam no processo de análise visual, possibilitando a percepção das várias classes e a separabilidade destas. No quadro 3 há recortes da cena *RapidEye* de cada área de estudo nas composições R3G2B1 e R5G3B1 e fotos de campo.

A Figura 15 ilustra as características espectrais e de textura das classes de uso e cobertura da terra para a área de estudo 1.

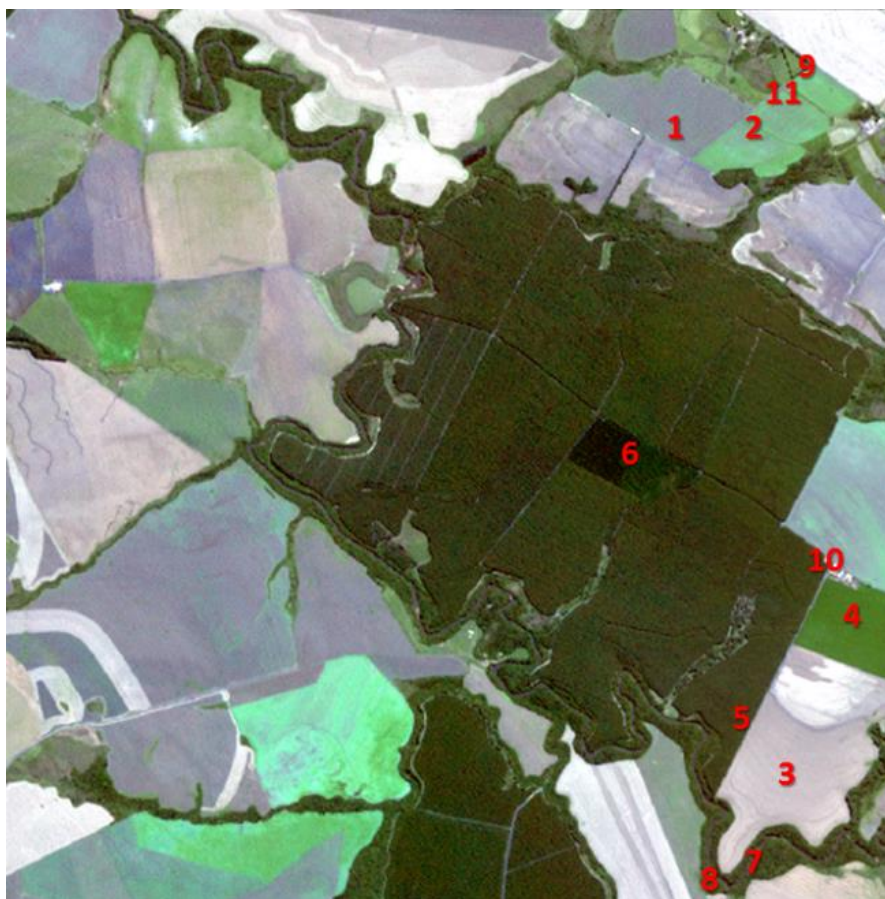
Figura 15 – Chaves de interpretação para área de estudo 1

Classe	R3G2B1	R5G3B1	Foto de campo
Cultivo 1 (milho)			
Cultivo 2 (aveia)			
Cultivo 3			
Cultivo 4 (milho 2)			
Cultivo 5 (trigo)			
Cultivo 6 (não definido)			
Cultivo 7 (não definido)			
Solo exposto			
Corpos hídricos			
Cobertura florestal			
Estradas/construções			

Fonte: a autora.

Para a segunda área de estudo foram definidas 11 classes: cultivo 1, cultivo 2, cultivo 3, cultivo 4, reflorestamento 1, reflorestamento 2, cobertura florestal, corpos hídricos, estradas, construções e solo exposto. A Figura 16 mostra áreas visitadas em campo para a área de estudo 2:

Figura 16- Áreas visitadas em campo (área de estudo 2)



Fonte: a autora.

Legenda: 1 – Cultivo 1; 2 – Cultivo 2; 3 – Cultivo 3; 4 – Cultivo 4; 5 – Reflorestamento 1; 6 – Reflorestamento 2; 7 – Corpos hídricos; 8 – Cobertura florestal; 9 – Estradas; 10 – Construções; 11 – Solo exposto.

A Figura 17 mostra as bandas monocromáticas da imagem *RapidEye* para a área de estudo 2.

Figura 17 – Bandas da imagem *RapidEye* para a área de estudo 2



A) Banda 1 (azul)



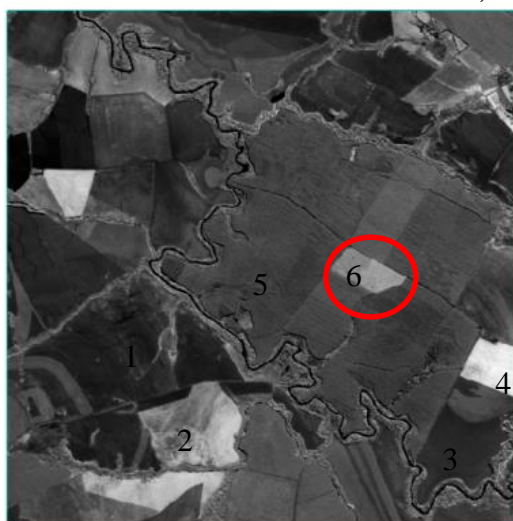
B) Banda 2 (verde)



C) Banda 3 (vermelho)



D) Banda 4 (*red-edge*)



E) Banda 5 (infravermelho próximo)

Fonte: a autora.

Legenda: 1 – Cultivo 1; 2 – Cultivo 2; 3 – Cultivo 3; 4 – Cultivo 4; 5 – Reflorestamento 1; 6 – Reflorestamento 2; 7 – Corpos hídricos; 8 – Cobertura florestal; 9 – Estradas; 10 – Construções; 11 – Solo exposto.

Na Figura 17A, que corresponde ao comprimento de onda do azul, pode-se notar diferentes tonalidades em cinza escura para os reflorestamentos e o cultivo 3 em nuances de cinza claro.

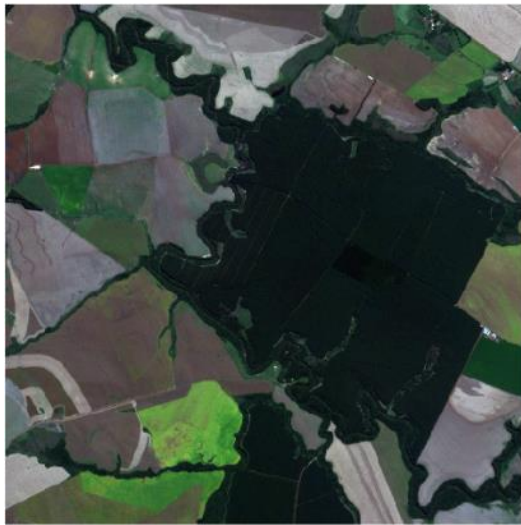
Na banda 2, relativa ao comprimento de onda do verde (Figura 17B), é possível diferenciar alguns cultivos e, sutilmente, o reflorestamento 1 e 2. As estradas e construções também são perceptíveis.

Na Figura 17C, correspondente ao comprimento de onda do vermelho, é possível notar os diferentes tipos de reflorestamento, pois o reflorestamento 2 aparece mais escuro que o 1, pois na faixa do visível a radiância refletida é mais alta. No entanto, não há uma boa distinção do reflorestamento 1 para o cultivo 4 e cobertura florestal. É possível notar as estradas e construções.

A Figura 17D é relativa ao comprimento de onda do *red-edge*. Neste recorte os corpos hídricos aparecem escuros e é possível diferenciar os cultivos 1 e 2. Nessa imagem é possível notar que o reflorestamento central é destacado mais claro, denotando o aumento da reflectância em direção do comprimento de onda se aproximar do infravermelho próximo.

A Figura 17E corresponde ao comprimento de onda do infravermelho próximo. Nesta banda evidenciam-se os corpos hídricos em cor bastante escura contrastando com a cobertura florestal no seu entorno que aparece numa tonalidade mais clara. Os cultivos 4 e 2 aparecem em tonalidade próximo do branco. Os reflorestamentos 1 e 2 são distinguíveis, pois o reflorestamento 2 aparece em tom mais claro que o reflorestamento 1.

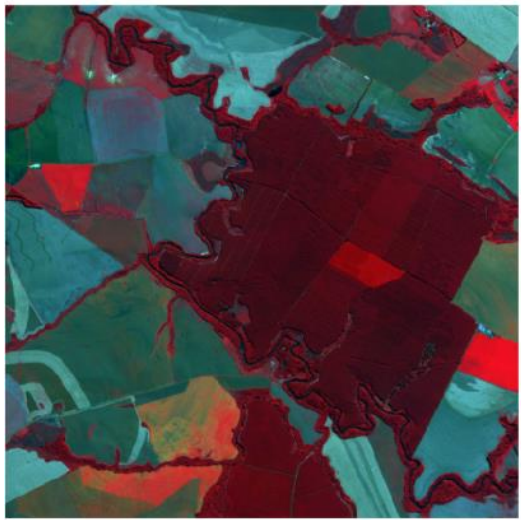
A Figura 18 traz diferentes composições coloridas da imagem *RapidEye* para a área de estudo 2.

Figura 18- Composições coloridas da imagem *RapidEye* para a área de estudo 2

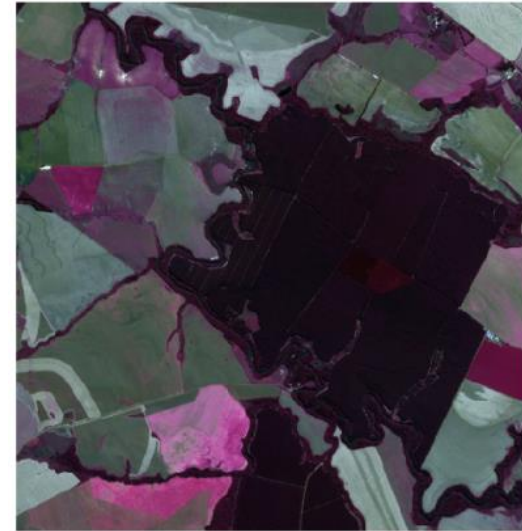
A) R3G2B1



B) R4G2B1



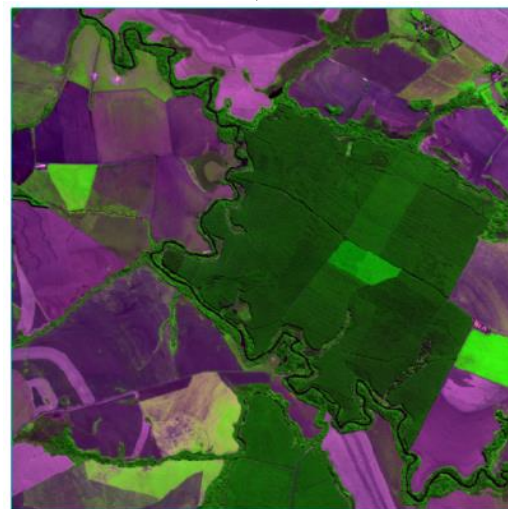
C) R5G3B1



D) R4G3B2



E) R5G2B1



F) R2G5B3

Fonte: a autora.

A Figura 18A corresponde à composição em verdadeira cor, composta pelos comprimentos de onda do vermelho, verde e azul. Nela são distinguíveis os tipos de cultivo, as estradas e construções e sutilmente o reflorestamento 1 e do 2.

As Figuras 18B, 18C, 18D, 18E, e 18F são composições falsa-cor. Na Figura 18B, o cultivo 2 aparece em nuances de amarelo e o cultivo 4 em marrom, dessa forma são distinguíveis dos outros usos.

Na Figura 18C os cultivos aparecem em diferentes nuances de cor, sendo bem separáveis. Os corpos hídricos aparecem escuros e os reflorestamentos e cobertura florestal em tons avermelhados. Nesse aspecto assemelha-se à composição 18E.










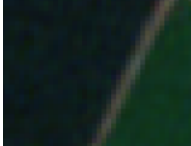





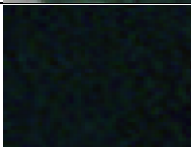
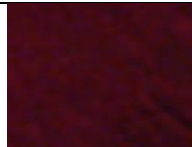




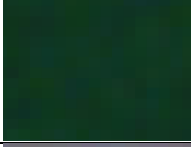


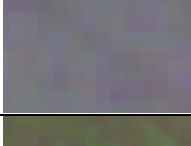
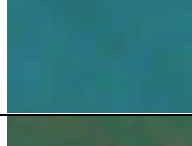

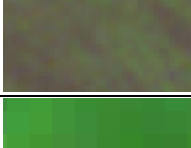



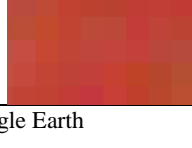

Na Figura 18D é possível distinguir as estradas e construções e os diferentes tipos de reflorestamentos. Os cultivos se apresentam bem distintos, podendo-se observar em verde o cultivo 1, o cultivo 2 em roxo claro, o cultivo 3 em verde bem claro e o cultivo 4 aparece em roxo escuro.

Na Figura 18E, é facilitada a delimitação dos corpos hídricos que aparecem bem escuros.

Após a definição das classes, coletaram-se amostras de treinamento para a classificação. Essas amostras são objetos geográficos criados a partir da segmentação.

A Figura 19 ilustra as características espectrais e de textura das classes de uso e cobertura da terra para a área de estudo 2.

Figura 19 – Chaves de interpretação para área de estudo 2

Classe	R3G2B1	R5G3B1	Foto de campo
Cobertura florestal			
Rio			
Construções			
Estradas			
Solo exposto			
Reflorestamento 1			
Reflorestamento 2			
Cultivo 1			
Cultivo 2			
Cultivo 3			
Cultivo 6			

*Imagem do Google Earth

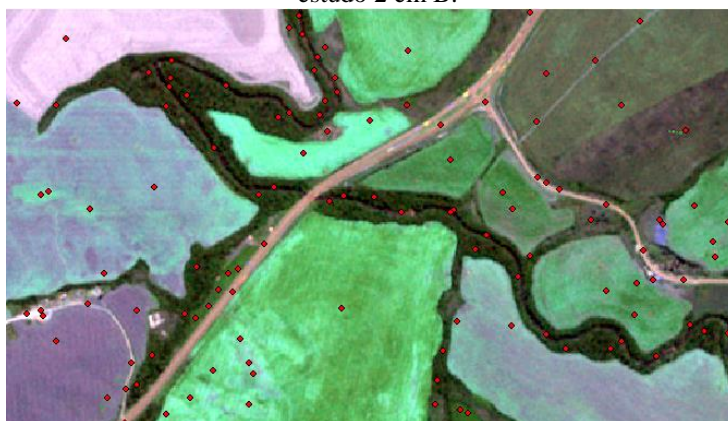
Fonte: a autora.

3.3.2.1 Coleta das amostras de treinamento

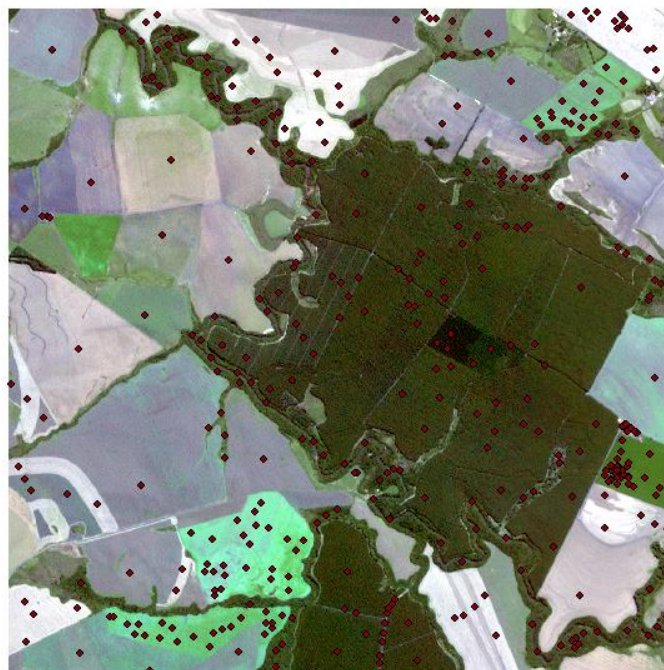
Sabe-se que a coleta das amostras de treinamento é crucial para a classificação supervisionada, pois nela o técnico apresenta ao algoritmo classificador o padrão que deve ser generalizado para a classificação.

Na abordagem orientada a objetos, as amostras de treinamento são exemplos de cada classe. A figura a seguir mostra a localização dos pontos que correspondem às amostras de treinamento. Para a área de estudo 1 foram selecionadas 119 amostras, enquanto para a área 2 foram selecionadas 419 amostras.

Figura 20- Pontos de localização das amostras de treinamento na área de estudo 1 em A e na área de estudo 2 em B.



A) Área de estudo 1



B) Área de estudo 2

Fonte: a autora.

Um banco de dados relacional contendo os valores de todos os descritores para cada amostra selecionada foi gerado. Esses dados foram submetidos à Análise de Componentes Principais e Análise de Agrupamentos para redução da dimensionalidade dos dados e seleção dos descritores para a classificação.

No próximo tópico serão explicadas as metodologias da Análise de Componentes Principais e Análise de Agrupamentos na seleção de descritores para a classificação.

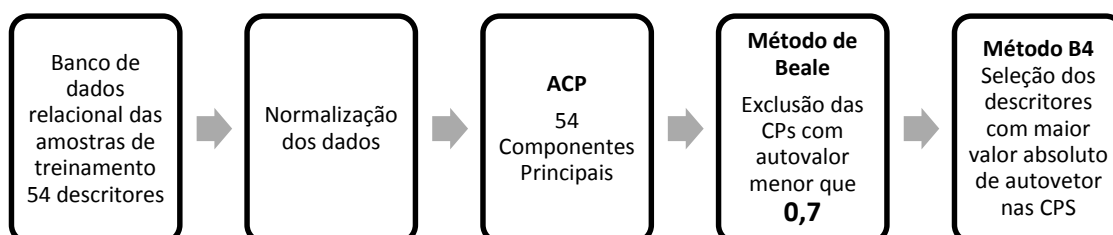
O algoritmo de classificação nesse trabalho corresponde ao *Support Vector Machine* (SVM) que representa um grupo de algoritmos de aprendizado de máquina teoricamente superiores. A atratividade do algoritmo SVM é a sua capacidade de minimizar o chamado risco estrutural ou os erros de classificação, minimizando a probabilidade de erros de classificação em uma distribuição prévia. O método SVM utiliza funções de decisão que aumentam o poder de classificação por flexibilizar a forma da superfície de separação (TSO E MATHER, 2009).

Foram utilizados os parâmetros pré-definidos pelo *software* no algoritmo SVM: Kernel base radial, Gamma na função *kernel* igual a 0,03, e parâmetro de penalidade (*penalty*) igual a 100. Mais informações a respeito do algoritmo SVM podem ser obtidas em Hsu, Chang, e Lin (2010).

3.3.3 Análise de Componentes Principais

Com o objetivo de reduzir a dimensionalidade dos dados e efetuar o descarte de variáveis que não fossem significativas, aplicou-se a Análise de Componentes Principais (ACP) no banco de dados relacional das amostras de treinamento obtido a partir da análise orientada a objeto geográfico. A Figura 21 ilustra os procedimentos adotados na ACP.

Figura 21 - Procedimentos para Análise de Componentes Principais



Fonte: A autora.

Os dados obtidos na GEOBIA possuem diferentes unidades de medida, portanto realizou-se a normalização dos mesmos com o propósito de minimizar os problemas decorrentes do uso de distintas unidades de medidas e dispersão entre as variáveis, atribuindo uma nova escala para o conjunto de dados em um valor especificado de 0,0 a 1,0.

Com os dados normalizados realizou-se a ACP, transformando os dados em um novo sistema de coordenadas com eixos ortogonais. A partir das n variáveis ortogonais são geradas n componentes principais (CPs), de modo que a variância explicada seja sempre maior nas primeiras CPs, não sendo necessário o uso de todas as CPs, pois em diversos conjuntos de dados existem variáveis redundantes que podem ser descartadas sem perdas, porque elas não fornecem informações ou porque são combinações de outras variáveis. (JOLLIFFE, 1972).

O autor Jolliffe (1972, 1973) investigou métodos de seleção de variáveis por meio da Análise de Componentes Principais, como os que serão utilizados nesse trabalho: Método de Beale e Método B4.

Consistindo que λ significa autovalor, o método de Beale considera que sendo λ_0 um valor limite para a inclusão de variáveis, a exclusão acontecerá para todas as CPs desde que $\lambda_i < \lambda_0$, Jolliffe (1972) indica que o valor de λ_0 deve ser igual a 0,70. Ao invés de usar λ_0 , pode-se utilizar k como sendo o número de variáveis onde a proporção da variabilidade acumulada seja $\alpha = 0,8$, no entanto, considera-se o método que usa λ_0 como o de melhor desempenho e por isso foi utilizado nesse trabalho (JOLLIFFE, 1972). O método de Beale foi utilizado para a definição de quais componentes principais utilizar, ou seja, todas as CPs que possuem autovalor a partir de 0,7.

O método B4 desenvolvido por Jolliffe (1972) explora a propriedade de que cada componente principal representa uma parte da variância dos dados para selecionar as variáveis mais relevantes para descrever a variabilidade do conjunto original. É útil na seleção das principais CPs e na identificação das variáveis originais que estão mais fortemente correlacionadas a elas. Envolve o uso das k CPs definidas pelo método de Beale, então para cada uma destas CPs seleciona-se somente a variável de maior valor absoluto de autovetor, no caso de a mesma variável apresentar maior valor absoluto de autovetor em CPs diferentes, selecionando-se a de segundo maior valor absoluto da CP correspondente. As variáveis não selecionadas são descartadas. Desta forma, o número de variáveis selecionadas é igual ao número k de CPs definidas a partir do método de Beale.

O procedimento B4 pode ser descrito como: execução da ACP sobre o conjunto de dados definido; definição do número k de variáveis para serem selecionadas; para cada uma das CPs: determinação e seleção da variável com maior coeficiente absoluto. Ou seja, seleciona-se para cada uma das CPs qual o descritor com maior valor absoluto para o autovetor.

Dessa forma, utilizou-se nesse trabalho a Análise de Componentes Principais para redução da dimensionalidade dos dados da matriz de variância, o método de Beale para definição das componentes principais a serem usadas, e o método B4 para definição das variáveis com maior relevância nas CPs definidas anteriormente e descarte do restante.

3.3.4 Análise de Agrupamentos

Conforme mencionado no referencial teórico a Análise de Agrupamentos visa agrupar variáveis com características comuns, definindo a estrutura dos dados e colocando as observações mais parecidas em grupos.

De acordo com Hair et al. (2009), a Análise de Agrupamentos (AA) envolve ao menos três etapas:

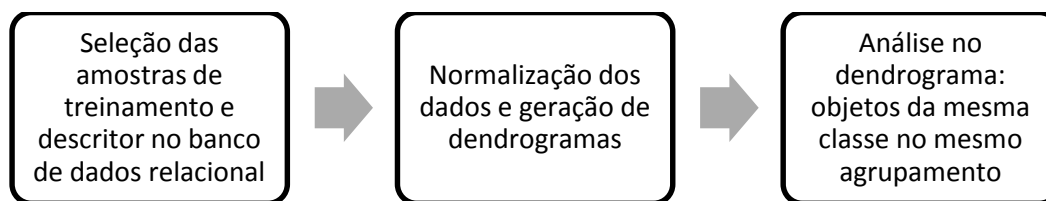
1. Medida de alguma forma de similaridade ou associação entre as variáveis para determinar quantos grupos existem na amostra.
2. Realização do agrupamento, em que as variáveis são divididas em grupos ou agrupamentos.
3. Estabelecer o perfil das variáveis para determinar a composição dos grupos.

Neste trabalho utilizou-se a AA com o intuito de identificar os descritores que possibilitavam melhor separabilidade entre classes e assim melhorar a classificação.

Para atender o objetivo, foi realizada a AA no banco de dados relacional das amostras de treinamento para cada descritor gerado, analisando nos dendrogramas quais descritores proporcionavam a formação de grupos heterogêneos das classes de uso e ocupação das terras.

A Figura 22 demonstra os procedimentos considerados para a AA. Extraiu-se do banco de dados relacional os valores correspondentes a cada um dos descritores selecionados, estes foram normalizados e processados para obtenção dos dendrogramas contendo os agrupamentos.

Figura 22 - Procedimentos para Análise de Agrupamentos



Fonte: A autora.

Utilizou-se nesse trabalho como medida de similaridade distância euclidiana que é uma medida de dissimilaridade. Quanto menor o valor da distância euclidiana entre duas variáveis, mais similares elas são, ao passo que quanto maior o valor, a similaridade entre as variáveis é menor.

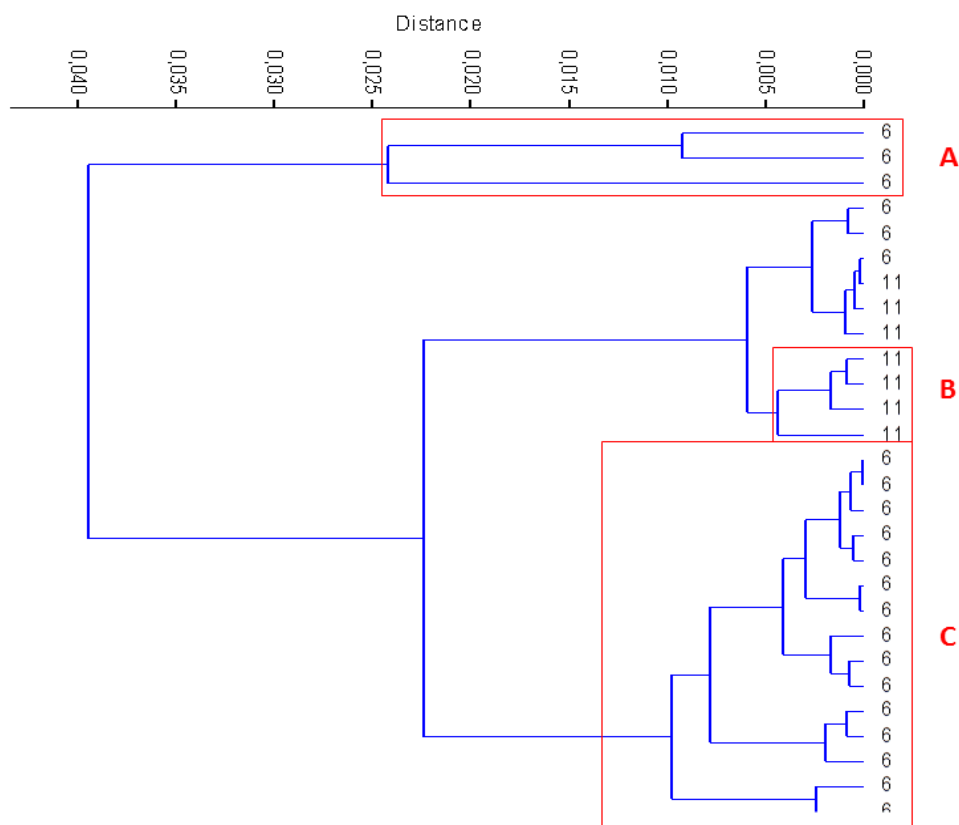
Existem três métodos de agrupamentos hierárquicos: o do vizinho mais próximo, do vizinho mais distante, e o método de Ward. Este último foi utilizado nesse trabalho e tem como premissa a distância euclidiana quadrática.

Como resultado, a Análise de Agrupamentos gera dendrogramas representando os grupos e subgrupos criados. Um corte no dendrograma a qualquer distância euclidiana produz classificação em k subgrupos. Um par de indivíduos que seja incluído numa mesma classe em qualquer etapa do processo não poderá mais ser separado em etapas posteriores, pois estas consistem em fusões de classes já existentes (ISA/UTL, 2010).

A análise do dendrograma pode ser realizada a partir de um corte uma distância euclidiana e a partir disso a verificação dos grupos. Nesse trabalho não considerou-se a distância euclidiana, mas sim se formavam grupos homogêneos de objetos de determinada classe.

A Figura 23 ilustra um exemplo de dendrograma. Nela, são destacados três grupos heterogêneos, A, B e C. Esses grupos são homogêneos em sua composição, pois possuem somente elementos de uma classe (A – classe 6; B – classe 11; C – classe 6). Já o grupo entre A e B não é homogêneo, pois é formado por duas classes de elementos distintos 6 e 11. Nota-se que os grupos A, B e C, são formados de acordo com os objetos que constituem grupos e não por uma determinada distância euclidiana.

Figura 23 - Dendrograma



Fonte: a autora.

3.3.5 Validação

A validação é o método de medir a confiabilidade dos resultados, importante para comprovação da qualidade da classificação digital.

A validação foi realizada avaliando-se a concordância dos resultados de classificação com as classes de uso e cobertura da terra levantados por amostras de campo, por meio da matriz de confusão. Foram avaliados a acurácia geral, e o índice *kappa*.

De acordo com Congalton (1991), a maneira mais comum para representar a validação de classificação de dados de Sensoriamento Remoto é sob a forma de uma matriz de erro, que é uma matriz quadrada de números estabelecida em linhas e colunas que expressam o número de unidades de amostragem atribuídas a uma categoria específica em relação à categoria real como verificado em solo. Nas colunas são apresentados os dados de referência, e nas linhas a classificação gerada.

A acurácia global da matriz de erros é calculada dividindo o total correto (isto é, a soma dos valores da diagonal) pelo número total de *pixels* na matriz de confusão. A precisão das categorias individuais pode ser calculada de forma semelhante, divide-se o número de *pixels* corretos pelo número total de *pixels* na linha ou coluna correspondente. No resultado da classificação distinguem-se dois tipos principais de erro: erros de omissão e erros de inclusão.

Moreira (2011) argumenta que a partir da matriz de confusão é possível obter a estatística *kappa* que tem como vantagem a inclusão de todos os elementos da matriz de confusão e não somente os elementos da diagonal principal. O autor argumenta que os valores acima de 0,60 são muito bons, e acima de 0,80 são considerados excelentes. Elabora-se a matriz de confusão coletando novas amostras de treinamento, que serão confrontadas com a imagem classificada.

A coleta das amostras para validação se baseou nas chaves de fotointerpretação analisadas nas imagens *RapidEye* e do *Google Earth*, e também de visita ao campo, conforme foi explicitado anteriormente quando apresentados as Figuras 18 e 19.

4 Resultados e discussão

Nessa seção são apresentados e discutidos os resultados obtidos para a Análise Orientada a Objetos Geográficos, seleção de descritores, e classificação supervisionada orientada a objetos.

4.1 Análise Orientada a Objeto Geográfico (GEOBIA – *Geographic Object Based Image Analysis*)

A segmentação é passo importante na Análise Orientada a Objeto Geográfico, dependendo dela o sucesso ou fracasso do processamento. Foram, portanto, testados diferentes valores ao nível de escala e de fusão de segmentação para definir qual os mais apropriados para a classificação das duas áreas de estudo.

A Figura 24 mostra testes com diferentes valores de segmentação (escala e fusão) para a área de estudo 1. À esquerda têm-se a imagem segmentada sobreposta pelos vetores, e à direita os vetores gerados.

Na Figura 24 observam-se os três testes realizados. No teste A com escala menor, muitos objetos pequenos foram gerados, sem a separação destes nas classes de interesse, como mostram as setas nas figuras A1 e B1.


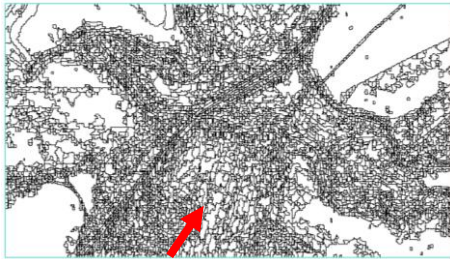

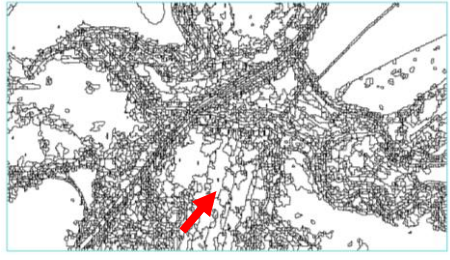


No teste B, foram gerados objetos de tamanho médio, indicados na setas nas figuras B1 e B2.

Buscando objetos maiores e que caracterizassem melhor as classes foi realizado o teste C, com escala 55 e fusão 75. Observam-se nas setas nas figuras C1 e C2 que os objetos geográficos foram agregados, caracterizando melhor as classes de interesse.

Portanto os valores do teste C da Figura 24 foram definidos como os mais adequados para a área de estudo 1. Gerou 1802 objetos geográficos e 54 descritores.

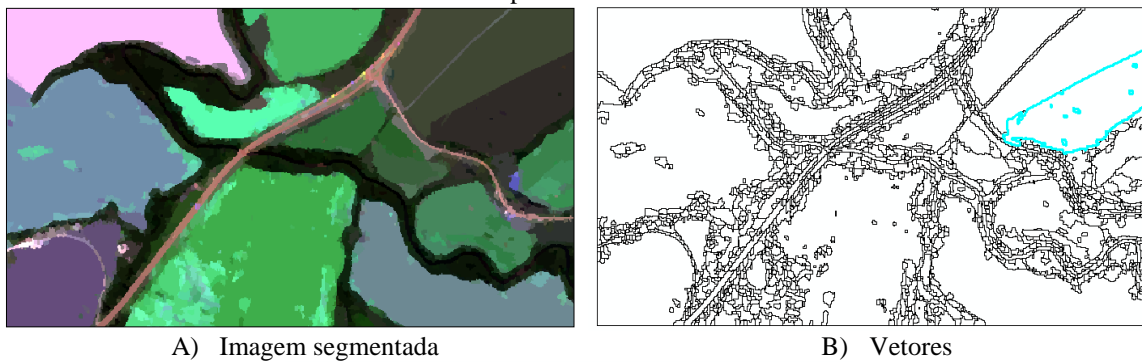
A Figura 25 ilustra a imagem segmentada e os vetores gerados para a área de estudo 1.

Figura 24 – Testes de segmentação com diferentes valores de escala e fusão (área de estudo 1)

Testes	Imagem segmentada + vetores	Vetores
A) Escala 30 Fusão 60	 A1	 A2
B) Escala 40 Fusão 60	 B1	 B2
C) Escala 55 Fusão 75	 C1	 C2

Fonte: A autora.

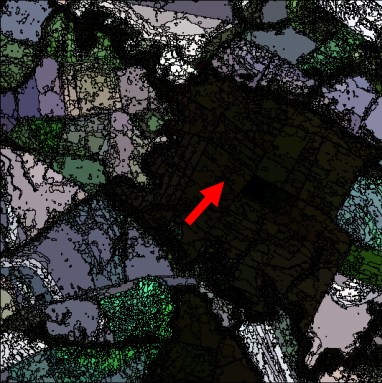
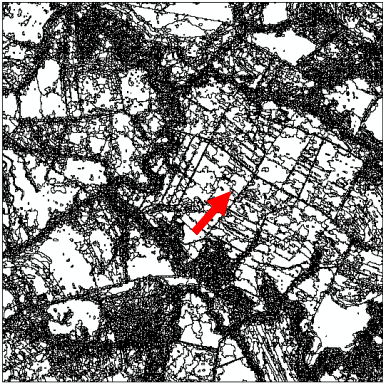
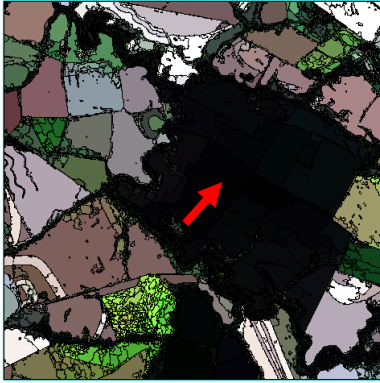
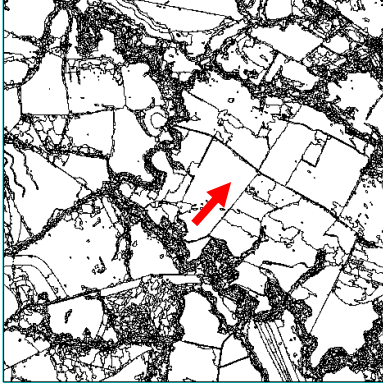

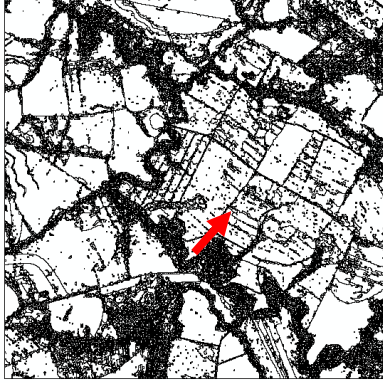
Figura 25 – Área de estudo 1: A) Imagem segmentada, B) vetores, e C) Imagem segmentada sobreposta por vetores



Fonte: a autora.

Na Figura 26 são apresentados os resultados da aplicação de diferentes testes de segmentação para a área de estudo 2. Foram realizados três testes, o primeiro com escala 30 e fusão 65 (Teste A), o segundo com escala 50 e fusão 80 (Teste B), e o terceiro com escala e fusão (Teste C).

Figura 26 – Testes de segmentação com diferentes valores de escala e fusão (área de estudo 1)

Testes	Imagem segmentada + vetores	Vetores
A) Escala 30 Fusão 65	 A1	 A2
B) Escala 50 Fusão 80	 B1	 B2
C) Escala 50 Fusão 40	 C1	 C2

Fonte: A autora.

No teste A foi gerado grande número de polígonos, ou seja segmentação excessiva da imagem, com muitos pequenos objetos geográficos, como nota-se na Figura 25 A1 e A2.

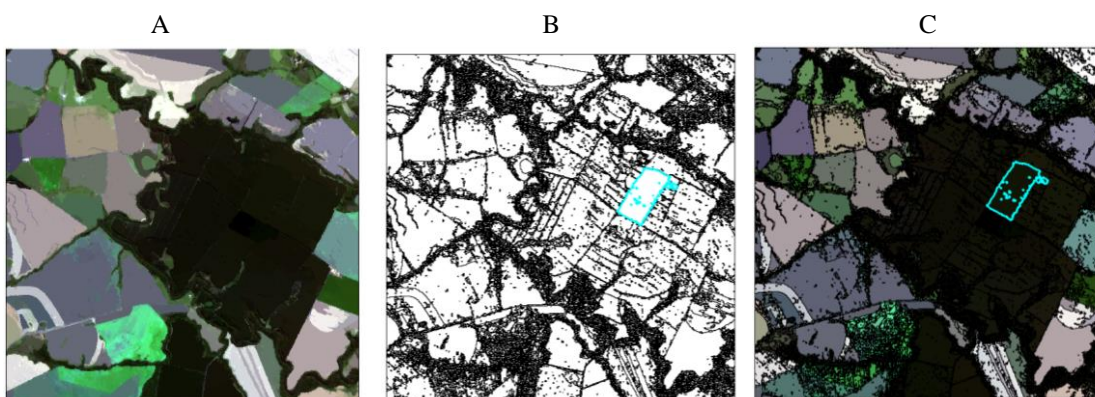
Em contra partida, o teste B com valores de escala e fusão respectivamente de 50 e 80, generalizou demais os objetos geográficos, como se vê nas figuras B1 e B2 do quadro 6. Foram agregados objetos geográficos de classes diferentes, por exemplo, as estradas ao reflorestamento 1, e corpos hídricos à cobertura florestal.

Desta forma foi feita a seleção da terceira segmentação com escala 50 e fusão 40 (Teste C). Nota-se nas figuras C1 e C2 que há objetos geográficos pequenos, mas em menor quantidade que o teste A, e menor agregação de objetos de classes diferentes como no teste B. Foram gerados para a área de estudo 2 um número igual a 2 30.846 objetos geográficos e banco de dados relacional com 54 descritores.

A Figura 27 ilustra para a área de estudo 2: a imagem segmentada (27A), os vetores gerados (27B), e a imagem sobreposta pelos vetores (27C).

Posterior à definição dos parâmetros e geração da imagem segmentada, vetores e banco de dados relacional foram coletadas as amostras de treinamento. O banco de dados das amostras foi submetido a análises para seleção dos descritores mais adequados à classificação, cujos resultados serão apresentados na próxima seção 6.2.

Figura 27 – Área de estudo 2: A) Imagem segmentada, B) vetores, e C) Imagem segmentada sobreposta por vetores



Fonte: a autora.

4.3 Seleção de descritores para a classificação

Os descritores utilizados na classificação orientada a objeto geográfico impactam no resultado e qualidade da classificação, portanto buscou-se selecionar os descritores para obter melhor acurácia na classificação.

Foram utilizados os métodos de Análise de Componentes Principais (ACP) e Análise de Agrupamentos (AA) para seleção de descritores.

4.3.1 Seleção de descritores mediante Análise de Componentes Principais (ACP)

Conforme mencionado na metodologia para redução da dimensionalidade no banco de dados das amostras de treinamento, foi realizada a Análise de Componentes Principais.

Abson et al. (2012), argumentam que a escolha do número de componentes principais a serem utilizadas em determinado trabalho é subjetiva, e geralmente baseia-se na facilidade da interpretação das CPs selecionadas.

A facilidade da interpretação de dados das CPs pode estar relacionada a complexidade da área de estudo e as várias classes existentes.

Dessa maneira, para definir quantas componentes principais utilizar e quais descritores originais se correlacionam a elas, adotou-se os métodos Beale e B4 propostos por Joliffe (1972,1973), pois estes métodos proporcionam interpretação clara dos dados das CPs.

Como apresentado na metodologia a seleção das componentes principais a partir do método Beale são definidas mediante o autovalor, que deve ser maior ou igual a 0,7. O método B4, conforme descrito na metodologia relaciona a cada componente principal uma variável original mediante a seleção do descritor de maior valor absoluto de autovetor de cada uma das componentes principais.

A seguir são apresentados os resultados da seleção dos descritores a partir da ACP para as áreas de estudo 1 e 2.

4.3.1.1 Área de estudo 1

O banco de dados relacional das amostras de treinamento da área de estudo 1 compõe-se de 54 descritores e 120 amostras de objetos geográficos. A ACP possibilitou

a geração de novas variáveis a partir dos 54 descritores originais, essas novas variáveis denominam-se de componentes principais.

A Tabela 2 traz os dados de autovalor, variância explicada e variância acumulada das componentes principais da área de estudo 1. Analisando a tabela é possível notar que as primeiras componentes principais possuem os maiores valores, e as últimas componentes principais têm os menores valores.

Tabela 1 - Autovalor e variância explicada e acumulada das componentes principais da área de estudo 1
(Continua)

CP	Autovalor	Variância explicada	Variância acumulada
CP1	35.113,00	25,820%	25,820%
CP2	21.500,00	9,680%	35,500%
CP3	19.969,00	8,350%	43,850%
CP4	178.801,00	6,695%	50,544%
CP5	150.909,00	4,769%	55,313%
CP6	142.818,00	4,271%	59,584%
CP7	131.169,00	3,603%	63,187%
CP8	125.324,00	3,289%	66,476%
CP9	119.780,00	3,005%	69,481%
CP10	114.551,00	2,748%	72,229%
CP11	112.334,00	2,643%	74,872%
CP12	104.729,00	2,297%	77,168%
CP13	100.977,00	2,135%	79,304%
CP14	0,9847	2,031%	81,334%
CP15	0,9207	1,775%	83,109%
CP16	0,9054	1,717%	84,826%
CP17	0,8496	1,511%	86,338%
CP18	0,8073	1,365%	87,703%
CP19	0,7846	1,289%	88,992%
CP20	0,7249	1,100%	90,090%
CP21	0,7158	1,073%	91,165%
CP22	0,6841	0,980%	92,140%
CP23	0,6643	0,924%	93,069%
CP24	0,6531	0,893%	93,962%
CP25	0,6074	0,773%	94,735%
CP26	0,5731	0,688%	95,422%
CP27	0,5477	0,628%	96,051%
CP28	0,5072	0,539%	96,589%
CP29	0,4872	0,497%	97,086%
CP30	0,4593	0,442%	97,528%
CP31	0,4404	0,406%	97,934%
CP32	0,3995	0,334%	98,268%
CP33	0,3955	0,328%	98,596%
CP34	0,3397	0,242%	98,838%

(Conclusão)			
CP	Autovalor	Variância explicada	Variância acumulada
CP35	0,3290	0,227%	99,064%
CP36	0,3093	0,200%	99,270%
CP37	0,2757	0,159%	99,424%
CP38	0,2483	0,129%	99,553%
CP39	0,2431	0,124%	99,677%
CP40	0,1974	0,082%	99,758%
CP41	0,1658	0,058%	99,816%
CP42	0,1473	0,045%	99,861%
CP43	0,1391	0,040%	99,900%
CP44	0,1161	0,028%	99,930%
CP45	0,1134	0,027%	99,957%
CP46	0,0986	0,020%	99,977%
CP47	0,0842	0,015%	99,992%
CP48	0,0334	0,002%	99,995%
CP49	0,0277	0,002%	99,996%
CP50	0,0268	0,001%	99,998%
CP51	0,0204	0,001%	99,999%
CP52	0,0201	0,001%	99,999%
CP53	0,0174	0,001%	100,000%
CP54	0,000000000000150	0,000%	100,000%

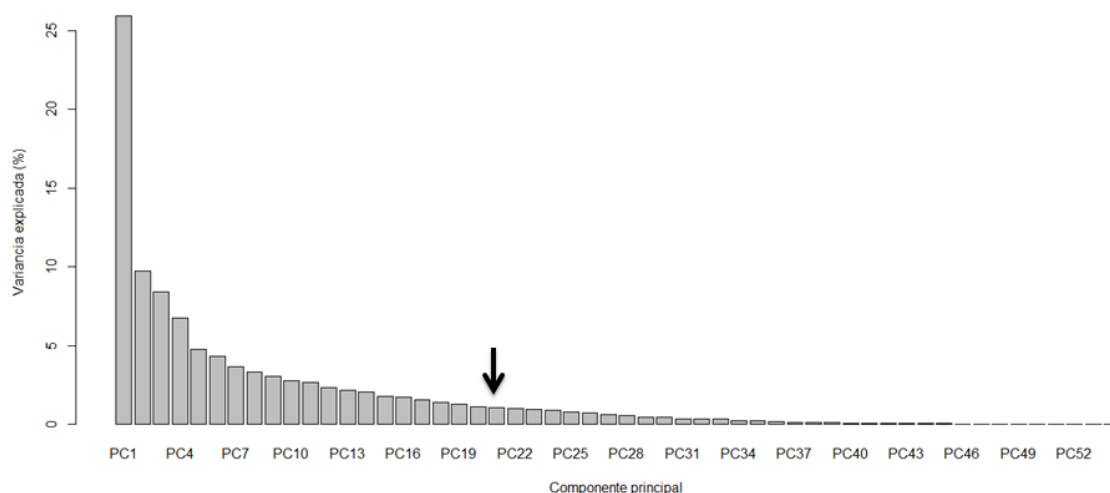
Fonte: a autora.

Utilizando os critérios do método de Beale foram selecionadas as primeiras 21 CPs da Tabela 2, pois essas possuem variância igual ou superior a 0,7, e juntas somam mais de 91% da variância explicada, sobrando para as outras 33 CPs somente 9% da variância explicada.

A Figura 28 mostra o gráfico de variância explicada pelas componentes principais da área de estudo 1.

Nota-se que o valor de variância vai decrescendo, sendo as primeiras CPs as que possuem maior variância. A seta na figura indica as 21 CPs selecionadas, percebe-se que o valor para as CPs de 22 a 54 são muito próximos de zero e portanto de acordo com a metodologia foram excluídas.

Figura 28 - Gráfico de variância explicada pela ACP (área de estudo 1)



Fonte: a autora.

Para o método B4 utiliza-se o critério de maior valor absoluto de autovetor para relacionar a cada componente principal uma variável original (descriptor).

A Tabela 2 mostra para cada uma das 54 CPs o autovetor com maior valor absoluto de cada descriptor. Nota-se na Tabela 2 que cada uma das 54 componentes principais corresponde a um descriptor provindo do banco de dados relacional da GEOBIA.

Tabela 2 - Descritores de maior autovetor absoluto correspondentes as componentes principais geradas

(Continua)

CP	Autovetor	Descriptor	Tipo
1	0,258	Entropia de textura - Banda 4	Textura
2	-0,387	Circuncidade	Espacial
3	0,416	Convexidade	Espacial
4	0,481	Média dos números digitais - Banda 5	Espetral
5	0,429	Variância de textura - Banda 4	Textura
6	-0,438	Direção principal	Espacial
7	0,403	Intervalo de textura - Banda 4	Textura
8	-1.950.003.421,540	Eixo menor	Espacial
9	9.426.361.844,717	Valor máximo de pixel - Banda 2	Espetral
10	-0,367	Desvio padrão dos números digitais - Banda 5	Espetral
11	-0,467	Variância de textura - Banda 1	Textura
12	0,522	Eixo maior	Espacial
13	-0,301	Intervalo de textura - Banda 5	Textura
14	0,387	Variância de textura - Banda 5	Textura
15	0,328	Retangularidade	Espacial
16	-0,390	Valor mínimo de pixel - Banda 5	Espetral
17	-0,548	Desvio padrão dos números digitais - Banda 3	Espetral
18	-0,445	Área	Espacial

				(Conclusão)
CP	Autovetor	Descritor	Tipo	
21	-0,299	Valor máximo de pixel - Banda 4		Espectral
22	-0,283	Desvio padrão dos números digitais - Banda 1		Espectral
23	0,418	Valor máximo de pixel - Banda 5		Espectral
24	-0,316	Entropia de textura - Banda 5		Textura
25	0,578	Entropia de textura - Banda 1		Textura
26	-0,365	Intervalo de textura - Banda 1		Textura
27	-0,304	Entropia de textura - Banda 2		Textura
28	0,483	Intervalo de textura - Banda 3		Textura
29	0,365	Desvio padrão dos números digitais - Banda 2		Espectral
30	-0,503	Intervalo de textura - Banda 2		Textura
31	0,312	Comprimento		Espacial
32	-0,258	Entropia de textura - Banda 3		Textura
33	-0,237	Fator de forma		Espacial
34	0,512	Solidez		Espacial
35	-0,509	Valor máximo de pixel - Banda 3		Espectral
36	0,345	Valor mínimo de pixel - Banda 1		Espectral
37	0,478	Valor máximo de pixel - Banda 1		Espectral
38	0,405	Valor mínimo de pixel - Banda 3		Espectral
39	0,235	Compacidade		Espacial
40	0,471	Número de furos		Espacial
41	-0,417	Valor mínimo de pixel - Banda 2		Espectral
42	-0,145	Média dos números digitais - Banda 3		Espectral
43	0,365	Alongamento		Espacial
44	-0,431	Valor mínimo de pixel - Banda 4		Espectral
45	-0,195	Textura média - Banda 1		Textura
46	0,317	Textura média - Banda 2		Textura
47	-0,363	Média dos números digitais - Banda 4		Espectral
48	0,500	Média dos números digitais - Banda 1		Espectral
49	0,579	Textura média - Banda 5		Textura
50	-0,603	Relação área/furos		Espacial
51	-0,327	Textura média - Banda 3		Textura
52	0,300	Textura média - Banda 4		Textura
53	0,012	Desvio padrão dos números digitais - Banda 4		Espectral
54	0,136	Variância de textura - Banda 2		Textura

Fonte: A autora.

Para cada uma das CPs selecionadas mediante o método de Beale selecionou-se um descritor por meio do método B4. Portanto, foram selecionados os 21 primeiros descritores da Tabela 2 para serem utilizados posteriormente.

Dos 21 descritores selecionados 7 são espectrais, 7 são de textura, e 7 são espaciais.

Dos descritores espectrais selecionados 3 correspondem ao comprimento de onda do infravermelho próximo: média dos números digitais – banda 5, desvio padrão dos números digitais – banda 5, e valor mínimo de pixel – banda 5. Um descritor corresponde ao comprimento de onda do *red-edge* (valor máximo de pixel – banda 4). Dois descritores correspondem ao comprimento de onda do verde: valor máximo de pixel – banda 2 e média dos números digitais – banda 2. Um descritor corresponde ao comprimento de onda do vermelho: desvio padrão dos números digitais – banda 3.

É provável que a seleção destes descritores ocorreu devido a área de estudo 1 possuir grandes áreas de vegetação, pois estes comprimentos de onda são caracterizados por alta reflectância para a vegetação. E também devido à baixa reflectância que destaca algumas classes como estradas, e corpos hídricos, isso é verificado principalmente pelo descritor valor mínimo de pixel na banda 5.

Com relação aos descritores de textura como mostra a Quadro 4: 1 corresponde ao comprimento de onda do azul (banda 1): variância de textura da banda 1. Um descritor corresponde ao comprimento de onda do vermelho (banda 3): 1 ao comprimento de onda do verde (banda 2), 1 ao comprimento de onda do vermelho, 3 ao comprimento de onda do *red-edge* (banda 4), e 2 ao comprimento de onda do infravermelho próximo (banda 5). Destaca-se que a banda 4 do comprimento de onda do *red-edge* é a que apresenta maior número de descritores.

Quadro 3 – Descritores de textura selecionados

Bandas espectrais	Descritores de textura selecionados		
	Variância de textura	Intervalo de textura	Entropia de textura
Banda 1	X	-	-
Banda 2	-	-	-
Banda 3	X	-	-
Banda 4	X	X	X
Banda 5	X	X	-

Fonte: a autora.

Percebe-se que não foram selecionados nenhum descritor de textura correspondentes ao comprimento de onda do verde, e o descritor de textura média não foi selecionado para nenhuma das classes.

No entanto, nota-se na segunda coluna do Quadro 4 o descritor variância de textura aparece para todas as bandas com exceção da banda 2. A banda 1 possui valores altos e intermediários de variância que caracterizam alguns polígonos de estradas e construções e cobertura florestal. A banda 3 não possui valores altos de variância, no

entanto possui valores intermediários que caracterizam polígonos da classe construções e estradas. A banda 4 possui valores altos de variância para os diversos tipos de cultivos. A banda 5 tem altos valores de variância para polígonos correspondentes a cobertura florestal, corpos hídricos e estradas, o que justifica-se mediante a análise visual da banda em questão, onde os corpos hídricos e estradas aparecem com textura lisa e cores escuras, e a cobertura florestal possui textura rugosa e em tons de cinza. A análise numérica aqui citada foi realizada no banco de dados relacional, onde os descritores são apresentados por região.

Possivelmente, o descritor de variância de textura correspondente a classe 2 não foi selecionado porque não caracteriza nenhuma classe especificamente.

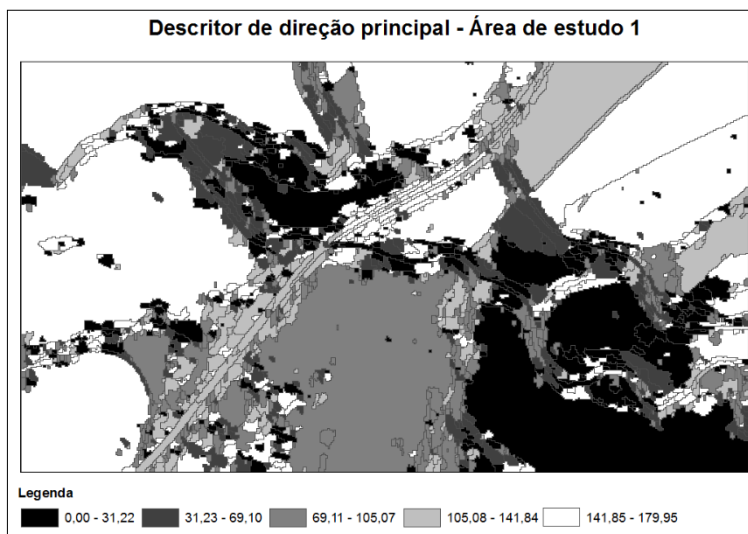
Já o descritor entropia de textura aparece somente para a banda 4, pois é nessa banda que aparecem os maiores valores de entropia em relação as outras bandas espectrais, e nessa podemos visualizar maior desorganização dos objetos. Os descritores de intervalo de textura para as classes 4 e 5 possivelmente foram selecionados devido a caracterizarem bem os corpos hídricos, cobertura florestal e estradas e construções.

Observa-se que dos descritores espectrais e de textura a maioria dos descritores selecionados correspondem às bandas dos comprimentos de onda do *red-edge* (banda 4) e infravermelho próximo (banda 5). Isso se deve ao fato desses comprimentos de onda possuírem alta reflectância para vegetação que se encontra em boa parte dessa área de estudo.

A discussão agora abordará os descritores espaciais, ou seja, àqueles que auxiliam nas formas dos objetos gerados por meio da segmentação. Foram selecionados 7 descritores espaciais selecionados foram: direção principal, convexidade, alongamento, área, eixo maior, eixo menor, e retangularidade.

O descritor espacial “direção principal” possui valores altos para algumas áreas de cultivo, caracterizando-as, como nota-se na representação deste descritor. Nota-se na Figura 29, as regiões em branco com o maior valor de direção principal.

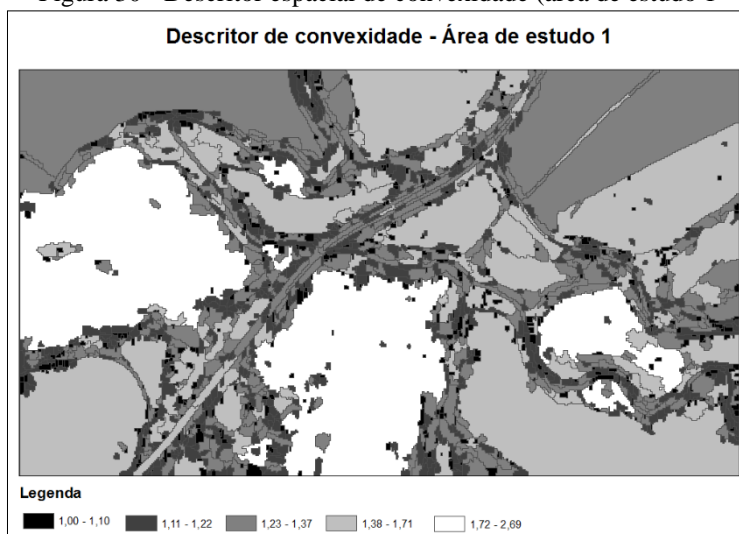
Figura 29 – Descritor espacial de direção principal - área de estudo 1



Fonte: a autora.

O descritor espacial “convexidade” caracteriza os cultivos e corpos hídricos, pois os cultivos aparecerem sempre com valores altos e os corpos hídricos sempre com valor baixo, branco e preto, respectivamente, conforme ilustra a Figura 30.

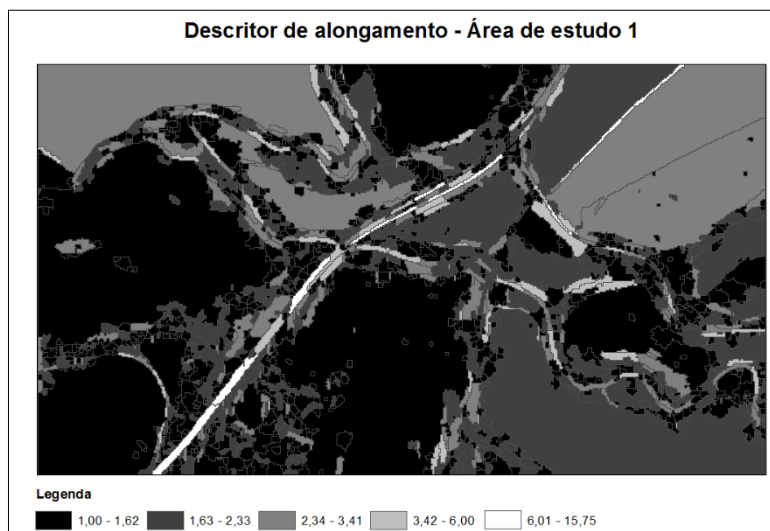
Figura 30 - Descritor espacial de convexidade (área de estudo 1



Fonte: a autora.

O descritor espacial de alongamento (Figura 31) caracteriza áreas de estradas e corpos hídricos que aparecem com os valores intermediários a altos. Destaca-se aqui as áreas de cultivo, pois estas são convexas e por isso tem baixo valor de alongamento.

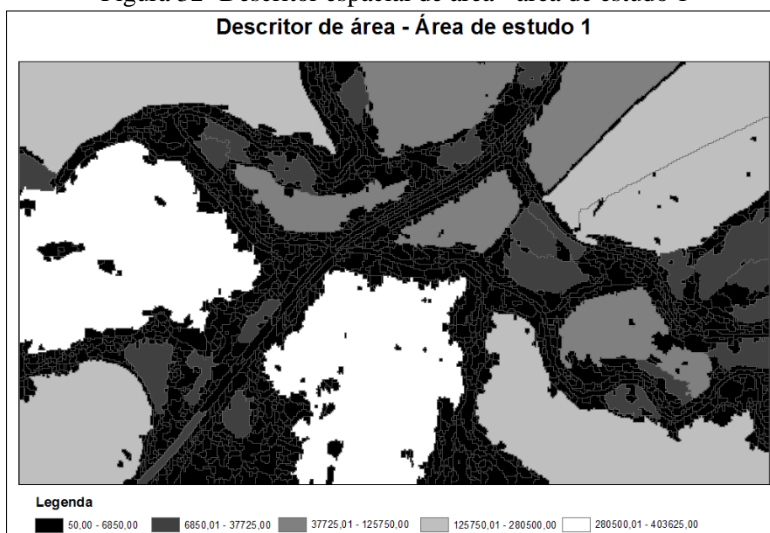
Figura 31 – Descritor espacial de alongamento – Área de estudo 1



Fonte: a autora.

O descritor espacial “área” caracteriza bem os cultivos, estes aparecem com os valores mais altos, representados na Figura 32 pelas cores mais claras (tons de cinza a branco). Os corpos hídricos, estradas, e cobertura florestal aparecem nas cores mais escuras, pois em sua maioria são polígonos com pequenas áreas.

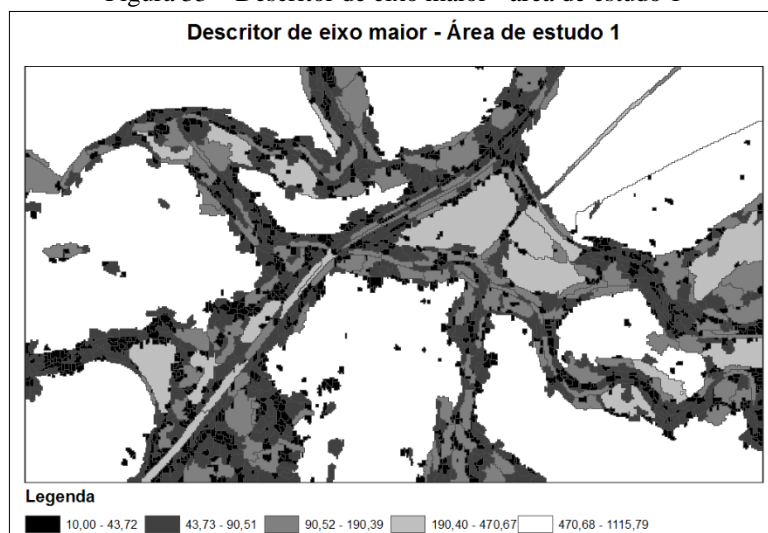
Figura 32- Descritor espacial de área - área de estudo 1



Fonte: a autora.

O descritor espacial “eixo maior” também caracteriza os cultivos, pois estes aparecem quase todos com valores altos, ilustrados em tons escuros na Figura 33.

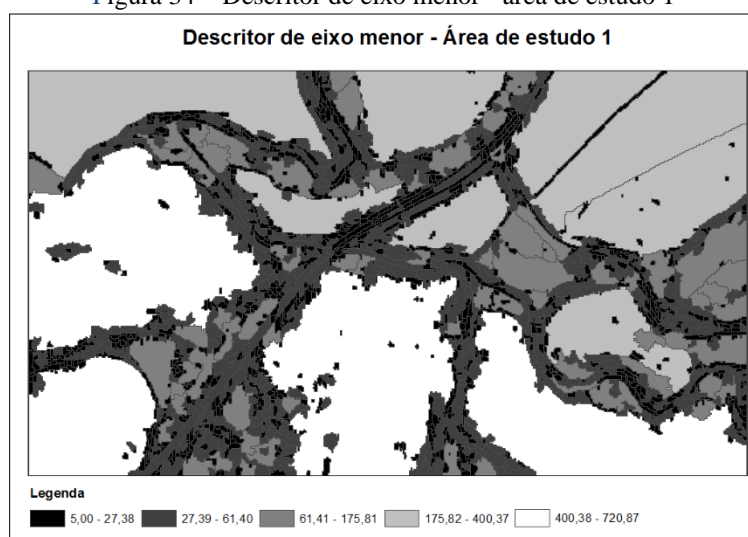
Figura 33 – Descritor de eixo maior - área de estudo 1



Fonte: a autora.

Para o descritor “eixo menor” os maiores valores são das áreas de cultivo, e os menores valores para estradas e corpos hídricos, caracterizando estes usos, conforme mostra a Figura 34.

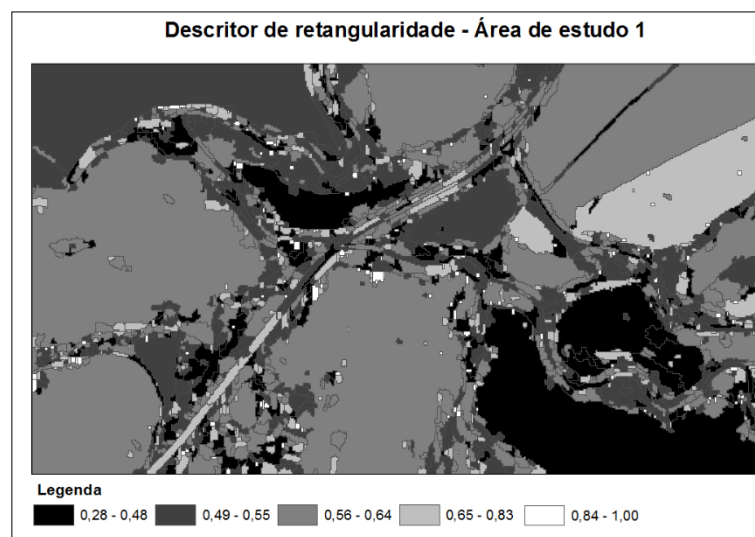
Figura 34 – Descritor de eixo menor - área de estudo 1



Fonte: a autora.

O descritor espacial de “retangularidade” (Figura 35) caracteriza as estradas, que em boa parte possuem valores altos nesse descritor, e também áreas de campos que nas classificações não tem boa separação com estradas e construções essas aparecem como pequenos retângulos brancos.

Figura 35 – Descritor de retangularidade - área de estudo 1



Fonte: a autora.

Nota-se que os descritores selecionados mediante a ACP, sejam eles espectrais, de textura, ou espaciais possuem relação a caracterização de classes da área de estudo 1 o que justifica a seleção desses descritores para a classificação dessa área de estudo.

4.3.1.2 Área de estudo 2

O banco de dados relacional das amostras de treinamento da área de estudo 2 compõe-se de 54 descritores e 400 amostras de objetos geográficos. Do mesmo modo que na área de estudo 1, a ACP possibilitou a geração de 54 componentes principais.

A Tabela 3 traz os dados de autovalor, variância explicada e variância acumulada das 54 componentes principais geradas para a área de estudo 2. Analisando a tabela é possível notar que as primeiras componentes principais possuem os maiores valores, e as últimas componentes principais têm os menores valores, o que é característico da ACP.

Tabela 3 - Autovalor e variância explicada e acumulada das componentes principais da área de estudo 2

(Continua)

CP	Autovalor	Variância explicada	Variância acumulada
CP1	30.913,000	20,390%	20,390%
CP2	21.495,000	9,860%	30,250%
CP3	190.773,000	7,766%	38,018%
CP4	178.401,000	6,792%	44,810%
CP5	150.854,000	4,856%	49,666%
CP6	133.844,000	3,823%	53,489%
CP7	129.556,000	3,582%	57,070%
CP8	125.197,000	3,345%	60,415%
CP9	122.213,000	3,187%	63,603%
CP10	115.217,000	2,833%	66,435%
CP11	111.373,000	2,647%	69,082%
CP12	108.871,000	2,529%	71,612%
CP13	100.451,000	2,153%	73,765%
CP14	0,974	2,022%	75,787%
CP15	0,959	1,963%	77,750%
CP16	0,948	1,918%	79,668%
CP17	0,915	1,788%	81,456%
CP18	0,885	1,673%	83,129%
CP19	0,822	1,441%	84,570%
CP20	0,790	1,333%	85,903%
CP21	0,782	1,304%	87,206%
CP22	0,753	1,209%	88,416%
CP23	0,743	1,177%	89,593%
CP24	0,728	1,130%	90,720%
CP25	0,702	1,053%	91,776%
CP26	0,686	1,004%	92,780%
CP27	0,671	0,960%	93,740%
CP28	0,592	0,748%	94,488%
CP29	0,576	0,708%	95,195%

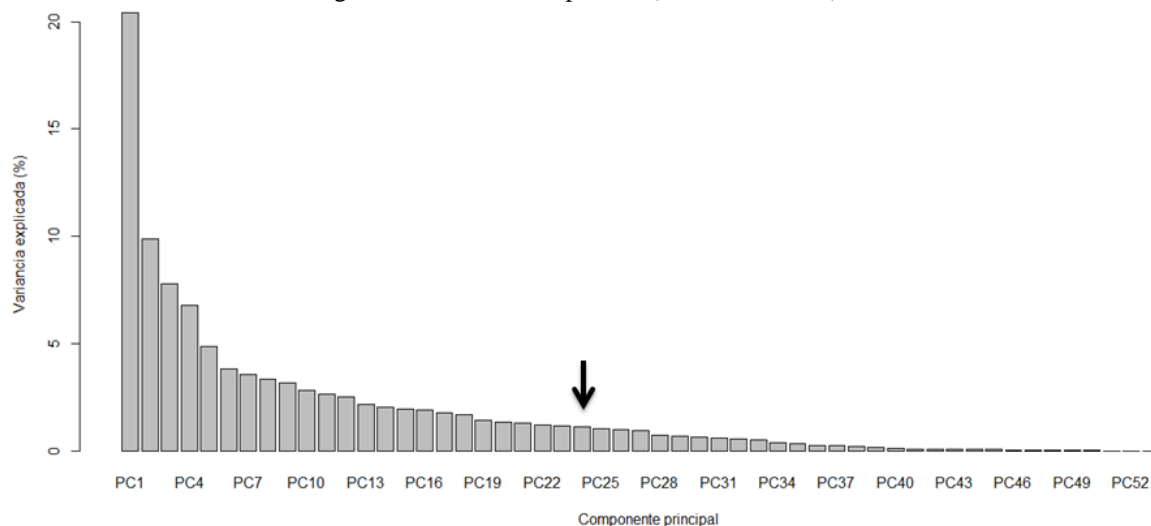
(Conclusão)			
CP	Autovalor	Variância explicada	Variância acumulada
CP30	0,563	0,677%	95,872%
CP31	0,536	0,613%	96,485%
CP32	0,514	0,563%	97,049%
CP33	0,485	0,501%	97,550%
CP34	0,439	0,412%	97,962%
CP35	0,394	0,330%	98,290%
CP36	0,352	0,265%	98,557%
CP37	0,338	0,244%	98,801%
CP38	0,323	0,223%	99,025%
CP39	0,288	0,177%	99,202%
CP40	0,232	0,115%	99,317%
CP41	0,218	0,101%	99,418%
CP42	0,209	0,094%	99,512%
CP43	0,202	0,087%	99,599%
CP44	0,195	0,081%	99,680%
CP45	0,178	0,068%	99,747%
CP46	0,168	0,060%	99,810%
CP47	0,156	0,052%	99,860%
CP48	0,150	0,048%	99,908%
CP49	0,135	0,039%	99,947%
CP50	0,117	0,029%	99,976%
CP51	0,085	0,015%	99,991%
CP52	0,052	0,006%	99,997%
CP53	0,036	0,003%	100,000%
CP54	0,000000000001	0,000%	100,000%

Fonte: a autora.

A aplicação do método de Beale no banco de dados das amostras de treinamento da área de estudo 2 selecionou as primeiras 25 componentes principais da Tabela 3, pois estas possuem autovalores iguais ou superiores a 0,7, e juntas somam mais de 91% da variância acumulada, restando as 29 CPs restantes excluídas menos de 9% de variância explicada.

A Figura 36 mostra o gráfico de variância explicada pelas CPs da área de estudo 2, observa-se que as primeiras CPs possuem os maiores valores e que estes vão decrescendo. A seta na figura indica que selecionaram-se até a 25ª CP. Percebe-se que as CPs não selecionadas possuem valores de variância explicada muito próximos de zero.

Figura 36- Variância explicada (área de estudo 2)



Fonte: a autora.

A Tabela 4 traz para cada CP gerada o maior valor absoluto de autovetor e o descritor correspondente, sendo este o critério do método de B4 para seleção de descritores.

Tabela 4 – Descritores de maior autovetor absoluto correspondentes as componentes principais geradas

CP	Autovetor	Descritor	Tipo
1	0,227	Textura média - Banda 2	Textura
2	0,398	Convexidade	Espacial
3	0,363	Intervalo de textura - Banda 3	Textura
4	0,401	Textura média - Banda 5	Textura
5	0,526	Circuncidade	Espacial
6	0,352	Intervalo de textura - Banda 4	Textura
7	0,393	Intervalo de textura - Banda 5	Textura
8	0,371	Variância de textura - Banda 4	Textura
9	5.580.877.847,113	Média dos números digitais - Banda 2	Espectral
10	0,414	Variância de textura - Banda 3	Textura
11	0,373	Variância de textura - Banda 5	Textura
12	0,436	Desvio padrão dos números digitais - Banda 5	Espectral
13	0,397	Eixo maior	Espacial
14	0,322	Desvio padrão dos números digitais - Banda 1	Espectral
15	0,270	Variância de textura - Banda 1	Textura
16	0,475	Direção principal	Espacial
17	0,549	Entropia de textura - Banda 1	Textura
18	0,472	Variância de textura - Banda 2	Textura
19	0,466	Desvio padrão dos números digitais - Banda 4	Espectral
20	0,342	Desvio padrão dos números digitais - Banda 2	Espectral
21	0,403	Entropia de textura - Banda 3	Textura
22	0,422	Entropia de textura - Banda 2	Textura
23	0,302	Entropia de textura - Banda 4	Textura

(continua)

			(Conclusão)
CP	Autovetor	Descritor	Tipo
24	0,265	Área	Espacial
25	0,452	Entropia de textura - Banda 5	Textura
26	0,380	Eixo menor	Espacial
27	0,177	Intervalo de textura - Banda 1	Textura
28	0,275	Valor máximo de pixel - Banda 2	Espectral
29	0,406	Valor mínimo de pixel - Banda 5	Espectral
30	0,355	Comprimento	Espacial
31	0,199	Intervalo de textura - Banda 2	Textura
32	0,205	Retangularidade	Espacial
33	0,436	Valor máximo de pixel - Banda 3	Espectral
34	0,283	Valor máximo de pixel - Banda 1	Espectral
35	0,369	Valor mínimo de pixel - Banda 3	Espectral
36	0,602	Solidez	Espacial
37	0,439	Média dos números digitais - Banda 3	Espectral
38	0,238	Relação área / furos	Espacial
39	0,398	Média dos números digitais - Banda 1	Espectral
40	0,311	Valor mínimo de pixel - Banda 1	Espectral
41	0,585	Valor máximo de pixel - Banda 4	Espectral
42	0,671	Textura média - Banda 1	Textura
43	0,607	Média dos números digitais - Banda 5	Espectral
44	0,314	Valor mínimo de pixel - Banda 2	Espectral
45	0,377	Compacidade	Espacial
46	0,235	Alongamento	Espacial
47	0,188	Fator de forma	Espacial
48	0,554	Valor mínimo de pixel - Banda 4	Espectral
49	0,262	Número de furos	Espacial
50	0,212	Textura média - Banda 3	Textura
51	0,029	Textura média - Banda 4	Textura
52	0,021	Média dos números digitais - Banda 4	Espectral
53	875.591.118,423	Desvio padrão dos números digitais - Banda 3	Espectral
54	0,136	Valor máximo de pixel - Banda 5	Espectral

Fonte: a autora.

Portanto mediante o método de Beale foram selecionadas as 25 primeiras componentes principais das Tabelas 3 e 4, e a partir do método B4 selecionou-se para cada uma das 25 CPs um descritor correspondente (Tabela 4).

Desses 25 descritores selecionados 5 são espectrais, 15 são de textura e 5 são espaciais.

Dos descritores espectrais selecionados o que se destacou foi o descritor de desvio padrão, sendo que foi selecionado para todas as bandas espectrais. A seleção

desses descritores pode ser justificada devido as médias de desvio padrão serem quase equivalentes nas 5 bandas espectrais.

A maioria dos descritores selecionados são de textura. A área de estudo 2 é composta em sua maior área por classes de reflorestamentos, cobertura florestal, e diferentes cultivos agrícolas. Para o ser humano distinguir visualmente essas classes utiliza aspectos de textura (textura lisa, rugosa, etc) de modo intrínseco. O fato de as técnicas selecionarem a maioria dos descritores de textura para essa área pode ser justamente por os descritores da GEOBIA tentarem imitar o processo de cognição humano em que o reconhecimento dessas classes se dá por meio de aspectos de textura.

O Quadro 5 mostra os descritores de textura selecionados:

Quadro 4 – Descritores de textura selecionados

Bandas espectrais	Descritores de textura selecionados			
	Variância de textura	Intervalo de textura	Entropia de textura	Textura média
Banda 1	X	-	X	-
Banda 2	X	-	X	X
Banda 3	X	X	X	-
Banda 4	X	X	X	-
Banda 5	X	X	X	X

Fonte: a autora.

Conforme o Quadro 5, dos descritores de textura selecionados: 5 são do comprimento de onda do infravermelho próximo, 3 do comprimento de onda do verde, e 3 do comprimento de onda do vermelho. Estas são bandas em que a reflectância da vegetação é alta, ocorrendo na área devido ao predomínio de vegetação e possivelmente por isso foram selecionados os descritores.

Os descritores espaciais selecionados para a área de estudo 2 foram: convexidade, circuncidade, eixo maior, direção principal, e área.

O descritor espacial convexidade mostra quais são os polígonos mais côncavos, sendo que quanto mais próximos ao valor 1,0, mais côncavos são os objetos geográficos, conforme mostra a Figura 37A. Esse descritor é característico da cobertura florestal que tem a forma côncava, e também das estradas e do cultivo 2.

Quanto ao descritor espacial circuncidade (Figura 37B) e direção principal (Figura 37C) não se nota relação direta entre estes descritores e as classes de uso e ocupação da terra, no entanto ambos foram selecionados a partir do método B4.

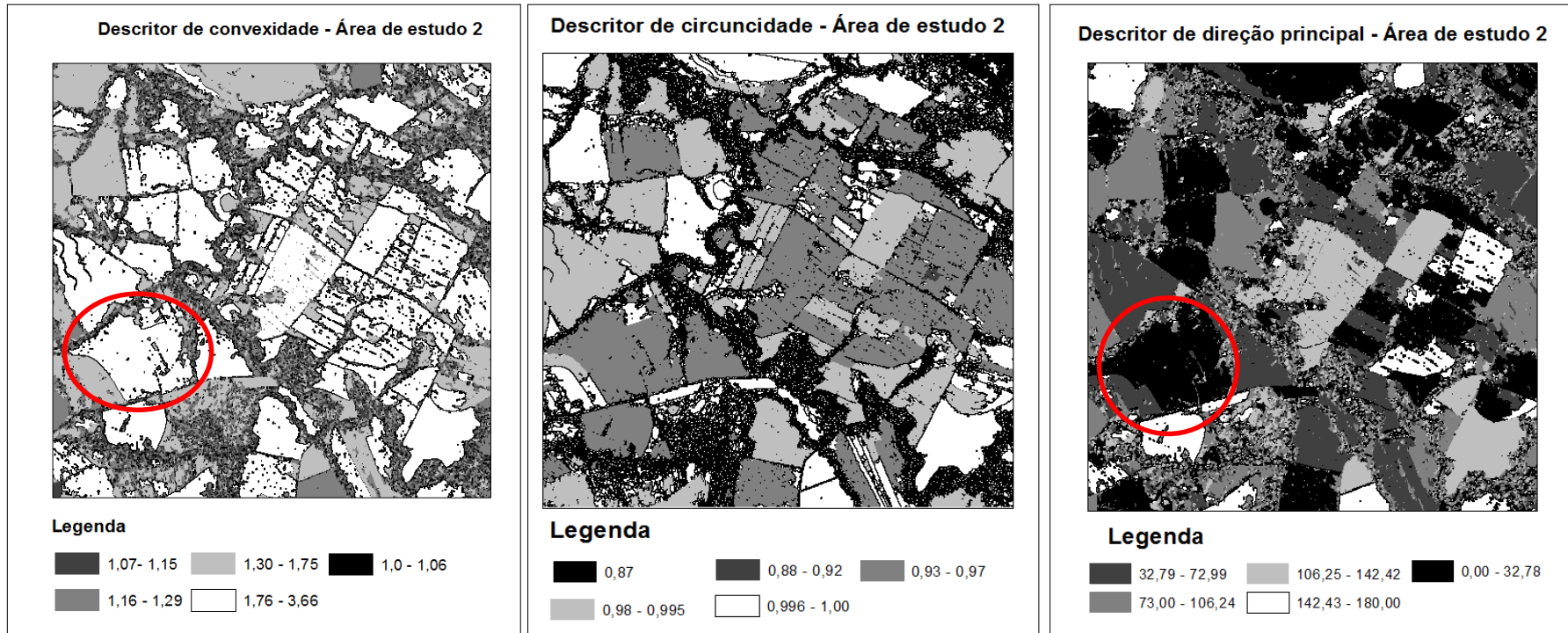
Nota-se que os descritores convexidade e direção principal se contrapõem conforme destacado na Figura 37A e 37C.

Figura 37 – Descritores espaciais: 37A - Descritor de convexidade; 37B – Descritor de circuncidade; 37C – Descritor de direção principal

37A

37B

37C

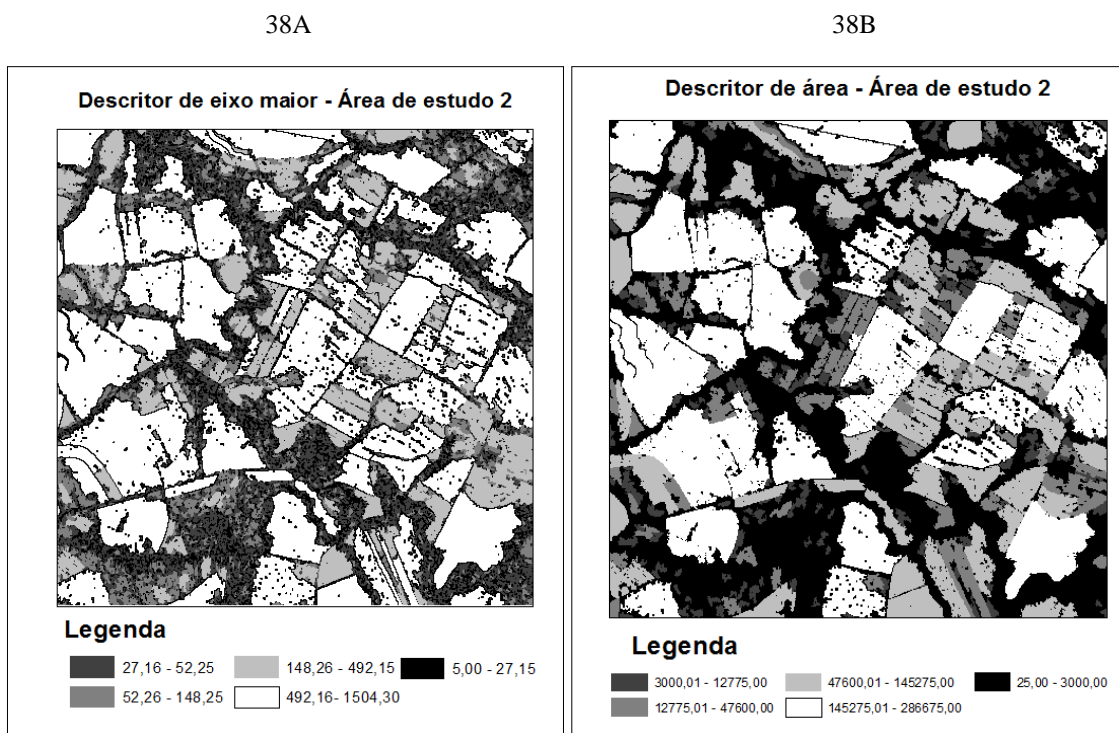


Fonte: a autora.

Conforme nota-se na Figura 38A, o descritor “eixo maior” caracteriza as áreas de corpos hídricos, cobertura florestal, estradas, construções e cultivo 2, devido aos menores valores destacados em cores escuras. Já os maiores valores destacados pelos tons claros não caracterizam classes específicas.

O descritor espacial de área é representado na Figura 38B, onde observa-se as nuances escuras para os valores baixos e claras para os maiores valores de área. Nota-se que os corpos hídricos e cobertura florestal são caracterizados nesse descritor por possuírem áreas pequenas, de baixo valor. Em contrapartida as áreas de cultivos em sua grande maioria têm valores intermediários a altos.

Figura 38 - Descritor espacial de convexidade



Fonte: a autora.

Observou-se isso para as duas áreas de estudo de acordo com os resultados apresentados que alguns descritores possuem relação direta com classes de uso e cobertura da terra, e alguns seleciona-se devido aos valores no banco de dados relacional serem ou muito divergentes ou muito parecidos,

O próximo tópico traz os resultados da seleção de descritores mediante a Análise de Agrupamentos que tem abordagem diferente que a Análise de Componentes Principais.

4.3.2 Seleção de descritores mediante a Análise de Agrupamentos (AA)

A Análise de Agrupamentos como mencionado nos procedimentos metodológicos tem por objetivo agrupar variáveis com características comuns em grupos.

A Análise de Componentes Principais foi realizada entendendo como variáveis os descritores do banco de dados relacional, de diferente forma na Análise de Agrupamentos analisou-se como variáveis os objetos geográficos observando os dendrogramas gerados separadamente para cada descritor ou determinado grupo de descritores.

A AA realizada dessa maneira possibilita que se reconheça em quais descritores há formação de grupos heterogêneos de objetos geográficos das classes determinadas para a classificação, auxiliando dessa forma na escolha de descritores para a classificação. Nesse trabalho optou-se por analisar os dendrogramas da AA a partir da visualização de grupos que se formavam com objetos de mesma classe.

Realizou-se a AA nos descritores que não foram selecionados a partir da ACP, tanto para a área de estudo 1 como para a área de estudo 2.

Segue a seguir os resultados da AA nas subseções seguintes.

4.3.2.1 Área de estudo 1

Para a área de estudo 1 a partir da Análise de Agrupamentos foram selecionados 19 descritores para a classificação. Destes 11 são espectrais, 1 espacial, e 7 de textura, conforme mostra o Quadro 6:

Quadro 5 – Descritores selecionados mediante Análise de Agrupamentos para a área de estudo 1

Tipo	Descritores
Espectral	Média dos números digitais na banda 1
Espectral	Valor mínimo (bandas 1,2,3,4,5)
Espectral	Valor máximo (banda 1,2,3,4,5)
Espacial	Fator forma
Textura	Entropia (bandas 1,3,5)
Textura	Intervalo (bandas 1,2,3)
Textura	Variância (banda 2)

Fonte: a autora.

Nas Figuras 39, 40, 41, 42, 43, 44 e 45 são representados os dendrogramas gerados dos descritores selecionados, destacando exemplos de agrupamentos que motivaram a seleção dos descritores para a classificação. Ressalta-se que optou-se por analisar os dendrogramas a partir de grupos que se formavam com objetos de mesma classe e não por uma determinada distância euclidiana.

É possível observar na Figura 39 o dendrograma do descritor “média dos números digitais na banda 1”, onde há destacados alguns agrupamentos das classes estradas/construções e corpos hídricos. Para a classe estradas/construções os agrupamentos ocorrem a distância euclidiana de 0,001, e para os corpos hídricos a 0,0005, distâncias pequenas que comprovam que os objetos geográficos em cada uma dessas classes são semelhantes.

No dendrograma gerado com os descritores de “valor mínimo das bandas 1, 2, 3, 4 e 5” (Figura 40), formaram-se agrupamentos das classes: cobertura florestal, corpos hídricos, solo exposto, cultivo 2, cultivo 3 e cultivo 5, e por isso foi selecionado para a classificação.

O dendrograma formado pelos descritores de “valor máximo das bandas 1, 2, 3, 4, e 5” (Figura 41) são encontrados agrupamentos para a classe cultivo 2 com distâncias euclidianas de 0,012 e 0,030 o que mostra que há grupos de objetos geográficos que são mais similares que outros nessa classe. Para o cultivo 3 os agrupamentos se formam a distância de 0,05, no entanto há apenas um agrupamento com elementos somente dessa classe. Para o cultivo 5 o dendrograma forma grupos a distâncias aproximadas de 0,015 e 0,018. Para a cobertura florestal há grupos com vários elementos a distância de 0,015, no entanto há grupos menores com distâncias menores dentro de agrupamentos maiores dessa classe, o mesmo ocorre para a classe de corpos hídricos a distância de 0,010. A classe estradas/construções formam agrupamentos com vários subgrupos objetos a uma distância aproximada de 0,030.

Com relação aos descritores “intervalo de textura nas bandas do azul, verde e vermelho” representados na Figura 42, formou-se grupos das classes estradas/construções, corpos hídricos, cobertura florestal e também de cultivo 2. Respectivamente a distâncias euclidianas de 0,2; 0,4; 0,4; e 0,4.

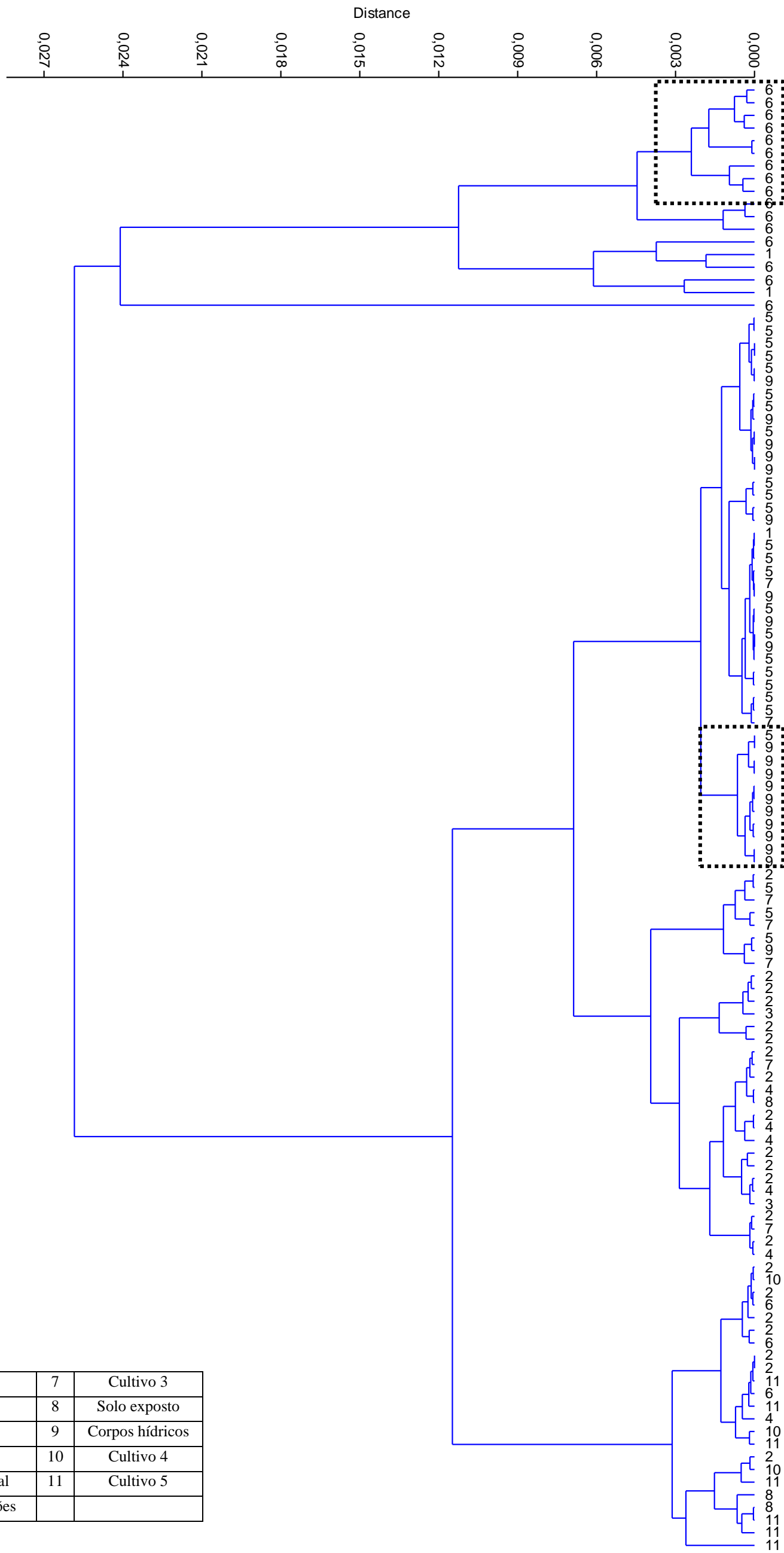
No dendrograma do descritor de “variância de textura” ilustrado na Figura 43, notam-se agrupamentos da classe “corpos hídricos”, em sua grande maioria com pequena distância euclidiana: menor que 0,08.

No dendrograma gerado a partir dos descritores de entropia de textura das bandas 1, 3 e 5 representado na Figura 44, são encontrados agrupamentos da classe cultivo 2 (aveia) a distância de 0,011, e do cultivo 3 (soja) a distância euclidiana de 0,015. Nota-se que essas distâncias são baixas, portanto as amostras são muito similares. Também há agrupamentos no dendrograma desse descritor relativo as classes de corpos hídricos e estradas e construções, demonstrando que os objetos dessas classes possuem boas separabilidade.

O descritor “fator forma” representado na Figura 45, não possui muitos agrupamentos heterogêneos. Esse descritor foi selecionado para a classificação por formar um agrupamento do cultivo 7, que outros descritores não formaram e dessa forma proporcionar maior separabilidade dessa classe na classificação.

Os descritores selecionados mediante a AA como veremos adiante na seção de classificação proporciona melhora a acuracidade da classificação, sendo pertinente as análises no dendrograma dos descritores gerados.

Figura 39 - Dendrograma do descritor espectral “média dos números digitais na banda 1”

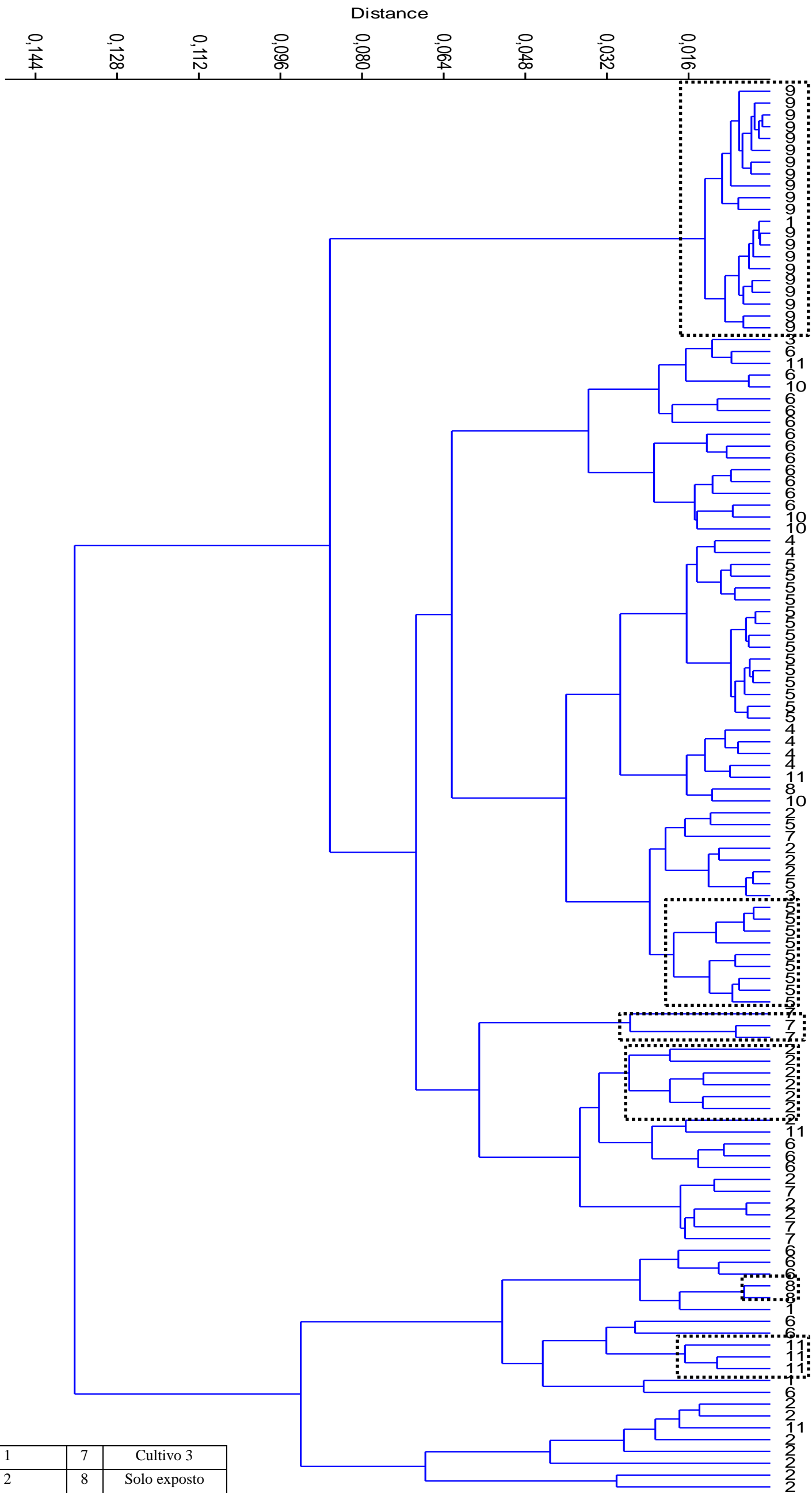


Legenda:

1	Cultivo 1	7	Cultivo 3
2	Cultivo 2	8	Solo exposto
3	Cultivo 7	9	Corpos hídricos
4	Cultivo 6	10	Cultivo 4
5	Cobertura florestal	11	Cultivo 5
6	Estradas/construções		

Fonte: a autora.

Figura 40 –Dendrograma dos descritores espectrais “valor mínimo de *pixel* nas bandas 1, 2, 3, 4, 5”

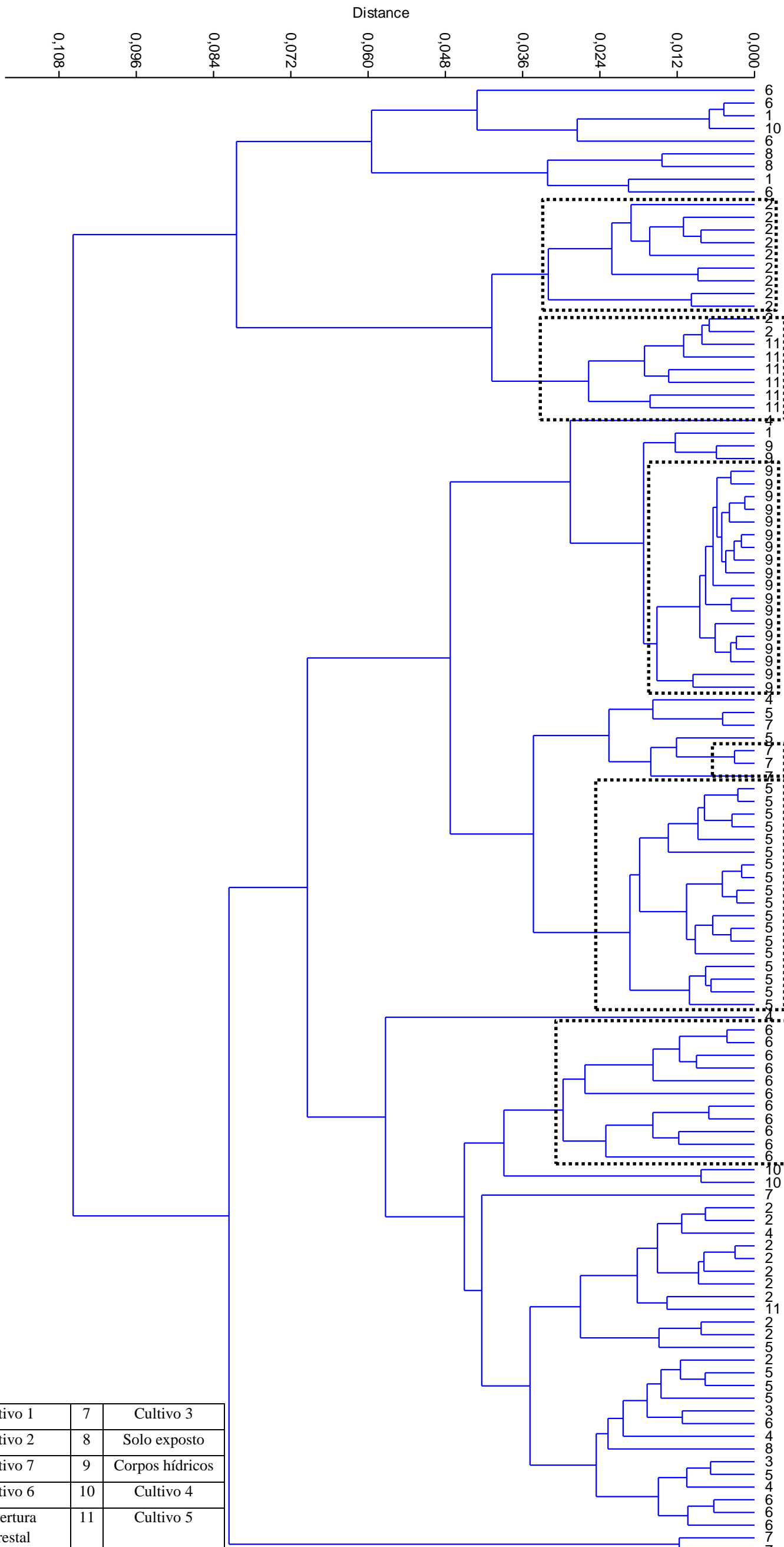


Legenda:

1	Cultivo 1	7	Cultivo 3
2	Cultivo 2	8	Solo exposto
3	Cultivo 7	9	Corpos hídricos
4	Cultivo 6	10	Cultivo 4
5	Cobertura florestal	11	Cultivo 5
6	Estradas/construções		

Fonte: a autora.

Figura 41 – Dendograma dos descritores “valor máximo de pixel nas bandas 1, 2, 3, 4, 5”

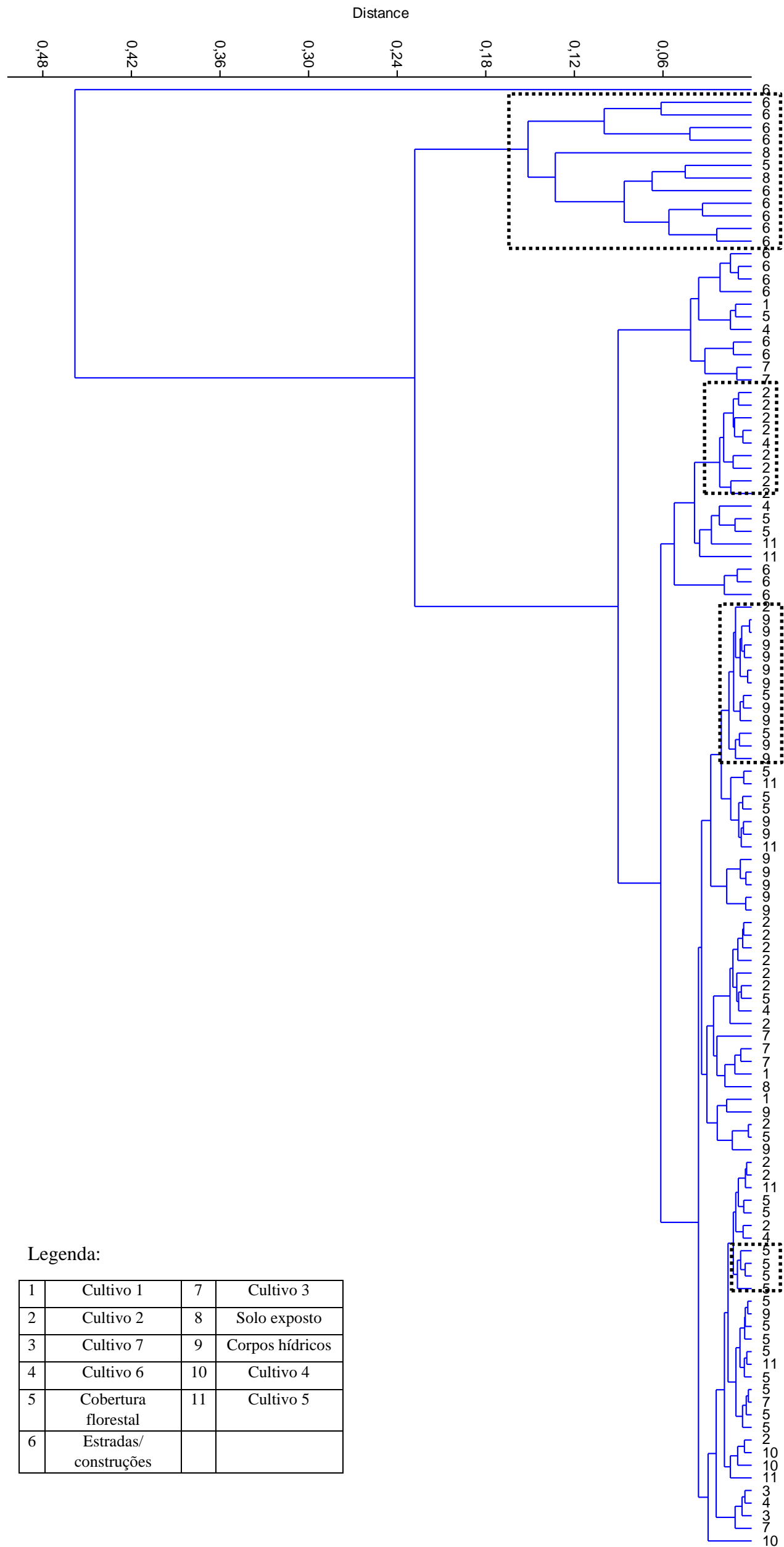


Legenda:

1	Cultivo 1	7	Cultivo 3
2	Cultivo 2	8	Solo exposto
3	Cultivo 7	9	Corpos hídricos
4	Cultivo 6	10	Cultivo 4
5	Cobertura florestal	11	Cultivo 5
6	Estradas/ construções		

Fonte: a autora.

Figura 42 – Dendrograma dos descritores de textura “intervalo de textura nas bandas 1,2 e 3”

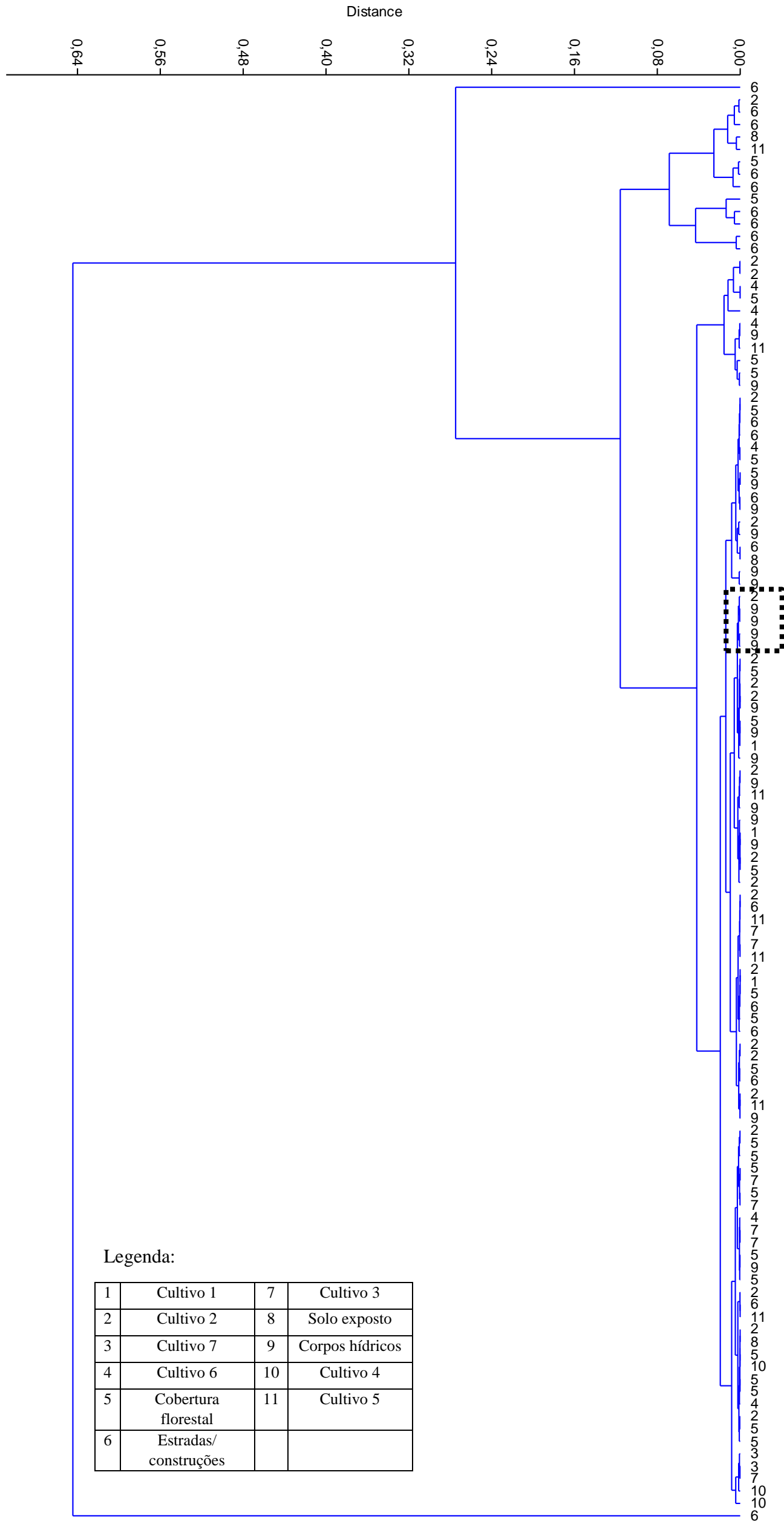


Legenda:

1	Cultivo 1	7	Cultivo 3
2	Cultivo 2	8	Solo exposto
3	Cultivo 7	9	Corpos hídricos
4	Cultivo 6	10	Cultivo 4
5	Cobertura florestal	11	Cultivo 5
6	Estradas/ construções		

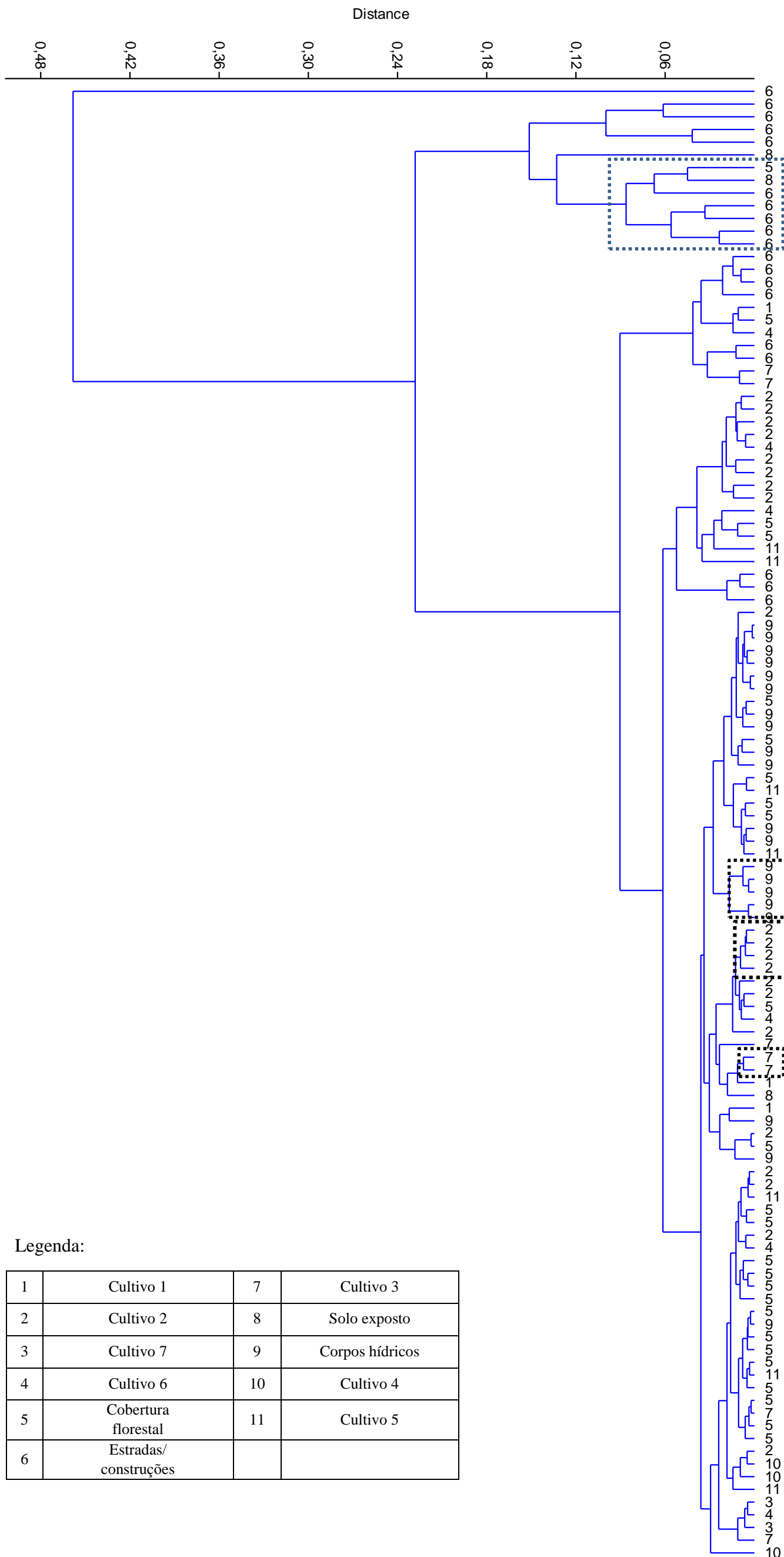
Fonte: a autora.

Figura 43– Dendrograma do descritores de “variância de textura na bandas 2”



Fonte: a autora.

Figura 44 – Dendrograma formado pelos descritores “entropia de textura nas bandas 1,3 e 5”

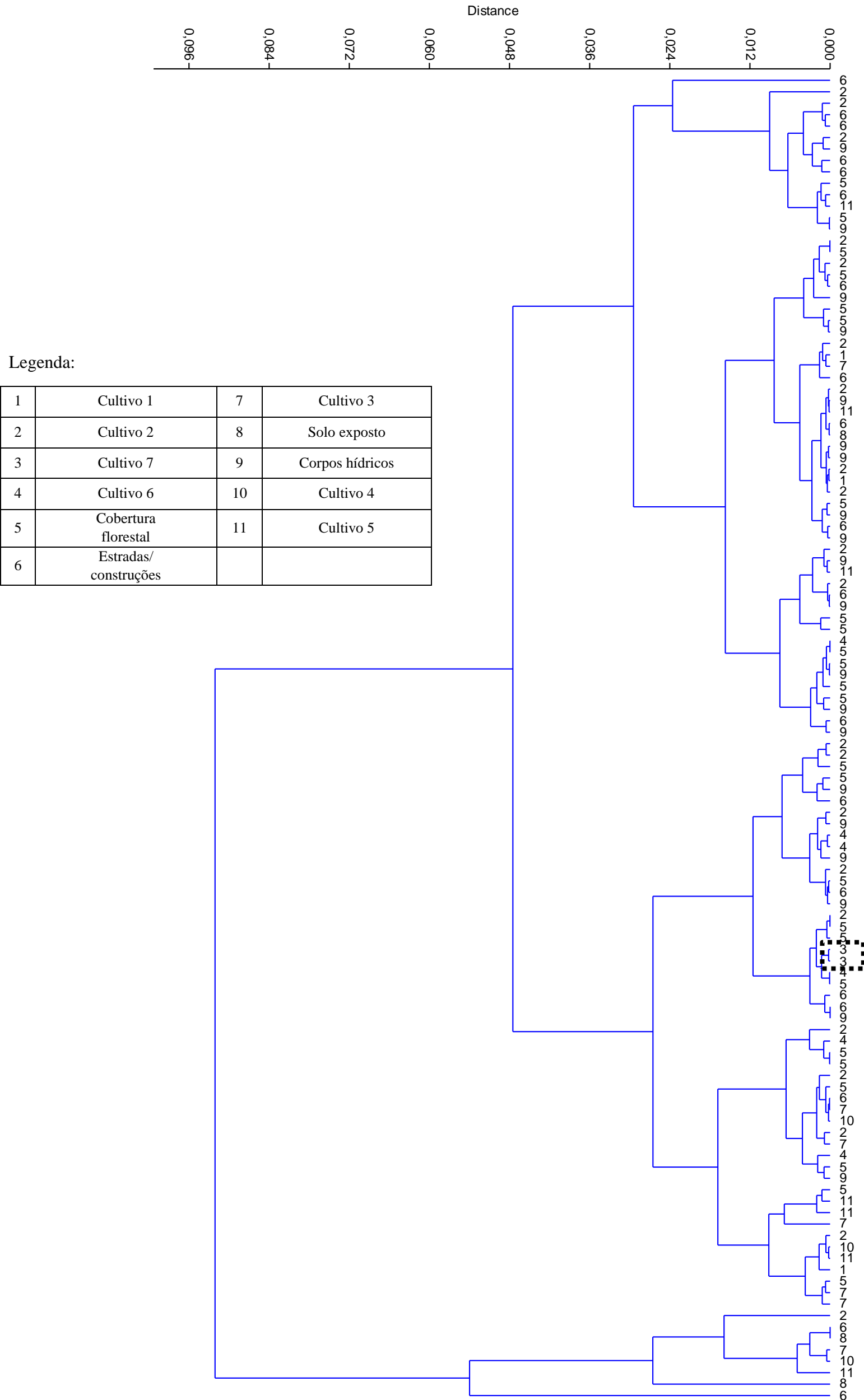


Legenda:

1	Cultivo 1	7	Cultivo 3
2	Cultivo 2	8	Solo exposto
3	Cultivo 7	9	Corpos hídricos
4	Cultivo 6	10	Cultivo 4
5	Cobertura florestal	11	Cultivo 5
6	Estradas/ construções		

Fonte: a autora.

Figura 45 – Dendrograma do descritor espacial “fator de forma”



Fonte: a autora.

4.3.2.2 Área de estudo 2

Para a área de estudo 2 mediante a Análise de Agrupamentos selecionaram-se 8 descritores para a classificação, 6 são espectrais, 1 espacial e 1 de textura conforme mostra o Quadro 7.

Quadro 6 – Descritores selecionados mediante Análise de Agrupamentos para a área de estudo 1

Tipo	Descritores
Espectral	Média dos números digitais das bandas 2 e 5
Espectral	Valor mínimo de pixel bandas 4 e 5
Espectral	Valor máximo de pixel bandas 1 e 4
Espacial	Comprimento
Textura	Textura média da banda 1

Fonte: a autora.

Nas Figuras 46, 47, 48, e 49 são representados os dendrogramas gerados dos descritores selecionados, contendo em “zoom” exemplos de agrupamentos que motivaram a seleção dos descritores para a classificação. Optou-se por analisar os dendrogramas a partir de grupos que se formavam com objetos de mesma classe e não por uma determinada distância euclidiana.

Na Figura 46 ilustra-se o dendrograma formado pelos descritores “média dos números digitais nas bandas do verde e infravermelho próximo”. Estes foram selecionados porque formou grupos heterogêneos que em sua composição continham maior número de objetos geográficos, o que contribui na classificação para a separação de classes.

A classe cultivo 1 forma grande agrupamento com vários objetos geográficos a distância euclidiana aproximada de 0,006, no entanto há grupos menores dentro desse a partir da distância de 0,001. O cultivo 2 possui dois grupos contendo todos os elementos da classe, o primeiro se agrupa a distância euclidiana de 0,009 e o segundo a 0,0012, no entanto o primeiro grupo se junta posteriormente a classe cultivo 4, e o segundo a classe cultivo 3. O cultivo 3 possui um agrupamento que a distância euclidiana de 0,008 engloba todos os seus 28 objetos geográficos.

A uma distância euclidiana de aproximadamente 0,012 todas as 29 amostras da classe cultivo 4 se agrupam. Ao passo que todas as amostras da classe construções se agrupam a distância euclidiana de 0,015.

A cobertura florestal possui grupos com distância euclidiana de 0,005 e 0,007. A classe corpos hídricos forma um agrupamento somente a distância de 0,012, no entanto contém todas as amostras da classe.

No dendrograma constituído pelos descritores valor mínimo de *pixel* das bandas do *red-edge* e infravermelho próximo (Figura 47), encontram-se grupos heterogêneos das classes cultivo 1, cultivo 2, cultivo 4, cobertura florestal e corpos hídricos. Isso devido à alta reflectância da vegetação nesses comprimentos de onda, e a baixa reflectância da água que faz com que os corpos hídricos se destaquem por conter valores mínimos.

As amostras da classe cultivo 1 se agrupam a uma distância euclidiana de 0,001 com poucos elementos. A classe cultivo 2 forma grupos com mais amostras a uma distância de 0,003; o mesmo ocorre com o cultivo 4 a distância euclidiana de 0,007. Os corpos hídricos formam grupos com várias amostras a uma distância de 0,005. O quadro seguinte mostra fragmentos do dendrograma que exemplificam agrupamentos para cada classe mencionada.

Figura 46 – Dendrograma do descritor espectral “Média dos números digitais das bandas 2 e 5”

Legenda:

1	Reflorestamento 1	7	Cultivo 2
2	Reflorestamento 2	8	Cultivo 3
3	Solo exposto	9	Cultivo 4
4	Cobertura florestal	10	Estradas
5	Corpos hídricos	11	Construções
6	Cultivo 1		

Zoom:

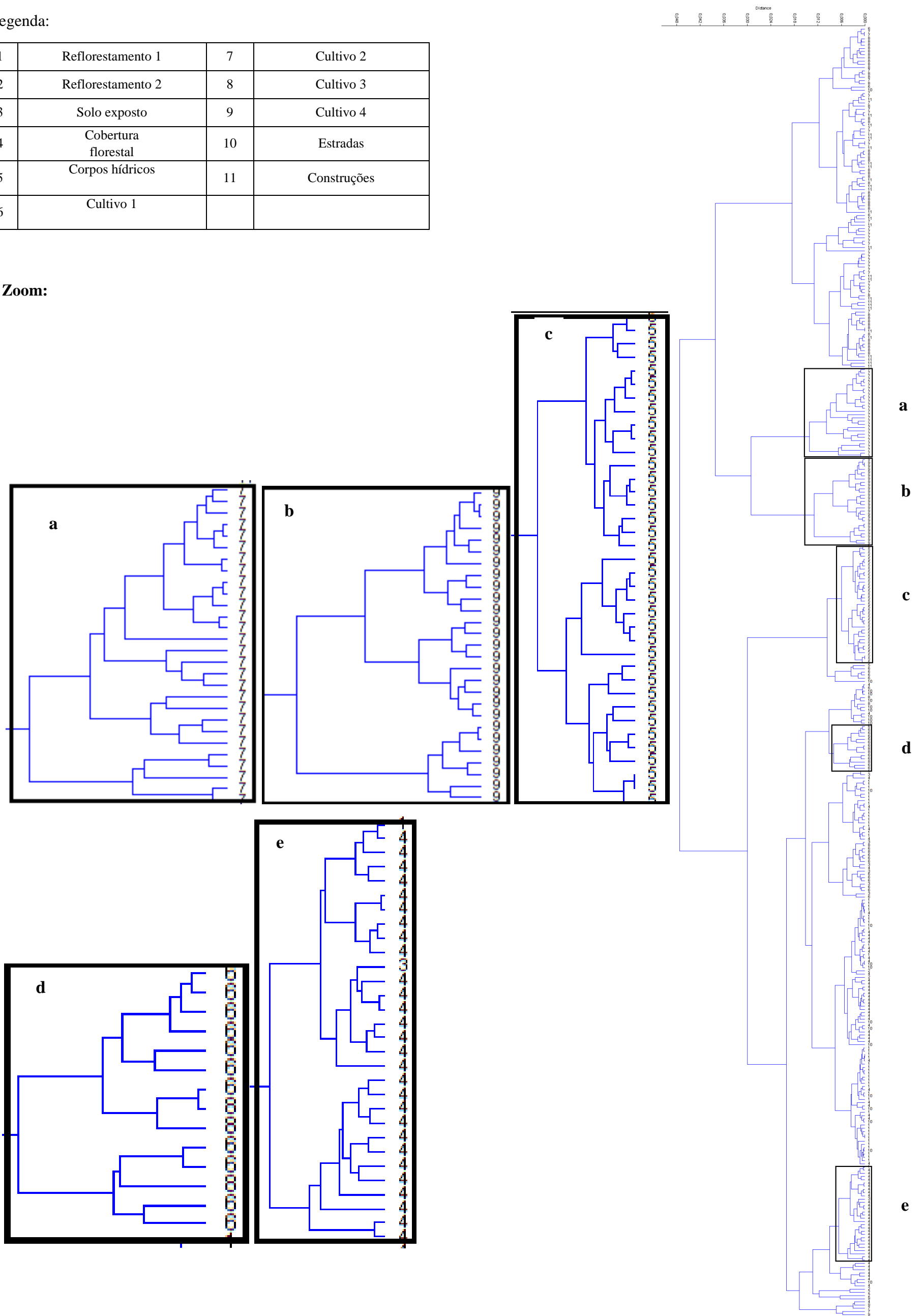


Figura 47 – Dendrograma dos descritores espectrais “valor mínimo de pixel nas bandas 4 e 5”

Legenda:

1	Reflorestamento 1	7	Cultivo 2
2	Reflorestamento 2	8	Cultivo 3
3	Solo exposto	9	Cultivo 4
4	Cobertura florestal	10	Estradas
5	Corpos hídricos	11	Construções
6	Cultivo 1		

Zoom:



No dendrograma formado pelos descritores “valor máximo de *pixel* nas bandas do azul e *red-edge*” (Figura 48), encontram-se agrupamentos das classes construções, cultivo 1, cultivo 2, cultivo 4, e cobertura florestal. Devido ao comprimento de onda do *red-edge* possuir alta reflectância para vegetação e as construções possuírem boa reflectância na banda do azul.

Nota-se que todas as amostras da classe construções se agrupam a distância euclidiana de 0,032, ou seja, nesses descritores essa classe possui boa separabilidade das outras.

O cultivo 1 possui um grupo a distância euclidiana aproximada de 0,006. A outro grupo a essa mesma distância que contém além das amostras dessa classe outras de reflorestamento 1.

O cultivo 2 possui um grupo com a maioria de suas amostras de treinamento a uma distância euclidiana aproximada de 0,005, o grupo deixa de ser homogêneo quando se junta a um agrupamento formado por amostras de cultivo 2 e uma única amostra de cultivo 1.

Há grupos com somente amostras da classe cultivo 3 a distância euclidiana de 0,003, posteriormente esses agrupamentos que contém amostras do cultivo 1.

As amostras do cultivo 4 estão todas agrupadas em somente um grupo a distância euclidiana de 0,002, o que demonstra que são similares e separáveis das outras classes.

A cobertura florestal possui vários pequenos grupos a distância euclidiana aproximada de 0,001, esses pequenos grupos deixam de ser homogêneos quando se juntam a agrupamentos que contém amostras de corpos hídricos, o que confirma que as duas classes não tem boa separabilidade.

Com relação ao dendrograma do descritor “comprimento” ilustrado na Figura 48, praticamente não há separação de classes entre os grupos, que se formam a distâncias euclidianas muito próximas de 0, como se pode notar no detalhamento a) da Figura 48, portanto nesse descritor não há separabilidade entre a maioria das classes. Formam-se agrupamentos com objetos das classes construções e cultivo como nota-se nos detalhamentos da Figura 49.

Figura 48 – Dendrograma formado pelos descritores “valor máximo de pixel nas bandas 1 e 4”

Legenda:

1	Reflorestamento 1	7	Cultivo 2
2	Reflorestamento 2	8	Cultivo 3
3	Solo exposto	9	Cultivo 4
4	Cobertura florestal	10	Estradas
5	Corpos hídricos	11	Construções
6	Cultivo 1		

Zoom:

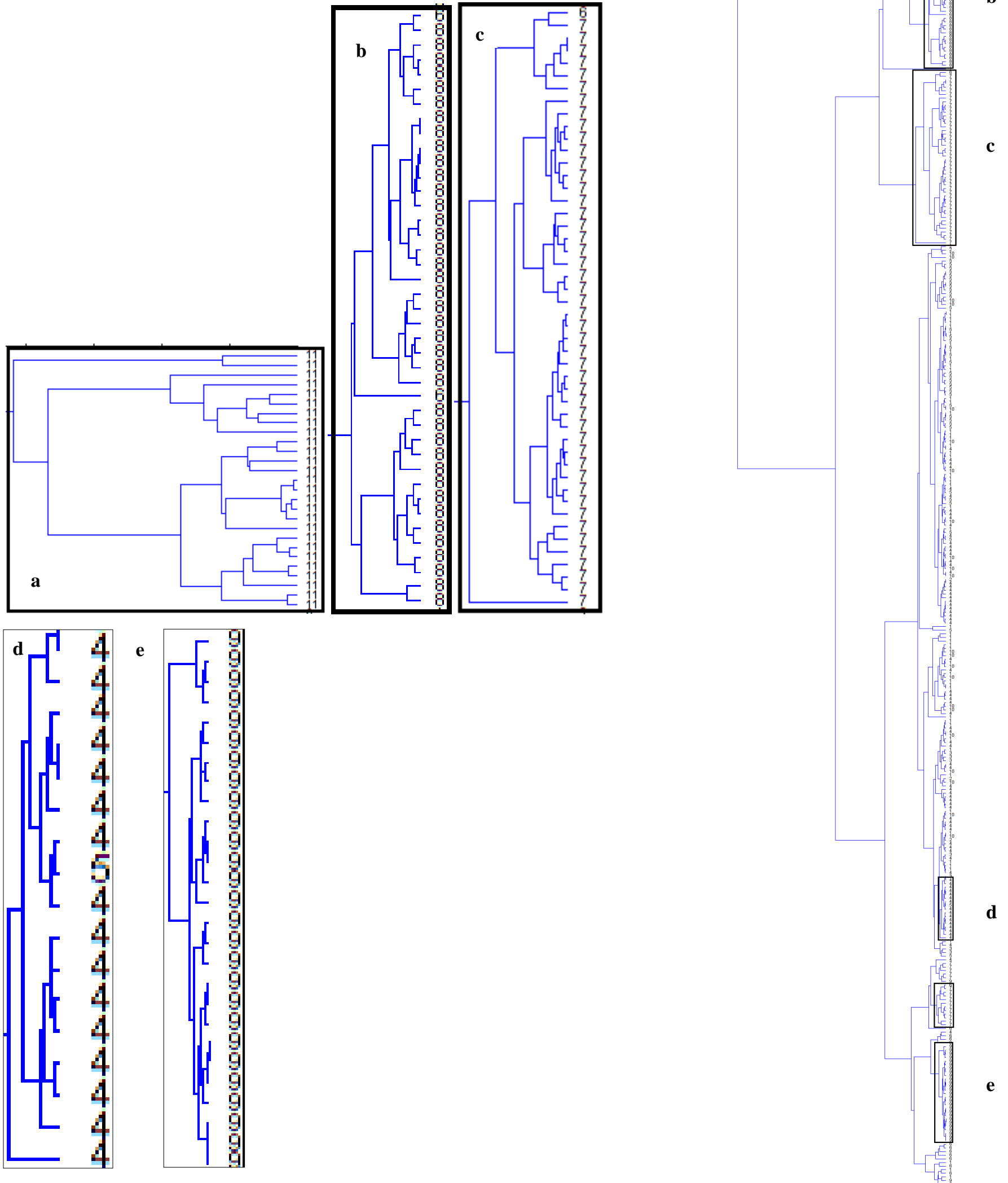
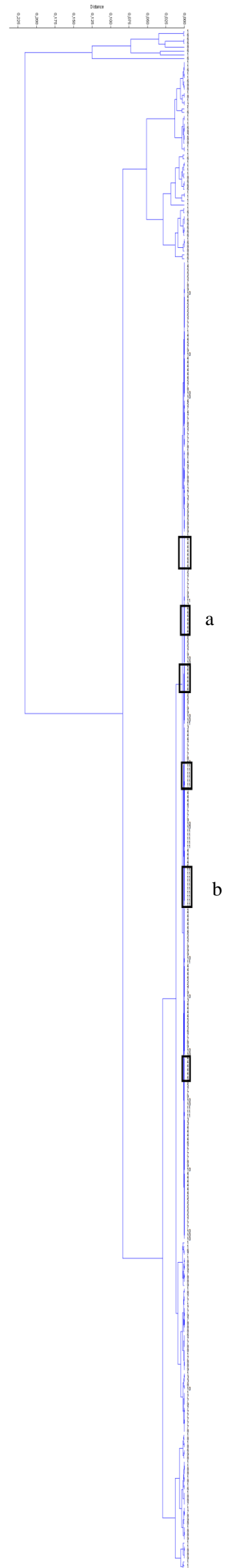
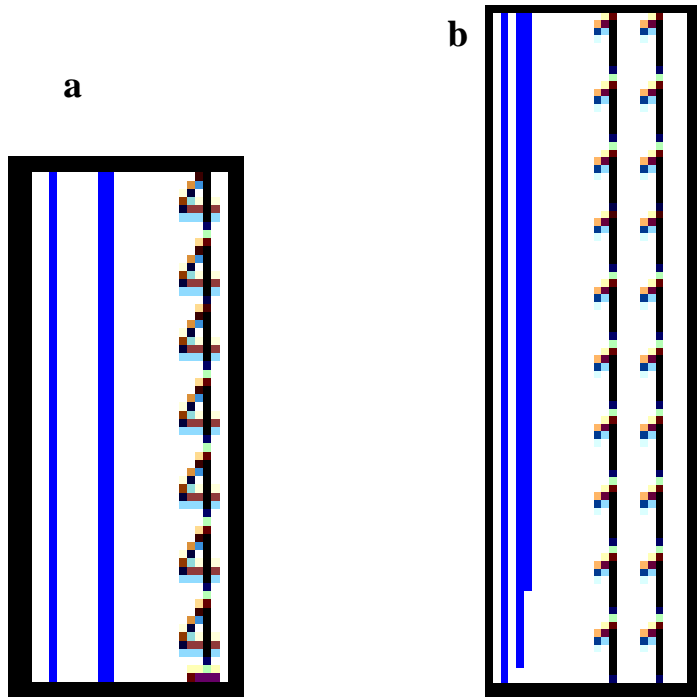


Figura 49 – Dendrograma formado pelos descritor espacial “comprimento”

Legenda:

1	Reflorestamento 1	7	Cultivo 2
2	Reflorestamento 2	8	Cultivo 3
3	Solo exposto	9	Cultivo 4
4	Cobertura florestal	10	Estradas
5	Corpos hídricos	11	Construções
6	Cultivo 1		

Zoom:



No próximo tópico são abordados os resultados para a classificação digital orientada a objetos.

4.4 Classificação e validação do uso e ocupação da terra

Para a classificação orientada a objetos geográficos há a possibilidade de selecionar os descritores a serem utilizados no processamento, o que pode ser uma vantagem, pois pode-se utilizar todos os descritores ou somente os que o técnico considerar mais adequado.

Com a finalidade de obter classificação ótima e foram realizados três testes de classificação para cada área de estudo, utilizando no primeiro descritores selecionados mediante ACP, no segundo ACP e AA, e no terceiro todos os descritores.

Para avaliar a acurácia dos três testes foi realizada a matriz de confusão, conforme citado na metodologia.

4.4.1 Classificação e validação do uso e ocupação da terra para a área de estudo 1

Para a classificação de uso e ocupação da terra para a área de estudo 1 considerou-se as 11 classes conforme mencionado na metodologia: cultivo 1, cultivo 2, cultivo 3, cultivo 4, cultivo 5, cultivo 6, cultivo 7, solo exposto, cobertura florestal e estradas/construções.

O Quadro 7 mostra quais descritores foram utilizados em cada teste de classificação para a área de estudo 1. No primeiro teste realizado utilizou-se somente os descritores selecionados a partir da ACP (segunda coluna do Quadro 7). No segundo teste além dos descritores selecionados mediante a ACP utilizou-se os selecionado a partir a AA. O terceiro teste foi realizado usando todos os descritores.

A Figura 50 mostra os testes de classificações geradas.

Quadro 7 – Testes e descritores para a classificação supervisionada orientada a objetos

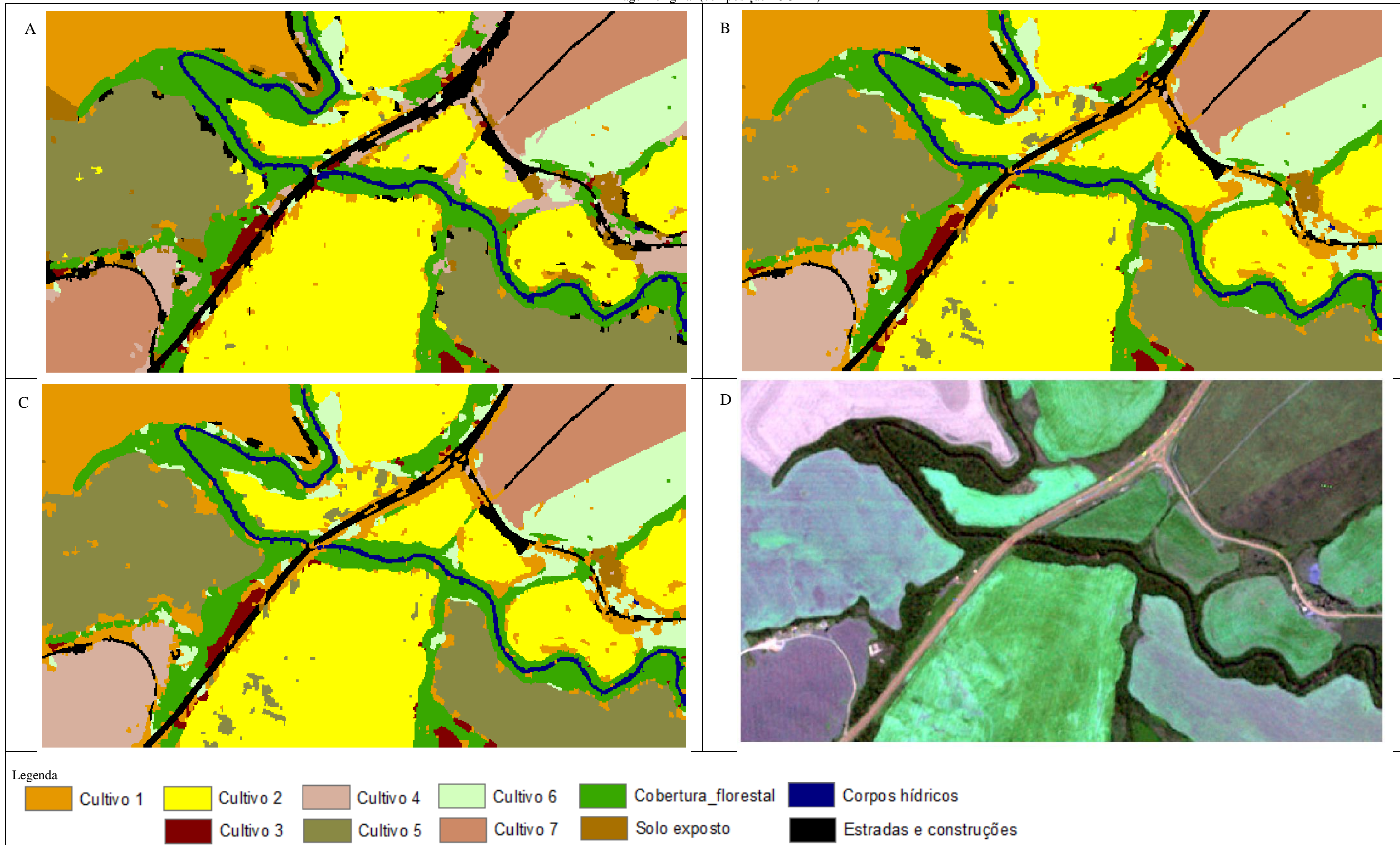
(Continua)

Descritores	ACP	ACP + AA	Todos
Desvio padrão dos números digitais - Banda 1			X
Desvio padrão dos números digitais - Banda 2			X
Desvio padrão dos números digitais - Banda 3	X	X	X
Desvio padrão dos números digitais - Banda 4			X
Desvio padrão dos números digitais - Banda 5	X	X	X
Média dos números digitais - Banda 1		X	X
Média dos números digitais - Banda 2	X	X	X
Média dos números digitais - Banda 3			X
Média dos números digitais - Banda 4			X
Média dos números digitais - Banda 5	X	X	X
Valor máximo de pixel - Banda 1		X	X
Valor máximo de pixel - Banda 2	X	X	X
Valor máximo de pixel - Banda 3		X	X
Valor máximo de pixel - Banda 4	X	X	X
Valor máximo de pixel - Banda 5		X	X
Valor mínimo de pixel - Banda 1		X	X
Valor mínimo de pixel - Banda 2		X	X
Valor mínimo de pixel - Banda 3		X	X
Valor mínimo de pixel - Banda 4		X	X
Valor mínimo de pixel - Banda 5	X	X	X
Área	X	X	X
Alongamento			X
Circuncidade	X	X	X
Compacidade			X
Comprimento			X
Convexidade	X	X	X
Direção principal	X	X	X
Eixo maior	X	X	X
Eixo menor	X	X	X
Fator de forma		X	X
Número de furos			X
Relação área/furos			X
Retangularidade	X	X	X
Solidez			X
Entropia de textura - Banda 1		X	X
Entropia de textura - Banda 2			X
Entropia de textura - Banda 3		X	X
Entropia de textura - Banda 4	X	X	X
Entropia de textura - Banda 5		X	X
Intervalo de textura - Banda 1		X	X
Intervalo de textura - Banda 2		X	X

(Conclusão)			
Descritores	ACP	ACP + AA	Todos
Intervalo de textura - Banda 3		X	X
Intervalo de textura - Banda 4	X	X	X
Intervalo de textura - Banda 5	X	X	X
Textura média - Banda 1			X
Textura média - Banda 2			X
Textura média - Banda 3			X
Textura média - Banda 4			X
Textura média - Banda 5			X
Variância de textura - Banda 1	X	X	X
Variância de textura - Banda 2		X	X
Variância de textura - Banda 3	X	X	X
Variância de textura - Banda 4	X	X	X
Variância de textura - Banda 5	X	X	X

Fonte: a autora.

Figura 50– Classificações geradas para área de estudo 1: A- Classificação com os descritores selecionados a partir da ACP; B - Classificação com descritores selecionados a partir da ACP e AA; C - Classificação com todos os descritores; D - Imagem original (composição R3G2B1)



Fonte: a autora.

A Figura 50A ilustra o teste de classificação com os 21 descritores selecionados a partir da ACP. A Tabela 5 mostra a matriz de confusão gerada para essa classificação. Na matriz de confusão esse teste obteve exatidão geral de 88,9% e índice *kappa* de 0,87, esses valores são considerados excelentes de acordo com Landis e Koch (1977, apud Moreira 2011).

Tabela 5 – Matriz de confusão da classificação com os descritores selecionados a partir da ACP (área de estudo 1)

	Culti vo 1	Culti vo 2	Estradas / construções	Culti vo 3	Corpos hídricos	Culti vo 4	Culti vo 5	Solo exposto	Culti vo 7	Culti vo 6	Cobertura florestal
Cultivo 1	100	0	0	6,72	1,61	0	0	5,13	0	0	0
Cultivo 2	0	100	0	0	0	0	0	0	0	0	0
Estradas/construções	0	0	87,5	0	0	0	0	10,26	0	0	0
Cultivo 3	0	0	0	91,6	0	0	0	0	0	0	0
Corpos hídricos	0	0	0	0	96,77	0	0	0	0	0	0
Cultivo 4	0	0	5	0	0	21,67	0	2,56	0	0,82	0,81
Cultivo 5	0	0	0	0	0	0	95	0	0	0	0
Solo exposto	0	0	0,83	0	0	5	5	82,05	0	0	0
Cultivo 7	0	0	2,5	0	0	73,33	0	0	100	0	0
Cultivo 6	0	0	4,17	0	0	0	0	0	0	99,18	0
Cobertura florestal	0	0	0	1,68	1,62	0	0	0	0	0	99,19
Total	100	100	100	100	100	100	100	100	100	100	100

Fonte: a autora.

Analisando a Figura 50A, a matriz de confusão e considerando o conhecimento obtido em visita ao campo é possível identificar erros na classificação entre os cultivos 4 e 7, solo exposto e estradas/construções.

Nota-se que o cultivo 4 teve mais de 70% de seus *pixels* classificados como pertencentes ao cultivo 7, e somente 21,67% de acerto. Quanto ao solo exposto e estradas/construções há áreas que foram classificadas erroneamente. Para as duas classes nota-se que as áreas classificadas de modo errado encontram-se nas bordas de cultivos, como pode-se observar no limite entre o “cultivo 1” e “cobertura florestal”.

Como resultado da classificação também há informações relativas a área de cada classe. A Tabela 6 traz essas informações relativas ao teste de classificação realizado com os descritores selecionados a partir da ACP.

Tabela 6 - Áreas (m²) da classificação com os descritores selecionados a partir da ACP (área de estudo 1)

Classe	nº de objetos	área média	área mínima	área máxima	total
Cultivo 1	201	1.688,06	50,00	240.275,00	339.300,00
Cultivo 2	266	4.083,55	100,00	372.150,00	1.086.225,00
Cultivo 3	55	733,18	75,00	8.800,00	40.325,00
Cultivo 4	146	1.091,09	50,00	24.225,00	159.300,00
Cultivo 5	32	22.321,87	150,00	403.625,00	714.300,00
Cultivo 6	33	6.344,69	75,00	167.375,00	209.375,00
Cultivo 7	3	156.875,00	112.175,00	188.525,00	470.625,00
Solo exposto	118	834,11	50,00	11.800,00	98.425,00
Corpos hídricos	103	692,23	50,00	3.375,00	71.300,00
Cobertura florestal	611	907,61	50,00	14.250,00	554.550,00
Estradas e construções	234	811,43	50,00	9.525,00	189.875,00
Total	1802				3.933.600,00

Fonte: a autora.

A Figura 50B ilustra a classificação realizada com os descritores selecionados pela ACP acrescidos dos selecionados a partir da AA totalizando 40 descritores. Em comparação visual aos resultados foram superiores ao primeiro teste somente com descritores selecionados mediante ACP. Comparando com o teste anterior as classes solo exposto e cultivos 4 e 7 tiveram resultados superiores.

A matriz de confusão com os valores de acurácia para essa classificação encontra-se na Tabela 7.

Tabela 7– Matriz de confusão da classificação com os descritores selecionados a partir da ACP + AA (área de estudo 1)

	Cultivo 1	Cultivo 2	Estradas / construções	Cultivo 3	Corpos hídricos	Cultivo 4	Cultivo 5	Solo exposto	Cobertura florestal	Cultivo 6	Cultivo 7
Cultivo 1	100	0	14,17	3,36	2,42	5	5	5,13	0	0,82	0
Cultivo 2	0	97,5	0	0	0	0	0	0	0	0	0
Estradas e construções	0	0	82,5	0	0	0	0	10,26	0	0	0
Cultivo 3	0	0	0	96,64	0	0	0	0	0	0	0
Corpos hídricos	0	0	0	0	95,97	0	0	0	0,81	0	0
Cultivo 4	0	0	0	0	0	95	0	2,56	0	0	0
Cultivo 5	0	2,5	0	0	0	0	95	0	0	0	0
Solo exposto	0	0	0	0	0	0	0	82,05	0	0	0
Cobertura florestal	0	0	0	0	1,61	0	0	0	99,19	0	0
Cultivo 6	0	0	0,83	0	0	0	0	0	0	99,18	0
Cultivo 7	0	0	2,5	0	0	0	0	0	0	0	100
Total	100	100	100	100	100	100	100	100	100	100	100

Fonte: a autora.

Observando as imagens classificadas mediante descritores da ACP + AA (Figuras 50B) e com todos os descritores disponíveis (Figura 50C), e as matrizes de confusão (Tabelas 7 e 8) nota-se que ambas as classificações são iguais. Obtiveram os mesmos resultados, possuem o mesmo número de objetos classificados em cada classe e as áreas são iguais (Tabelas 9 e 10) ou seja, com 40 descritores se obteve o mesmo resultado que com todos os descritores, demonstrando que alguns desses são redundantes e não fazem diferença na classificação.

Para ambos os teste de classificação mediante descritores selecionados por meio da ACP e AA, e com todos os descritores, o resultado de exatidão geral na matriz de confusão foi 95,68% de exatidão e 0,95 no índice *kappa*. O cultivo 4 ficou com 95% de acerto nessas classificações, e o valor mais baixo foi para solo exposto com 82,05% e estradas e construções com 87,5%.

Tabela 8 – Matriz de confusão da classificação com os 54 gerados (área de estudo 1)

	Cultivo 1	Cultivo 2	Estradas e construções	Cultivo 3	Corpos hídricos	Cultivo 4	Cultivo 5	Solo exposto	Cobertura florestal	Cultivo 6	Cultivo 7
Cultivo 1	100	0	14,17	3,36	2,42	5	5	5,13	0	0,82	0
Cultivo 2	0	97,5	0	0	0	0	0	0	0	0	0
Estradas e construções	0	0	82,5	0	0	0	0	10,26	0	0	0
Cultivo 3	0	0	0	96,64	0	0	0	0	0	0	0
Corpos hídricos	0	0	0	0	95,97	0	0	0	0,81	0	0
Cultivo 4	0	0	0	0	0	95	0	2,56	0	0	0
Cultivo 5	0	2,5	0	0	0	0	95	0	0	0	0
Solo exposto	0	0	0	0	0	0	0	82,05	0	0	0
Cobertura florestal	0	0	0	0	1,61	0	0	0	99,19	0	0
Cultivo 6	0	0	0,83	0	0	0	0	0	0	99,18	0
Cultivo 7	0	0	2,5	0	0	0	0	0	0	0	100
Total	100	100	100	100	100	100	100	100	100	100	100

Fonte: a autora.

Tabela 9- Áreas (m²) das classificações com os descritores selecionados a partir da ACP+AA e todos os 54 descritores obtidos mediante a GEOBIA (área de estudo 1)

(Continua)

Classe	nº de objetos	área média	área mínima	área máxima	total
Cultivo 1	559	946,06	50,00	240.275,00	528.850,00
Cultivo 2	254	4.253,84	50,00	372.150,00	1.080.475,00
Cultivo 3	50	751,50	50,00	8.800,00	37.575,00
Cultivo 4	12	16.606,25	75,00	169.925,00	199.275,00
Cultivo 5	87	8.598,28	75,00	403.625,00	748.050,00

					(Conclusão)
Cultivo 6	110	2.732,05	125,00	167.375,00	300.525,00
Cultivo 7	2	150.350,00	112.175,00	188.525,00	300.700,00
Solo exposto	8	1.937,50	75,00	11.625,00	15.500,00
Corpos hídricos	83	795,18	75,00	3.375,00	66.000,00
Cobertura florestal	536	1.010,82	50,00	14.250,00	541.800,00
Estradas e construções	101	1.137,13	50,00	9.525,00	114.850,00
Total	1802				3.933.600,00

Fonte: a autora.

Embora a validação tenha valores satisfatórios sabemos que há algumas confusões na classificação. É o caso do cultivo 1 que aparece no entorno de outros cultivos e também da estrada, isso ocorre em virtude da vegetação campestre que não foi considerada como classe devido à semelhança com os cultivos.

4.4.2 Classificação e validação do uso e ocupação da terra para a área de estudo 2

Para a classificação do uso e ocupação da terra na área de estudo 2 foram consideradas as classes: cultivo 1, cultivo 2, cultivo 3, cultivo 4, reflorestamento 1, reflorestamento 2, corpos hídricos, cobertura florestal, solo exposto, estradas e construções.

Do mesmo modo que na área de estudo 1, foram realizados três testes de classificação para a área de estudo 2, sendo o primeiro utilizando somente os descritores selecionados por meio da ACP, o segundo utilizando descritores selecionados mediante ACP e AA, e o terceiro usando todos os descritores disponíveis. O Quadro 8 traz os descritores utilizados em cada teste de classificação

Quadro 8- Descritores utilizados para cada teste de classificação (ACP, ACP+AA, Todos) para a área de estudo 2

(Continua)

Descritores	ACP	ACP + AA	Todos
Desvio padrão dos números digitais - Banda 1	X	X	X
Desvio padrão dos números digitais - Banda 2	X	X	X
Desvio padrão dos números digitais - Banda 3			X
Desvio padrão dos números digitais - Banda 4	X	X	X
Desvio padrão dos números digitais - Banda 5	X	X	X
Entropia de textura - Banda 1	X	X	X
Entropia de textura - Banda 2	X	X	X
Entropia de textura - Banda 3	X	X	X
Entropia de textura - Banda 4	X	X	X
Entropia de textura - Banda 5	X	X	X
Média dos números digitais - Banda 1			X
Média dos números digitais - Banda 2	X	X	X
Média dos números digitais - Banda 3			X
Média dos números digitais - Banda 4			X
Média dos números digitais - Banda 5		X	X
Valor máximo de pixel - Banda 1		X	X
Valor máximo de pixel - Banda 2			X
Valor máximo de pixel - Banda 3			X
Valor máximo de pixel - Banda 4		X	X
Valor máximo de pixel - Banda 5			X
Valor mínimo de pixel - Banda 1			X
Valor mínimo de pixel - Banda 2			X
Valor mínimo de pixel - Banda 3			X
Valor mínimo de pixel - Banda 4		X	X
Valor mínimo de pixel - Banda 5		X	X
Área	X	X	X
Alongamento			X
Circuncidade	X	X	X
Compacidade			X
Comprimento			X
Convexidade	X	X	X
Direção principal	X	X	X
Eixo maior	X	X	X
Eixo menor			X
Fator de forma			X
Número de furos			X
Relação área/furos			X
Retangularidade			X
Solidez			X
Intervalo de textura - Banda 1			X
Intervalo de textura - Banda 2			X

(Conclusão)

Intervalo de textura - Banda 3	X	X	X
Intervalo de textura - Banda 4	X	X	X
Intervalo de textura - Banda 5	X	X	X
Textura média - Banda 1			X
Textura média - Banda 2	X	X	X
Textura média - Banda 3			X
Textura média - Banda 4			X
Textura média - Banda 5	X	X	X
Variância de textura - Banda 1	X	X	X
Variância de textura - Banda 2	X	X	X
Variância de textura - Banda 3	X	X	X
Variância de textura - Banda 4	X	X	X
Variância de textura - Banda 5	X	X	X

Fonte: a autora.

Na próxima página, a Figura 51 ilustra os testes de classificação realizados para a área de estudo 2.

Os três testes de classificação apresentam um bom resultado visual comparando com a imagem original. É possível notar a delimitação de cultivos, a diferenciação entre os dois tipos de reflorestamentos, bem como a cobertura florestal e os corpos hídricos.

Os resultados obtidos de acurácia mediante a matriz de confusão para os três testes de classificação são considerados excelentes, o quadro a seguir mostra esses valores.

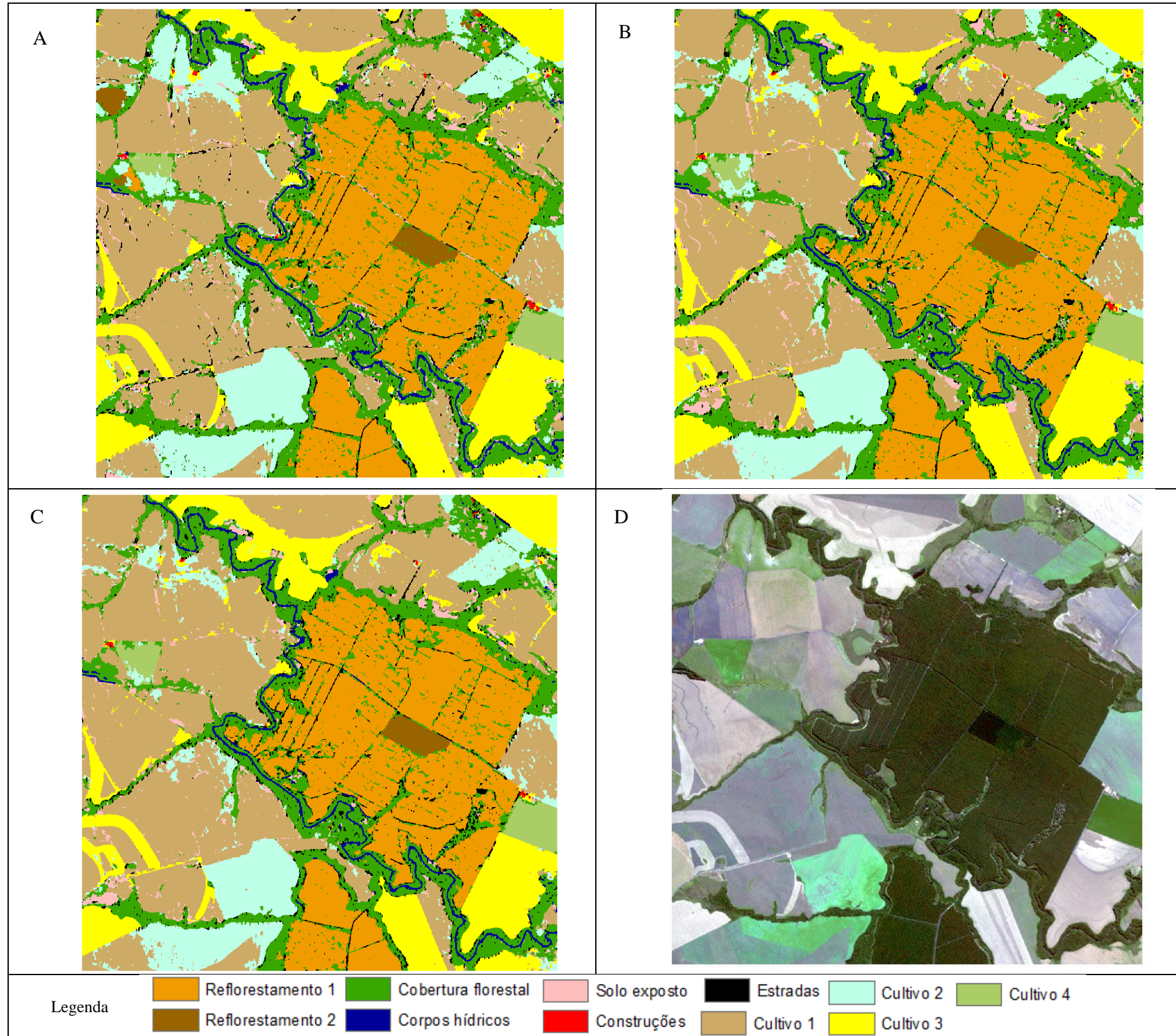
Quadro 9 – Acurácia dos testes de classificação para a área de estudo 2

	Matriz de confusão	Índice Kappa
ACP	91,22	0,90
ACP + AA	93,85	0,93
Todos	92,5	0,90

Fonte: a autora.

Nota-se que o melhor resultado em relação a acurácia foi do teste de classificação utilizando os descritores selecionados mediante a Análise de Componentes Principais e Análise de Agrupamentos, seguido do teste realizado com todos os descritores e posteriormente com os descritores selecionados somente a partir da ACP.

Figura 51 – Classificações da área de estudo 2: A - Classificação com descritores selecionados a partir da ACP; B - Classificação com descritores selecionados a partir da ACP e AA; C - Classificação com os 54 descritores gerados
 D - Imagem original (composição R3G2B1)



Fonte: a autora.

A Tabela 10 mostra a matriz de confusão para o teste de classificação utilizando os descritores selecionados somente mediante a ACP.

Tabela 10 – Matriz de confusão da classificação realizada com os descritores selecionados mediante a ACP para a área de estudo 2

	Cobertura floresta 1	Corpos hídricos	Construções	Estradas	Reflorestamento 1	Reflorestamento 2	Cultivo 1	Cultivo 2	Cultivo 3	Cultivo 4	Solo exposto
Cobertura florestal	91,67	2,04	0	22	0	0	0	0	0	5,88	12,5
Corpos hídricos	0	95,92	3,85	2	0	0	0	0	0	0	0
Construções	0	0	94,23	0	0	0	0	0	0	0	0
Estradas	8,33	0	0	66	0	0	0	0	0	0	0
Reflorestamento 1	0	0	0	8	100	0	0	0	0	0	0
Reflorestamento 2	0	0	0	0	0	100	0	0	0	0	0
Cultivo 1	0	0	0	0	0	0	100	0	0	0	25
Cultivo 2	0	0	0	0	0	0	0	100	0	5,88	0
Cultivo 3	0	0	1,92	0	0	0	0	0	100	0	0
Cultivo 4	0	0	0	0	0	0	0	0	0	88,24	0
Solo exposto	0	2,04	0	2	0	0	0	0	0	0	62,5
	100	100	100	100	100	100	100	100	100	100	100

Fonte: a autora.

As classes com maiores acertos são reflorestamentos 1 e 2, e cultivos 1, 2 e 3. As classes solo exposto e cultivo 4 aparecem com os menores valores de acertos. Observando a matriz de confusão e analisando visualmente a Figura 51A nota-se alguns erros de classificação, é o caso da confusão entre estradas e solo exposto, bem como estradas e cobertura florestal. Também há confusão entre cultivo 1 e o reflorestamento 2, conforme se pode observar na área superior esquerda da figura 32A.

A Tabela a seguir traz os resultados de área para cada classe obtidos partir da classificação com os descritores selecionados mediante a ACP.

Tabela 11 - Áreas da classificação com os descritores selecionados mediante ACP (área de estudo 2)

(Continua)

	nº de objetos geográficos	Área média	Área mínima	Área máxima	Área total
Cobertura florestal	15.379	283,97	25,00	74.750.000,00	4.367.200,00
Corpos hídricos	1.335	262,49	50,00	74.500.000,00	350.425,00
Construções	148	207,60	50,00	16.750.000,00	30.725,00

(Conclusão)

Estradas	3.267	198,87	50,00	23.250.000,00	649.700,00
Reflorestamento 1	268	20.385,73	10,00	627.050,00	5.463.375,00
Reflorestamento 2	75	3.343,33	50,00	69.625.000,00	250.750,00
Cultivo 1	225	33.512,00	75,00	813.750,00	7.540.200,00
Cultivo 2	4.642	496,95	50,00	91.125.000,00	2.306.850,00
Cultivo 3	2.273	1.254,54	25,00	468.350,00	2.851.575,00
Cultivo 4	313	1.138,82	50,00	63.625.000,00	356.450,00
Solo exposto	2.920	250,98	25,00	79.500.000,00	732.850,00
Total	30.845				24.900.100,00

Fonte: a autora.

O teste de classificação que apresentou melhores resultados para a área de estudo 2 foi o que utilizou os descritores selecionados mediante a Análise de Componentes Principais e Análise de Agrupamentos, como observado anteriormente no Quadro 8.

Analisando visualmente as classificações ilustradas nas Figuras 51A e 51B junto ao conhecimento de campo e observando a imagem original, constatou-se que a classificação relativa ao segundo teste de classificação tem menos erros de classificação que a primeira.

Na matriz de confusão do segundo teste de classificação (Tabela 12) observa-se que as classes com maior percentual de acerto foram reflorestamentos, e cultivos 1, 2 e 3, estes com 100% de acerto. E as classes com maiores erros são estradas, essa que possui o percentual de 66% considerado insatisfatório, e solo exposto.

Observa-se comparando a matriz de confusão do primeiro e segundo testes de classificação que já há melhora nos resultados para a classe solo exposto que aumentou o percentual de acerto.

A Tabela 13 apresenta os resultados de área relativos ao segundo teste de classificação.

Tabela 12 – Matriz de confusão da classificação realizada com descritores selecionados mediante ACP + AA para a área de estudo 2

	Cobertura florestal	Corpos hídricos	Construções	Estradas	Reflorestamento 1	Reflorestamento 2	Cultivo 1	Cultivo 2	Cultivo 3	Cultivo 4	Solo exposto
Cobertura florestal	95,83	2,04	0	22	0	0	0	0	0	5,88	0
Corpos hídricos	0	95,92	0	2	0	0	0	0	0	0	0
Construções	0	0	92,31	0	0	0	0	0	0	0	0
Estradas	4,17	0	0	66	0	0	0	0	0	0	0
Reflorestamento 1	0	0	0	8	100	0	0	0	0	0	0
Reflorestamento 2	0	0	0	0	0	100	0	0	0	0	0
Cultivo 1	0	0	0	0	0	0	100	0	0	0	12,5
Cultivo 2	0	0	0	0	0	0	0	100	0	1,96	0
Cultivo 3	0	0	7,69	0	0	0	0	0	100	0	0
Cultivo 4	0	0	0	0	0	0	0	0	0	92,16	0
Solo exposto	0	2,04	0	2	0	0	0	0	0	0	87,5
	100	100	100	100	100	100	100	100	100	100	100

Fonte: a autora.

Tabela 13 - Áreas (m²) da classificação com os descritores selecionados mediante ACP e AA (área de estudo 2)

	nº de objetos geográficos	Área média	Área mínima	Área máxima	Área total
Cobertura florestal	14.794	287,37	25,00	8.400,00	4.251.425,00
Corpos hídricos	1.295	263,98	50,00	7.450,00	341.850,00
Construções	92	225,82	50,00	1.675,00	20.775,00
Estradas	3.182	190,78	50,00	2.325,00	607.050,00
Reflorestamento 1	399	13.734,90	50,00	627.050,00	5.480.225,00
Reflorestamento 2	73	2.209,59	75,00	60.200,00	161.300,00
Cultivo 1	1.313	6.078,98	50,00	813.750,00	7.981.700,00
Cultivo 2	4.758	441,07	50,00	22.900,00	2.098.600,00
Cultivo 3	2.664	1.115,82	25,00	468.350,00	2.972.550,00
Cultivo 4	365	988,97	50,00	22.900,00	360.975,00
Solo exposto	1.910	326,52	25,00	11.250,00	623.650,00
Total	30.845				24.900.100,00

Fonte: a autora.

O terceiro teste de classificação realizado para a área de estudo 2 utilizou todos os 54 descritores disponíveis. A Figura 51C ilustra a classificação, e a Tabela 14 a matriz de confusão.

Tabela 14 – Matriz de confusão para classificação com todos os descritores para a área de estudo 2

	Cobertura florestal 1	Corpos hídricos	Construções	Estradas	Reflorestamento 1	Reflorestamento 2	Cultivo 1	Cultivo 2	Cultivo 3	Cultivo 4	Solo exposto
Cobertura florestal	95,83	2,04	0	24	0	0	0	0	0	5,88	0
Corpos hídricos		95,92	0	2	0	0	0	0	0	0	0
Construções			86,54		0	0	0	0	0	0	0
Estradas	4,17	2,04	7,69	66	0	0	0	0	0	0	0
Reflorestamento 1	0	0	0	8	100	0	0	0	0	0	0
Reflorestamento 2	0	0	0	0	0	100	0	0	0	0	0
Cultivo 1	0	0	0	0	0	0	100	0	0	0	25
Cultivo 2	0	0	0	0	0	0	0	100	0	0	0
Cultivo 3	0	0	5,77	0	0	0	0	0	100	0	0
Cultivo 4	0	0	0	0	0	0	0	0	0	94,12	0
Solo exposto	0	0	0	0	0	0	0	0	0	0	75
	100	100	100	100	100	100	100	100	100	100	100

Fonte: a autora.

O resultado do terceiro teste tem a segunda melhor acurácia conforme o Quadro 8, no entanto este teste utilizado com 54 descritores não tem melhor resultado que realizado com somente 33 descritores, demonstrando que há alguns que causam prejuízo a classificação.

A Tabela 15 mostra os resultados de área para o terceiro teste de classificação:

Tabela 15 - Áreas (m²) da classificação com os 54 descritores gerados (área de estudo 2)

	nº de objetos geográficos	Área média	Área mínima	Área máxima	Área total
Cobertura florestal	14.651	282,28	25,00	8.400,00	4.135.750,00
Corpos hídricos	1.297	266,48	50,00	7.450,00	345.625,00
Construções	78	197,44	50,00	900,00	15.400,00
Estradas	3.113	187,25	50,00	2.175,00	582.900,00
Reflorestamento 1	283	19.260,34	75,00	627.050,00	5.450.675,00
Reflorestamento 2	78	2.103,53	75,00	60.200,00	164.075,00
Cultivo 1	2.626	3.240,12	50,00	813.750,00	8.508.550,00
Cultivo 2	4.433	444,50	50,00	21.300,00	1.970.450,00
Cultivo 3	2.654	1.089,44	25,00	468.350,00	2.891.375,00
Cultivo 4	384	1.021,29	50,00	63.625,00	392.175,00
Solo exposto	1.248	355,07	25,00	7.950,00	443.125,00
	30.845				24.900.100,00

Fonte: a autora.

Observa-se que para cada classificação o número de objetos geográficos e as áreas diferem em virtude do resultado da classificação utilizando diferentes descritores.

Para a área de estudo 2 admite-se a classificação com os descritores selecionados mediante AA e ACP como o melhor resultado, pois possibilitou melhora na acurácia com menor quantidade de descritores na classificação.

5 Considerações finais:

Verificou-se nesse trabalho que a GEOBIA, traduzida como Análise Orientada a Objetos Geográficos, é uma metodologia para o Processamento de Imagens Digitais de Sensoriamento Remoto bastante relevante, principalmente se tratando do uso de imagens de alta resolução como é o caso da *RapidEye*.

Para essa nova abordagem no processamento de imagens nota-se que o conhecimento do intérprete é importante, pois é a partir da fotointerpretação que serão definidos os parâmetros para a segmentação (escala e fusão), se estes se adequam a realidade da área de estudo, e também é o interprete que definirá as classes de uso que culminarão na classificação. Ou seja, antes mesmo de iniciar os processamentos é importante que o técnico analise as imagens em diversas composições coloridas e também que faça trabalho de campo, deste modo conseguirá melhor conhecimento da área de estudo que culminará depois em parâmetros da GEOBIA melhor definidos.

A GEOBIA gera banco de dados relacional com diversos descritores espaciais, espectrais e de textura. O uso destes descritores é crucial na extração de informações, no entanto devido constituírem banco de dados com diversas pode haver por parte do usuário indecisão em relação a quais descritores utilizar, quais deles de fato fará diferença significativa em seu trabalho ou que poderá acarretar em redundâncias ou ruídos.

Para resolver a indecisão de quais descritores utilizar nesse trabalho recorreu-se às Análises Multivariadas de Componentes Principais e de Agrupamentos, com o objetivo de selecionar descritores para obter melhor acurácia para as classificações supervisionadas da área de estudo, por isso buscou-se na literatura metodologias que possibilitassem essa seleção.

Como o banco de dados gerados mediante a GEOBIA possui diversas variáveis, utilizou-se a Análise de Componentes Principais para diminuição da dimensionalidade de dados. Com os métodos de Beale e B4 propostos por Joliffe (1972) foi possível a associação para cada uma das componentes principais geradas um descritor proveniente da GEOBIA, além da exclusão e seleção de descritores.

Os descritores selecionados mediante os métodos de Beale e B4 da Análise de Componentes Principais foram utilizados em testes para classificar as duas áreas de estudo. Em ambas as áreas os resultados são considerados excelentes de acordo com Landis (1977) *apud* Moreira (2011).

No entanto, utilizando a metodologia a partir da ACP são excluídos os descritores associados as componentes principais que possuem menor variância dos dados, e analisando os dendrogramas gerados mediante a Análise de Agrupamentos verificou-se que há descritores excluídos pela ACP que possibilitam a separabilidade de algumas classes.

Portanto, utilizou-se no segundo teste de classificação para ambas as áreas os descritores selecionados tanto pela ACP quanto os que notou-se na AA melhor separabilidade das classes. Em ambas as áreas de estudo obteve-se melhor resultado com essa combinação para a seleção de descritores.

Um terceiro teste utilizando todos os descritores disponíveis foi realizado nas duas áreas de estudo para comparação com os testes anteriores. Verificou-se com isso que o resultado aliando ACP e AA foi igual ou superior aos que utilizavam todos os descritores, que os que utilizavam só a ACP embora considerado pela literatura como excelente sempre era menor que as duas outras condições.

As três classificações geradas para as duas áreas de estudo obtiveram resultados excelentes, no entanto possuem erros de classificação que correspondem a falta de separabilidade entre classes, essa que foi menor ou maior dependendo do teste de classificação, mas sempre existente e demonstrando que ainda carece estudos para uma classificação fidedigna à realidade. Foi possível perceber também que as bandas 4 e 5 dos comprimentos de onda do *red-edge* e infravermelho próximo, respectivamente, são as que mais se destacarão na seleção de descritores.

Considera-se os testes de classificação utilizando ACP e AA os mais adequados e com esse resultado pôde-se conhecer do uso e ocupação das duas diferentes áreas de estudo.

6 Referências:

ABBAS S., QAMER F. M., RANA A. D., HUSSAIN N., SALEEM R. **Application Of Object Based Image Analysis For Forest Cover Assessment Of Moist Temperate Himalayan Forest In Pakistan**. ISPRS .GEOBIA 2010: Geographic Object-Based Image Analysis. Volume XXXVIII-4/C7, 2010. IEEE Transaction on Geoscience and Remote Sensing. v. 39, n. 2, p. 2680-2688, 2010.

ABSON, D. J.; DOUGILL, A. J.; STRINGER, L. C. Using Principal Component Analysis for information-rich socio-ecological vulnerability mapping sin Southern Africa. **Applied Geography**, v. 35, 2012, p. 515-524.

ADAMI, M. **Estudo da dinâmica espaço-temporal do bioma Pantanal por meio de imagens MODIS**. *Pesq. agropec. bras.* [online]. 2008, vol.43, n.10, pp. 1371-1378. ISSN 1678-3921. Acesso em 15/01/2015.

ALENCAR, B. J., **A análise multivariada aplicada ao tratamento da informação espacial: uma abordagem matemático-computacional em análise de agrupamentos e análise de componentes principais**. Tese (Doutorado) /. PUC: Belo Horizonte, 2009. 200f.

ANTUNES, A. F. B. **Classificação de Ambiente Ciliar Baseada em Orientação a Objeto em Imagens de Alta Resolução Espacial**. 2003. 145 fl. Tese (Doutorado) – Universidade Federal do Paraná, Curitiba, 2003.

ARROYO, L.A; K JOHANSEN, K.; PHINN, S. **MAPPING LAND COVER TYPES FROM VERY HIGH SPATIAL RESOLUTION IMAGERY: AUTOMATIC APPLICATION OF AN OBJECT BASED CLASSIFICATION SCHEME**. ISPRS .GEOBIA 2010: Geographic Object-Based Image Analysis. Volume XXXVIII-4/C7, 2010.

BLASCHKE, T.; GLÄSSER, C. e LANG, S.: **Processamento de Imagens num ambiente Integrado SIG / Sensoriamento Remoto – Tendências e Consequencias**. In **Sensoriamento e SIG avançados: novos sistemas sensores: métodos inovadores / versão brasileira atualizada e organizada por Thomas Blaschke e Hermann Kux, tradução Hermann Kux**. – 2ª ed. – São Paulo: Oficina de Textos, 2007.

BORATTO, I. M. P.; AVELLAR, G., **Espacialização, quantificação e classificação das áreas verdes do município de Sete Lagoas, utilizando as ferramentas de modelagem (NDVI) e de técnicas matemáticas - estatística de componentes principais e a análise de agrupamento.** Anais XVI Simpósio Brasileiro de Sensoriamento Remoto - SBSR, Foz do Iguaçu, PR, Brasil, 13 a 18 de abril de 2013, INPE.

BOROUCHE, J. M., SAPORTA. G. **Análise de dados.** Zahar Editores. Rio de Janeiro, 1982. 116 p.

CÂMARA, G. et al. **Anatomia de Sistemas de Informação Geográfica.** Campinas: Instituto de Computação, UNICAMP, 1996. 197p.

CÂMARA, G.; DAVIS. C.; Introdução. In: CÂMARA, G.; DAVIS. C.; MONTEIRO, A. M.; Introdução à Ciência da Geoinformação. São José dos Campos, INPE, 2001. Disponível em: <http://www.dpi.inpe.br/gilberto/livro/introd/>. Acesso em 20/12/2014.

CÂMARA, G.; DAVIS. C.; MONTEIRO, A. M.; Introdução à Ciência da Geoinformação. São José dos Campos, INPE, 2001. Disponível em: <http://www.dpi.inpe.br/gilberto/livro/introd/>. Acesso em 25/09/2013.

CÂMARA, G; MONTEIRO, A.M.; Conceitos básicos em Ciência da Geoinformação. In: CÂMARA, G.; DAVIS. C.; MONTEIRO, A. M.; Introdução à Ciência da Geoinformação. São José dos Campos, INPE, 2001. Disponível em: <http://www.dpi.inpe.br/gilberto/livro/introd/>. Acesso em 20/12/2014.

CENTENO, J. A. S. **Sensoriamento remoto e processamento de imagens digitais.** Curitiba: Ed. Curso de Pós Graduação em Ciências Geodésicas, Universidade Federal do Paraná, 2004, 219 p.

CARMO, M. R. B do, MORO, R. S.; NOGUEIRA, M. K. F. S., KACZMARECH, R. A. vegetação ripária ao longo do rio Pitangui. In: GEALH, A. M.; MELO, M. S.; MORO, R. S. (Orgs.). **Pitangui, rio de contrastes: seus lugares, seus peixes, sua gente.** Ponta Grossa:Ed. UEPG, 2010. 199p.

CLARO, L. O. T. **Texturas de imagens utilizando conceitos de geomorfologia matemática.** Porto Alegre: CEP SRM/CPGSR – UFRGS, 1995.

CONCI, A. Segmentação de Imagem. Material de aula. Disponível em: <http://www2.ic.uff.br/~aconci/CV.html>. Acesso em 18/12/2014.

CONGALTON, R. A review of assessing the accuracy of classifications of remotely sensed data. **Remote Sensing of Environment**, v. 49, n. 12, p. 1671–1678, 1991.

CRUZ, G. C. F. Alguns aspectos do clima dos Campos Gerais. In: MELO, M.S.; MORO, R.S.; GUIMARÃES, G.B. (Eds.). **Patrimônio Natural dos Campos Gerais do Paraná**. Ponta Grossa: Ed. UEPG, 2007. p. 59-72.

EMBRAPA. **Centro Nacional de Pesquisa de Florestas. Zoneamento Ecológico para Plantios Florestais no Estado do Paraná**. Brasília: EMBRAPA-DDT, 2006.

EXELISVIS. Feature extraction – List of Attributes. Disponível: <http://www.exelisvis.com/docs/attributelist.html>. Acesso em 20/08/2014.

FAN, G.; XIA, X.. **A Joint Multicontext and Multiscale Approach to Bayesian Image Segmentation**. IEEE Transaction on Geoscience and Remote Sensing. v. 39, n. 2, p. 2680-2688, 2001.

FELIX, I. M.; KAZMIERCZAK, M. L.; ESPINDOLA, G. M. de. RapidEye: a nova geração de satélites de Observação da Terra. In: Anais XIV Simpósio Brasileiro de Sensoriamento Remoto, Natal, Brasil, 25-30 abril, 2009. INPE, p. 7619-7622.

FIGUEIREDO, Divino. Conceitos Básicos de Sensoriamento Remoto. 2005. Disponível em:

http://www.conab.gov.br/conabweb/download/SIGABRASIL/manuais/conceitos_sm.pdf . Acesso em: 05 set 2014..

FRITZSONS, E; MANTOVANI, L. E.; **Os Substratos Geológicos e os Coeficientes Morfométricos em Bacias Hidrográficas do Carste Dolomítico no Estado do Paraná**. Revista Brasileira de Geografia Física, 03 (2010) 181-189.

GEALH, A.M.; MELO, M.S.; MORO, R.S. (Orgs.). **Pitangui, Rio de Contrastes: seus lugares, seus peixes, sua gente**. Ponta Grossa: Ed. UEPG, 2010. 199p.

GELLI, G. Apresentação. In: IBGE. **Introdução ao Processamento Digital de Imagens**. Rio de Janeiro: IBGE, 2001, 94p.

GONZALEZ, R.; WOODS, R. **Processamento de Imagens Digitais**. 3. ed. São Paulo: Edgard Blucher Ltda, 2000.

GORE, Paul A. Jr. Cluster Analysis. **Handbook of Applied Multivariate Statistics and Mathematical Modeling 2000** by Academic Press.

GUIMARÃES, G. B. **As rochas Granitóides do Complexo Granítico Cunhaporanga, Paraná: Aspectos Geológicos, Geofísicos, Geoquímicos e Mineralógicos**. 2000, 230 fl. Tese (Doutorado em Geociências) – Instituto de Geociências, Universidade de São Paulo, São Paulo, 2000.

HAIR, J. F. et al. **Análise Multivariada de dados**. 6. ed. Porto Alegre: Bookman, 2009.

Hay, G.J., Castilla, G., 2008. **Geographic Object-Based Image Analysis (GEOBIA): A new name for a new discipline**. In: Blaschke, T., Lang, S., Hay, G. (Eds.), *Object Based Image Analysis*. Springer, Heidelberg, Berlin, New York, pp. 93–112.

Castilla, G. Hay, G.J. (2008) **Image objects and geographic objects**. In: *Object-Based Image Analysis: Spatial Concepts for Knowledge-Driven Remote Sensing Applications* (eds. Blaschke, T., Lang, S. and Hay, G.J.). Springer, Berlin, 817 pp.

Hsu, C.-W., C.-C. Chang, and C.-J. Lin. (2010). *A practical guide to support vector classification*. National Taiwan University. <http://ntu.csie.org/~cjlin/papers/guide/guide.pdf>

ISA/UTL. **Análises classificatórias (clustering)**. Mestrado em Matemática – Estatística Multivariada – 2009/2010. Disponível em: <http://www.isa.utl.pt/dm/mestrado/mmach/UCs/em/webEMCap4.pdf>. Acesso em: 20 de agosto 2014.

JENSEN, J. R. **Sensoriamento remoto do ambiente: uma perspectiva em recursos terrestres**. 2nd. ed. São José dos Campos: [s.n.], 2009.

JESUS, S. C.; EPIPHANIO, J. C. N. Sensoriamento Remoto multissensores para a avaliação temporal da expansão agrícola municipal. **Bragantia**, Campinas, v.69, n. 4, p. 946-956, 2010.

JOHNSON, R.A.; WICHERN, D.W. **Applied multivariate statistical analysis**. New Jersey: Prentice Hall, 5ª ed., 2002. 767p.

JOLLIFFE, I.T.. Discarding variables in a principal component analysis. I. Artificial data. *Appl. Stat.*, 21(2):160-173. 1972.

JOLLIFFE, I.T.. Discarding variables in a principal component analysis. II. Real data. *App. Stat.*, 21(1):21-31. 1973.

JUNGES, A. H.; FONTANA, D. C. e PINTO, D. G. Identification of croplands of winter cereals in Rio Grande do Sul state, Brazil, through unsupervised classification of normalized difference vegetation index images. **Eng. Agríc.** , vol.33, n.4 [cited 2013-11-14], 2013, p. 883-895 .

Koepfler, G., Lopez, C., and Morel, J. M. (1994). **A multiscale algorithm for image segmentation by variational method**. *SIAM J. Numer. Anal.*, 31(1):282–299.

LANG, S.; BLASCHKE, T.; Análise da Paisagem com SIG. Tradução Hermann Kux. São Paulo : Oficina de Textos, 2009.

LANDIS, J. R.; KOCH, G. G. The measurement of observer agreement for categorical data. *Biometrics*, v. 33, n. 1, p. 159-74, 1977.

LARSON, R; FARBER, B. Estatística Aplicada. 4 ed. São Paulo: Person Prentice Hall, 2010.

LAUREANO, P. **Atlas Del Agua. Los conocimientos tradicionales para combatir la desertificación**. UNESCO-Laia Libros.Barcelona.2005.

MAACK, R. Geografia Física do Estado do Paraná. Ponta Grossa, Ed: UEPG, 4ed. 2012.

MARQUES, J. Notas de aula. 2014.

MARQUES, J. M.; MENDES MARQUES, A. As componentes principais no descarte de variáveis em um modelo de regressão múltipla. *Revista FAE*, Curitiba, v.8, n.1, p.93-101, jan./jun. 2004.

MATHER, P. M. **Computer processing of remotely-sensed images: an introduction**. 2. ed. Chichester: John Wiley & Sons. 1999. 292 p.

MELO JÚNIOR, J. C. F; SEDIYAMA, G. C. ; FERREIRA, Paulo A.; LEAL, Brauliro G. **Determinação de regiões homogêneas quanto à distribuição de frequência de**

chuvas no leste do Estado de Minas Gerais. Rev. bras. eng. agríc. ambient. vol.10 no.2 Campina Grande Apr./June 2006.

MELO, M. S.; GUIMARÃES, G. B.; SANTANA, A. C. Fisiografia da bacia do rio Pitangui. In: GEALH, A.M et al. **Pitangui, rio de contrastes: seus lugares, seus peixes, sua gente.** Ponta Grossa:Ed. UEPG, 2010. 199p

MELO, M. S. et al. Relevo e hidrografia dos Campos Gerais. In: MELO, M. S.; MORO, R. S.; GUIMARÃES, G. B. **Patrimônio Natural dos Campos Gerais do Paraná.** Ponta Grossa: Editora UEPG, 2007. p. 49-58.

MELO, M. S. de. **Formas Rochosas do Parque Estadual de Vila Velha.** Ponta Grossa: UEPG, 2006. 154p.

MENEGUZZO, I. S.; MELO, M. S. **Dicionário Histórico e Geográfico dos Campos Gerais.** Disponível em: <<http://www.uepg.br/dicion/index.htm>>Acesso em: 10 agosto 2012.

MINEROPAR, Minerais do Paraná. **Atlas Geomorfológico do Paraná.** 2006.

MOREIRA, Maurício Alves. **Fundamentos do sensoriamento remoto e metodologias de aplicação.** 4. ed. atual. e ampl. Viçosa: UFV, 2011. 422 p.

NETO, José Machado Moita; MOITA, Graziella Ciaramella **Uma Introdução À Análise Exploratória De Dados Multivariados.** QUÍMICA NOVA, 21(4) (1998).

OLIVEIRA, F. P.; FERNANDES FILHO, E. I.; SOARES, V. P.; SOUZA, A. L. **Mapeamento de fragmentos florestais com monodominância de aroeira a partir da classificação supervisionada de imagens Rapideye.** Rev. *Árvore* [online]. 2013, vol.37, n.1, pp. 151-161. ISSN 0100-6762.

PETRISOR, A.; IANOS, I.; IUREA, D.; VAIDIANU, M. Applications of Principal Component Analysis integrated with GIS. **Procedia Environmental Sciences**, v. 14, 2012, p. 247-256.

REGAZZI, A. J. INF 766 - Análise multivariada. Viçosa: Universidade Federal de Viçosa, Centro de Ciências Exatas e Tecnológicas. Departamento de Informática, 2001. 166p. Apostila de disciplina.

REGO, Luiz Felipe Guanaes. **O uso de sistemas de informações geográficas para o estudo do espaço geográfico.** In Paisagem, espaço e sustentabilidades: uma perspectiva multidimensional da Geografia. Organização João Rua. Rio de Janeiro: Ed. PUC-Rio, 2007.

REZENDE, M. L., FERNANDES, L. P. S., SILVA, A. M. R.; **Utilização da análise fatorial para determinar o potencial de crescimento econômico em uma região sudeste do Brasil.** Revista Economia e Desenvolvimento, nº 19, 2007.

RIBEIRO, Selma Regina Aranha. **Integração de Imagens Landsat TM e dados auxiliares para a delimitação automática de unidade fisionômicas (paisagem) e geológicas usando redes neurais artificiais.** Tese (Doutorado) – Universidade Federal do Paraná, Curitiba, 2003.

RIBEIRO, Selma Regina Aranha. **Comparação entre classificações com rede neural artificial em diferentes áreas de estudo no Paraná.** Mercator. Fortaleza, v.12, n.27, p. 159-168, jan./ abr. 2013

RICHARDS, John. A. **Remote Sensing Digital Image Analysis - An Introduction.** 2nd ed. Springer-Verlag, Berlin. 1993. 281p

ROBINSON, D.; REDDING, N.; CRISP, D. **Implementation of a Fast Algorithm for Segmenting SAR Imagery.** [S.l.]: DSTO - Electronics and Surveillance Research Laboratory, 2002.

ROCHA, C. H.; WEIRICH, P. H. N. Padrões de uso das terras e implicações ambientais. In: GEALH, A.M et al. **Pitangui, rio de contrastes: seus lugares, seus peixes, sua gente.** Ponta Grossa:Ed. UEPG, 2010. 199p.

ROSA, R. **Geotecnologias na Geografia Aplicada.** Revista do Departamento de Geografia da USP, São Paulo, nº 16 (2005), 81-90.

SCHOWENGERDT, R. A. **Spectral Transforms in: Remote Sensing: Models and Methods for Image Processing.** London: Academic Press, 1997. 522p.

SEARA, D. M. ; ELIZANDRO, G. C. S. E. Encontrando a linha divisória: detecção de bordas. In: WANGENHEIM, A. **Introdução à Visão Computacional.** PPGCC-INE-UFSC. 2005. Disponível em <http://www.inf.ufsc.br/~visao/>. Acesso em 19/12/2014.

SILVA, J. dos S. V. da, CARVALHO, J. R. P, SANTOS, R. dos, FELGUEIRAS, C. A.; **Espacialização da homogeneidade e heterogeneidade do meio biofísico para planejamento ambiental utilizando análise multivariada.** Anais 1º Simpósio de Geotecnologias no Pantanal, Campo Grande, Brasil, 11-15 novembro 2006, Embrapa Informática Agropecuária/INPE, p.691-699.

SOBEL, I. E. Camera Models and Machine Perception. Dissertação Ph.D. Stanford University, Palo Alto, California.

SOLBERG, A.H.S. et al. **A Markov random field model for classification of multisource satellite imagery.** IEEE Transactions on Geosciences and Remote Sensing. 1996. v.34, n.1, p. 100-112.

SWAIN, P.H.; DAVIS, S.M., ed. **Remote sensing: the quantitative approach.** New York, McGraw-Hill, 1978. 396 p.

TRENTIN, C. B.; PARDI LACRUZ, M. S.; JÚNIOR, M. A. S.; SILVA, V. O. **Análise da Ocorrência de Enchentes na Área de Abrangência do Rio Jacuí/RS com a Utilização de Imagens MODIS e Dados SRTM.** In: V SEMINÁRIO LATINO-AMERICANO E I SEMINÁRIO IBERO-AMERICANO DE GEOGRAFIA FÍSICA, 2008, Santa Maria. Anais... 2008. p. 3803-3815. DVD.

TSO, B.; MATHER, P. **Classification methods for remotely sensed data.** 2. ed. [S.l.]: CRC Press, 2009.

VENTURIERI, A.; SANTOS, J.R. Técnicas de Classificação de Imagens para Análise de Cobertura Vegetal. In: ASSAD, E.D.; SANO, E.E. (Ed.). **Sistema de Informações Geográficas.** Brasília: Embrapa-SPI/Embrapa-CPAC. Capítulo 18, p. 351-371, 1998.

VICINI, Lorena. **Análise Multivariada da teoria à prática.** Santa Maria: UFSM, CENE, 2005. 215p.

VITTE, Antonio Carlos. **Contribuições à história e à epistemologia da Geografia.** Rio de Janeiro: Bertrand Brasil, 2007.

VELOSO, H. P.; RANGEL FILHO, A. L.; LIMA, J. C. A. **Classificação da Vegetação Brasileira, adaptada a um sistema universal.** Rio de Janeiro: IBGE, 1991. 123p. Disponível em: <http://biblioteca.ibge.gov.br/visualizacao/monografias/GEBIS%20-%20RJ/classificacaovegetal.pdf>. Acesso em: 18 ago. 2014

VETORAZZI, C. A.; Técnicas de Geoprocessamento no monitoramento de áreas florestadas. Série Técnica IPEF, Piracicaba: IPEF. v.10, n.29, p.45-51. 1996.

WANGENHEIM, A. **Introdução à Visão Computacional**. PPGCC-INE-UFSC. 2005.
Disponível em <http://www.inf.ufsc.br/~visao/>. Acesso em 19/12/2014.

7 Anexo

Script da ACP no software R

```
#ANALISE DE COMPONENTES PRINCIPAIS NO R
```

```
#Criado por: Kelly Lais Wiggers
```

```
#Mestranda em Computacao Aplicada
```

```
#UEPG - Universidade Estadual de Ponta Grossa (Brasil)
```

```
#Agradecimentos: CAPES pela concessão da minha bolsa de estudos.
```

```
#e-mail: kellylaiswiggers@hotmail.com
```

```
#Carrega arquivo de dados
```

```
dados <- read.table
```

```
#Adiciona o nome das variaveis
```

```
variaveis <-
```

```
("SAVGB_1","SAVGB_2","SAVGB_3","SAVGB_4","SAVGB_5","SSTDB_1","SSTDB_2","SSTDB_3","SSTDB_4","SSTDB_5","SMINB_1","SMINB_2","SMINB_3","SMINB_4","SMINB_5","SMAXB_1","SMAXB_2","SMAXB_3","SMAXB_4","SMAXB_5","TRANB_1","TRANB_2","TRANB_3","TRANB_4","TRANB_5","TAVGB_1","TAVGB_2","TAVGB_3","TAVGB_4","TAVGB_5","TVARB_1","TVARB_2","TVARB_3","TVARB_4","TVARB_5","TENTB_1","TENTB_2","TENTB_3","TENTB_4","TENTB_5","FXAREA","FXLENGTH","FXCOMPACT","FXCONVEX","FXSOLID","FXROUND","FXFORMFAC","FXELONG","FXRECTFIT","FXMAINDIR","FXMAJAXLN","FXMINAXLN","FXNUMHOLE","FXHOLESOL")
```

```
names(dados) <- variaveisprint(dados)
```

```
#Transforma os dados de data.frame para matriz
```

```
dadosMatriz <- data.matrix(dados, rownames.force = NA)
```

```
#Pre-processamento dos dados
```

```
media <- apply(dadosMatriz,1,mean)
```

```
desvio <- apply(dadosMatriz,1,sd)
```

```
preproc <- (dadosMatriz - media)/desvio
```

```
#Cria a matriz de correlacao com os dados pre-processados
```

```
matcor <- cor(preproc)
```

```
#Inicio da ACP
```

```
#Realiza o processo de ACP
```

```
acCPor <- prcomp(preproc)
```

```
summary(acCPor)
#Mostra variancia explicada
varianciaExplicada <- summary(acCPor)
str(varianciaExplicada)
varianciaExplicada <- varianciaExplicada$importance[2,] * 100
#print(varianciaExplicada)
#Biplot da ACP
#biplot(acCPor,col = c("gray", "black"))
barplot(varianciaExplicada, xlab = "Componente principal ", ylab = "Variancia
explicada (%)")
#Loadings
loadings <- acCPor$rotation
#Scores
scores <- acCPor$x
#Salva arquivo com os loadings
write.table(loadings, file)
#Salva arquivo com os scores
write.table(scores, file)
```