MINISTÉRIO DA CIÊNCIA, TECNOLOGIA, INOVAÇÕES E COMUNICAÇÕES
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

sid.inpe.br/mtc-m21b/2016/08.23.15.02-TDI

COMPARAÇÃO ENTRE CLASSIFICAÇÕES DE COBERTURA DO SOLO URBANO DERIVADAS DO WV-2 QUANTO AO NÍVEL DA LEGENDA DE CLASSIFICAÇÃO: ESTUDO DE CASO PARA UM SETOR DA UNICAMP, SP

David Guimarães Monteiro França

Dissertação de Mestrado do Curso de Pós-Graduação em Sensoriamento Remoto, orientada pela Dra. Cláudia Maria de Almeida, aprovada em 31 de agosto de 2016.

URL do documento original:
<http://urlib.net/8JMKD3MGP3W34P/3MAPN7E>

INPE
São José dos Campos
2017

PUBLICADO POR:

Instituto Nacional de Pesquisas Espaciais - INPE
Gabinete do Diretor (GB)
Serviço de Informação e Documentação (SID)
Caixa Postal 515 - CEP 12.245-970
São José dos Campos - SP - Brasil
Tel.:(012) 3208-6923/6921
Fax: (012) 3208-6919
E-mail: pubtc@inpe.br

COMISSÃO DO CONSELHO DE EDITORAÇÃO E PRESERVAÇÃO DA PRODUÇÃO INTELECTUAL DO INPE (DE/DIR-544):
Presidente:
Maria do Carmo de Andrade Nono - Conselho de Pós-Graduação (CPG)
Membros:
Dr. Plínio Carlos Alvalá - Centro de Ciência do Sistema Terrestre (CST)
Dr. André de Castro Milone - Coordenação de Ciências Espaciais e Atmosféricas (CEA)
Dra. Carina de Barros Melo - Coordenação de Laboratórios Associados (CTE)
Dr. Evandro Marconi Rocco - Coordenação de Engenharia e Tecnologia Espacial (ETE)
Dr. Hermann Johann Heinrich Kux - Coordenação de Observação da Terra (OBT)
Dr. Marley Cavalcante de Lima Moscati - Centro de Previsão de Tempo e Estudos Climáticos (CPT)
Silvia Castro Marcelino - Serviço de Informação e Documentação (SID)
BIBLIOTECA DIGITAL:
Dr. Gerald Jean Francis Banon
Clayton Martins Pereira - Serviço de Informação e Documentação (SID)
REVISÃO E NORMALIZAÇÃO DOCUMENTÁRIA:
Simone Angélica Del Ducca Barbedo - Serviço de Informação e Documentação (SID)
Yolanda Ribeiro da Silva Souza - Serviço de Informação e Documentação (SID)
EDITORAÇÃO ELETRÔNICA:
Marcelo de Castro Pazos - Serviço de Informação e Documentação (SID)
André Luis Dias Fernandes - Serviço de Informação e Documentação (SID)

MINISTÉRIO DA CIÊNCIA, TECNOLOGIA, INOVAÇÕES E COMUNICAÇÕES
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

sid.inpe.br/mtc-m21b/2016/08.23.15.02-TDI

COMPARAÇÃO ENTRE CLASSIFICAÇÕES DE COBERTURA DO SOLO URBANO DERIVADAS DO WV-2 QUANTO AO NÍVEL DA LEGENDA DE CLASSIFICAÇÃO: ESTUDO DE CASO PARA UM SETOR DA UNICAMP, SP

David Guimarães Monteiro França

Dissertação de Mestrado do Curso de Pós-Graduação em Sensoriamento Remoto, orientada pela Dra. Cláudia Maria de Almeida, aprovada em 31 de agosto de 2016.

URL do documento original:
<http://urlib.net/8JMKD3MGP3W34P/3MAPN7E>

INPE
São José dos Campos
2017

Dados Internacionais de Catalogação na Publicação (CIP)

França, David Guimarães Monteiro.

F844c Comparação entre classificações de cobertura do solo urbano derivadas do WV-2 quanto ao nível da legenda de classificação: Estudo de caso para um setor da Unicamp, SP / David Guimarães Monteiro França. – São José dos Campos : INPE, 2017.

xxvi + 217 p. ; (sid.inpe.br/mtc-m21b/2016/08.23.15.02-TDI)

Dissertação (Mestrado em Sensoriamento Remoto) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2016.

Orientadora : Dra. Cláudia Maria de Almeida.

1. Sensoriamento remoto urbano. 2. GEOBIA. 3. Modelo digital de altura. 4. CART. 5. AdaBoost.M1. I.Título.

CDU 528.854

Esta obra foi licenciada sob uma Licença Creative Commons Atribuição-NãoComercial 3.0 Não Adaptada.

This work is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported License.

Aluno (a): *David Guimarães Monteiro França*

"COMPARAÇÃO ENTRE CLASSIFICAÇÕES DE COBERTURA DO SOLO URBANO DERIVADAS DO WV-2 QUANTO AO NÍVEL DA LEGENDA DE CLASSIFICAÇÃO: ESTUDO DE CASO PARA UM SETOR DA UNICAMP, SP"

Aprovado (a) pela Banca Examinadora em cumprimento ao requisito exigido para obtenção do Título de ***Mestre*** em

Sensoriamento Remoto

Dra. Cláudia Maria de Almeida

Presidente / Orientador(a) / INPE / São José dos Campos - SP

Dr. Hermann Johann Heinrich Kux

Membro da Banca / INPE / SJCampos - SP

Dr. Thales Sehn Körting

Membro da Banca / INPE / São José dos Campos - SP

Dr. Raul Queiroz Feitosa

Convidado(a) / PUC- RIO / Rio de Janeiro - RJ

Este trabalho foi aprovado por:

( ) maioria simples

(X) unanimidade

São José dos Campos, 31 de agosto de 2016

"Se você só fizer o que sabe, nunca será mais do que é agora."

- Mestre Shifu

A meus pais e irmão, Alexandre, Leila e Rafae, Ubuntu.

AGRADECIMENTOS

Agradeço aos amigos que tive o prazer de conhecer durante minha passagem na DSA-CPTEC do INPE de Cachoeira Paulista e na FATEC-GT. Especialmente à Gisele Mira, Mayra, Nelsinho, Tadeu, Daniel Villa, Fred e Simone Sievert, por todo incentivo em cursar o mestrado no SERE.

Agradeço a orientação técnica, profissional, pessoal e por todo tempo cedido pelos pesquisadores: Dr. Thales Körting, Dra. Elisabete Caria Moraes, Dra. Liana, Dr. Sidnei Sant'Anna, Dr. Rafael Santos, Dr. Camilo Rennó, Dr. Luiz Aragão, Dr. Hermann, Dr. Lênio, Dr. Flavio Ponzonni, Dra. Evlyn Novo, Dr. Eduardo Camargo, Dr. Leonardo Bins, Dr. Raul Feitosa, Dr. Antônio Miguel e todos aqueles que, quando saí pelos corredores, abriram suas portas quando bati.

A todo quadro de funcionários do SERE, especialmente à turma de 2014! Aline, Ana, Anielli, Bruna Pechini, Bruna Braga, Bruna Neves, Cesare, Daniel, David Costa, Diogo, Eder, Evelyn, Helô, Hugo, João Bosco, João Felipe, Laura, Omar, Rani, Rodolfo e Vanessa. A ordem é alfabética porque a sua maneira, todos vocês foram muito especiais para mim. Mas quem eu realmente tive o orgulho de conhecer foram os remoteiros "camp trip", em "ubas" e oxalá! Vocês sabem quem são. Fizeram parte da minha formação profissional e pessoal, me deram conselhos, compartilharam conquistas, angústias, afeto e união. Um obrigado especial para o Sacha, Pompeu e Marujo, pelas linhas de código citadas nos apêndices e para o Maurano da Divisão de Processamento de Imagens (DPI) pelo apoio nos pontos de validação dos mapas.

Agradeço à CAPES pela concessão da bolsa e também ao INPE-SJC, por ceder a infraestrutura. E quando falo em infraestrutura, acho que devo incluir os bolos e o café servido religiosamente pelas meninas da Perfect Clean: Ercília, Sandra e Cris. Toda equipe de segurança, em especial o Seu Orlando, Aragão, Carneiro, Dimas e Paulo, tanto pelo zelo, quanto pela prosa. E também à turma da lanchonete ADC, pelos quitutes e caprichos: Lú, Cassia, Cris, Carminha, Bru e Magrão.

Por fim, aos meus irmãos e amigos: Jesus, Fronza, Lia, Karinne, Rafael Le Masson, Fabíola, Salman, Chris Stephan, Alexandre Prado, Danilo Franco, Dani Goltz, Yoshi, Claudinho, Ronald e a toda turma do ex-Sigma, LTID e ASA. Muito obrigado!

RESUMO

Áreas urbanas possuem importância estratégica por abrigarem a imensa maioria da população brasileira, mas essas áreas são dinâmicas e demandam contínua atualização cartográfica. Tal necessidade posiciona o sensoriamento remoto orbital (SR) como ferramenta chave para o mapeamento e atualização dessas áreas, oferecendo cobertura sistemática e elevada qualidade geométrica. Mas a complexidade dos alvos urbanos é tema persistente na literatura e vai além das limitações espectrais. Neste contexto, este trabalho explorou as potencialidades e limitações do sensor WorldView-2 (WV-2) para a classificação fina de alvos urbanos, utilizando-se de algoritmos não-paramétricos e tendo como objeto de estudo uma área contendo grande diversidade de materiais de cobertura do solo, localizada em uma área-teste do campus da Universidade Estadual de Campinas (UNICAMP), São Paulo. De forma a cumprir essa tarefa, foram explorados métodos de mineração de dados, refinados pelo algoritmo híbrido de *boosting* (AdaBoost.M1) e conjugados a análise de imagens baseada em objetos geográficos (GEOBIA). Foram analisados três algoritmos de árvore de decisão: C4.5 (J48), C5.0 (See5) e CART em dois níveis de legenda de classificação, com e sem o auxílio de um modelo digital de altura (MDA). Como resultado, foram gerados 12 cenários de classificação para uma mesma imagem da área de estudo. O algoritmo de *boosting* se mostrou eficiente nos 12 experimentos de classificação, auxiliando as árvores de decisão a atingirem valores de exatidão temática de até 0,74 para 12 classes de legenda (Nível I) e 0,72 para 36 classes do Nível II. No que diz respeito à análise comparativa dos algoritmos, o CART com auxílio do MDA gerou o modelo de árvore de decisão com o menor número de regras dentre os algoritmos aqui apresentados. Por fim, conclui-se que a metodologia apresentada se mostrou satisfatória na mineração de dados visando à classificação supervisionada baseada em objeto da imagem WV-2 referente à área de estudo analisada neste trabalho.

Palavras-chave: Sensoriamento remoto urbano. GEOBIA. mineração de dados. KDD. C4.5. J48. C5.0. See5. CART. Modelo Digital de Altura. AdaBoost.M1. boosting. WEKA. eCognition. QGIS. [R].

COMPARISON OF LAND COVER CLASSIFICATIONS OF WV-2 IMAGES REGARDING THE LEGEND LEVEL: A CASE STUDY IN THE UNICAMP-SP SECTOR

ABSTRACT

Urban areas have strategic importance for hosting the vast majority of the population, but these areas are dynamic and require continuous and updated mapping. As a consequence of this need, remote sensing (RS) plays a key role in mapping and updating these areas, providing systematic coverage and high geometrical quality, but the complexity of urban targets is a persistent theme in the literature and goes beyond spectral limitations. In this context, this work explored comparatively the potential and constraints of the WorldView-2 sensor (WV-2) in the fine classification of urban targets by means of non-parametric algorithms and having as a study object an area containing a large diversity of land cover materials, located at the campus of Campinas State University (UNICAMP), belonging to Sao Paulo state, Brazil. Data mining methods where explored in order to fulfill this task, refined by the boosting algorithm (AdaBoost.M1) they where also coupled to the geographical object-based image analysis (GEOBIA). Three algorithms were also analyzed for two legend levels: C4.5 (J48), C5.0 (See5) and CART, with and without the help of a digital height model (DHM). As a result, 12 classification scenarios where generated for a single image of the study area. The boosting algorithm displayed its efficiency in the 12 experiments by means of assisting the decision trees to reach thematic accuracy values of up to 0.74 to the first 12 legend classes (Level I) and up to 0.72 to 36 classes of the Level II. Regarding the comparative analysis among the algorithms, the CART algorithm with the help of the DHM, generated the decision tree model with the smallest number of rules among them. Lastly, we can conclude that the methodology hereby presented proved to be satisfactory in the data mining targeted to the object-based supervised classification of the WV-2 image related to the study area analyzed in this work.

Keywords: Urban remote sensing. GEOBIA. data mining. KDD. C4.5. J48. C5.0. See5. CART. digital height model. AdaBoost.M1. boosting. WEKA. eCognition. QGIS. [R].

LISTA DE FIGURAS

Pág.

2.1 Posições das bandas espectrais do sensor WV-2. 6
2.2 Resoluções espectral e temporal do sensor WorldView-2, lançado em 2009, em relação a outros sistemas imageadores. 7
2.3 Ilustração conceitual dos limiares de proporção sugeridos para realização de fusão de imagens. 10
2.4 Produtos derivados de modelos digitais de elevação. 11
2.5 Ilustração conceitual das distorções nas imagens de sensores orbitais e sua variação após o procedimento de ortorretificação. 12
2.6 Abstração hierárquica dos alvos na segmentação multirresolução. 14

3.1 WV-2 Composição RGB (5,3,2) da área de estudo em Campinas-SP. . . . 19
3.2 Fluxograma de etapas metodológicas do trabalho. 21
3.3 ENVI 5.1: Ferramenta para conversão dos níveis de cinza da imagem em valores de radiância (L). 23
3.4 ENVI 5.1: interface e parâmetros da ferramenta FLAASH. 24
3.5 Sub-campos do FLAASH para definição de propriedades multiespectrais do sensor WV-2. 25
3.6 Interpolação matemática aplicados em fusão de imagens: Convolução Bicúbica (A), Bilinear (B) e Interpolação por Vizinho mais Próximo (C). . 26
3.7 Ilustração conceitual das etapas metodológicas de confecção do MDA. . . 27
3.8 Modelo Digital de Altura proveniente da subtração do MDT pelo MDS. . 28
3.9 Ferramenta de segmentação e parâmetros de entrada no eCognition. . . . 29
3.10 Experimento realizado durante o trabalho, que apresenta as quadras de tênis com segmentos bem definidos (A) e segmentos enfatizados em amarelo representando mistura espectral (B). 30
3.11 Área de estudo segmentada por um fator de escala de 104, gerando um total de 21101 segmentos. 32
3.12 Histograma de segmentos amostrais coletados por classe. 39
3.13 Esquema de coleta de amostra para alvos urbanos que apresentam comportamento não-lambertiano (A) e galpão com segmentos coletados como amostras para classe Galvalume A, demarcados em vermelho (B). 40
3.14 Espectro das amostras coletadas para classe Galvalume A. 40
3.15 Totalidade dos segmentos amostrados para as 36 classes do Nível de Legenda de classificação II. 41

3.16 Interface de entrada dos atributos e amostras para os modelos C4.5 e CART na plataforma WEKA, para mineração de dados e confecção das árvores. 42
3.17 Modelo de árvore de decisão concebido por meio de C4.5. 49
3.18 Comportamento da seleção de atributos por algoritmos greedy. 53
3.19 Conceitualização das regras de particionamento dos dados sobre um conjunto de atributos bidimensionais sendo recursivamente definidas. 54
3.20 Resultado do sorteio de pontos de validação estratificados por classe. . . 63

4.1 Alvos de interesse sensíveis ao processo de segmentação. 67
4.2 Legenda das classes temáticas no Nível de Legenda II para os diagramas de espalhamento. 72
4.3 Comportamento das amostras para valores de NDVI, apresentando *outliers* em azul e um recorte sugerindo *zoom* para a área de interesse em vermelho. 73
4.4 Valores das amostras para os atributos de NDVI 1 e 2. 74
4.5 Diagrama de espalhamento dos atributos Max_Pixel_MDA x BRT. . 75
4.6 Contraste dos alvos de alta reflectância no Azul (B2) e Vermelho (B5), explorados pelo índice de Leonardi (2010) adaptado para Vermelho Limítrofe (B6). 76
4.7 Atributo proveniente da oitava componente principal versus o atributo de brilho, que juntos apresentam potencial de separabilidade entre algumas classes. 77
4.8 Atributos resultantes da análise por componentes principais sobre as oito bandas multiespectrais do sensor WV-2. 80
4.9 Acertos (a), erros (b) e concentração por área(c), observados para os pontos de validação, em relação à classificação C4.5 Nível II com MDA. . 83
4.10 Reflectância para as classes de Asfalto e Concreto e interferência das janelas de absorção atmosférica por vapor d´água (H_2O). 89
4.11 Transecto perfilado sobre a área de estudo (a), descrevendo os níveis de cinza do MDA em uma curva de tendência em vermelho (b). 97
4.12 Reflectância espectral mensurada no eixo Y para as amostras de Cerâmica (a) e Solo Exposto (b) com base nas oito bandas do WV-2. 100
4.13 Reflectância mensurada em laboratório (a) versus a reflectância normalizada para os valores da imagem WV-2 (b), ambas para a classe Cerâmica. 101
4.14 Representantes da classe de Vegetação Arbórea que potencialmente contribuem para confusão com as classes de Vegetação Herbácea e Solo Exposto. 102
4.15 Similaridade espectral entre alvos de vegetação seca e solo exposto. . . . 103

4.16 Comparação da reflectância (eixo Y) para as amostras de Solo (A) e Vegetação Herbácea (B). 103
4.17 Contraste entre as velhas telhas de fibrocimento amianto e as novas de CRFS, dispostos em um recorte do transecto na área de estudo. 105
4.18 Classificações que apresentaram os maiores coeficientes Kappa com e sem o MDA para o Nível de Legenda I. 107
4.19 Classificações acusadas como mais distintas através do teste Z com $p < 0,05$ para o Nível de Legenda II. 108
4.20 Mapa de divergência entre os classificadores dos Níveis I e II. 109

D.1 Imagens temáticas classificadas para as legendas do Nível I, que foram omitidas por apresentarem índices Kappa inferiores aos dados justificados no capítulo de discussão dos resultados. 181
D.2 Imagens temáticas classificadas para o Nível de Legenda I que foram omitidas da discussão de resultado por apresentarem alta similaridade estatística no teste Z. 182

E.1 Galvalume Ondulado. 183
E.2 Galvalume Trapezoidal. 184
E.3 Galvalume Trapezoidal Tamanho 40 com Pintura Eletrostática Branca. . 185
E.4 Aço Galvanizado Trapezoidal Tamanho 25. 186
E.5 Galvalume tipo sanduíche com pintura eletrostática branca. 187
E.6 Telha de fibrocimento CRFS. 188
E.7 Telha cerâmica resinada tipo portuguesa. 189
E.8 Quartzito amarelo (Pedra mineira). 190
E.9 Telha de concreto. 191

LISTA DE TABELAS

Pág.

3.1 Disposição hierárquica das legendas de classificação. 33
3.2 Tipologia e espectro das classes do Nível de Legenda II. 34
3.3 Atributos customizados não-nativos da plataforma eCognition 43
3.4 Atributos utilizados como descritores relativos à forma, espectro, propriedades estatísticas e classes dos segmentos da imagem. 46
3.5 Sistema de decisão por regras de classificação para C5.0. 52
3.6 Descrição em pseudocódigo para um algoritmo conceitual de boosting. . 58

4.1 Histórico sequencial dos parâmetros utilizados na segmentação. 66
4.2 Ranking dos 10 atributos mais utilizados por cada classificador no Nível de Legenda II, com e sem o MDA. 69
4.3 Atributos descartados pelos classificadores nas seis primeiras árvores dos algoritmos C4.5, C5.0 e CART com e sem atributos do MDA. 70
4.4 Atributos mais utilizados pelas árvores de decisão para as classificações do Nível de Legenda II. 71
4.5 Atributos menos utilizados pelas árvores de decisão para as classificações do Nível de Legenda II. 71
4.6 Percentual de contribuição das bandas espectrais do WV-2 na composição da Análise por Principais Componentes. 78
4.7 Resultados da avaliação dos coeficientes de concordância *Kappa*. 81
4.8 Distribuição por classe dos pontos de validação. 84
4.9 Matriz de confusão do algoritmo C4.5 fazendo uso de MDA, modelo que apresentou o maior valor de *Kappa* (0,74) para o Nível de Legenda I. . . 85
4.10 Matriz de confusão CART, modelo que apresentou o maior *Kappa* (0,74) para o Nível de Legenda I sem a necessidade de um MDA. 86
4.11 Acurácia do Produtor, Usuário e Kappa-classe individual para todos os cenários de classificação do Nível de Legenda I. 88
4.12 Matriz de confusão para classificação por CART com MDA para o Nível de Legenda II. 92
4.13 Matriz de confusão para classificação por C5.0 com MDA para o Nível de Legenda II. 93
4.14 Acurácia do Produtor, Usuário e Kappa-classe individual para todos os cenários de classificação do Nível de Legenda II. 95

C.1 Matriz de confusão para classificação por CART Nível I com MDA. . . . 173

C.2 Matriz de confusão para classificação por C4.5 Nível I sem MDA. 174
C.3 Matriz de confusão para classificação por C5.0 Nível I com MDA. 174
C.4 Matriz de confusão para classificação por C5.0 Nível I sem MDA. 175
C.5 Matriz de confusão para classificação por CART Nível II sem MDA. . . . 176
C.6 Matriz de confusão para classificação por C4.5 Nível II com MDA. 177
C.7 Matriz de confusão para classificação por C4.5 Nível II sem MDA. 178
C.8 Matriz de confusão para classificação por C5.0 Nível II sem MDA. 179

LISTA DE ABREVIATURAS E SIGLAS

ASD	–	*Analytical Spectral Devices*
BRDF	–	*Bidirectional Reflectance Distribution Function* (Função de Distribuição da Reflectância Bidirecional)
CART	–	*Classification And Regression Tree* (Árvore de Classificação e Regressão)
CSV	–	*Comma Separated Values* (Valores Separados por Vírgula)
DCBD	–	Descoberta de Conhecimento em Bancos de Dados
DLT	–	*Direct Linear Transformation* (Transformação Linear Direta)
DM	–	*Data Mining* (Prospecção de Dados)
ES	–	*Expert System* (Sistema Especialista)
FLAASH	–	*Fast Line-of-sight Atmospheric Analysis of Hypercubes*
GCP	–	*Ground Control Point* (Pontos de Controle no Terreno)
GEOBIA	–	*Geographic Object-Based Image Analysis* (Análise de Imagens Baseada em Objetos Geográficos)
GIS	–	*Geographic Information System*
IBGE	–	Instituto Brasileiro de Geografia e Estatística
IC	–	Inteligência Computacional
IFOV	–	*Instantaneous Field of View* (Campo de Visada Instantâneo)
KDD	–	*Knowledge Discovery in Databases*
LARAD	–	Laboratório de Radiometria do DPA/INPE
MDA	–	Modelo Digital de Altura
MDE	–	Modelos Digitais de Elevação
MDS	–	Modelo Digital de Superfície
MDSn	–	Modelo Digital de Superfície Normalizado
MDT	–	Modelo Digital de Terreno
ML	–	*Machine Learning* (Aprendizado de Máquina)
MNDWI	–	*Modified Normalized Difference Water Index* (Índice Modificado de Diferença Normalizada para Água)
MSI	–	Bandas Multiespectrais
NC	–	Nível de cinza
NDBI	–	*Normalized Difference Built-Up Index* (Índice de Diferença Normalizada para Áreas Construídas)
NDSI	–	*Normalized Difference Snow Index*

		(Índice de Diferença Normalizada para Neve)
NDVI	–	*Normalized Difference Vegetation Index*
		Índice de Vegetação de Diferença Normalizada
NN	–	*Nearest Neighbor*
		(Vizinho Mais Próximo)
PAN	–	Banda Pancromática
PDI	–	Processamento Digital de Imagens
REM	–	Radiacao Eletromagnética
RFM	–	*Rational Function Model*
		(Modelo de Funções Racionais)
RPC	–	*Rational Polynomial Coefficients*
		(Coeficientes Polinomiais Racionais)
SIG	–	Sistema de Informação Geográfica
SR	–	Sensoriamento Remoto
TOA	–	*Top of Atmosphere*
		(Topo da Atmosfera)
UNICAMP	–	Universidade Estadual de Campinas
UNPD	–	*United Nations Population Division*
		(Divisão de População das Nações Unidas)
WEKA	–	*Waikato Environment for Knowledge Analysis*
		(Ambiente Waikato para Análise de Conhecimento)
WV-BI	–	*WorldView Built-Up Index*
		(Índice de Área Construída para WorldView)
WV-NHFD	–	*WorldView Non-Homogeneous Feature Difference*
		(Diferença de Características Não-Homogeneas para WorldView)
WV-NII	–	*WorldView New Iron Index*
		(Novo Índice de Ferro para WorldView)
WV-SI	–	*WorldView Soil Index*
		(Índice de Solos para WorldView)

SUMÁRIO

Pág.

1 INTRODUÇÃO **1**
1.1 Objetivo Geral 2
1.2 Objetivos Específicos 3

2 FUNDAMENTAÇÃO TEÓRICA **5**
2.1 Sensoriamento Remoto Urbano de Alta Resolução Espacial 5
2.1.1 O Satélite WorldView-2 6
2.2 Processamento de Imagens Digitais 8
2.2.1 Pré-Processamento 8
2.2.1.1 Correção Atmosférica 8
2.2.1.2 Conversão de Nível de Cinza (NC) em Reflectância 8
2.2.1.3 Fusão de Imagens 9
2.2.2 Geração de Modelos Digitais de Elevação (MDE) a partir de um Estéreo-Par 10
2.2.3 Ortorretificação 11
2.2.4 Segmentação Multirresolução 13
2.2.5 Métodos de Classificação de Imagens 14
2.2.5.1 Sistemas Especialistas e GEOBIA 14
2.2.5.2 Mineração de Dados 15
2.2.5.3 A Maldição da Dimensionalidade 16
2.2.6 Análise da Acurácia da Classificação 16

3 MATERIAL E MÉTODOS **19**
3.1 Área de Estudo 19
3.2 Material 20
3.3 Metodologia de Trabalho 21
3.4 Conversão de NC para Reflectância e Correção Atmosférica 22
3.5 Fusão de Imagens 25
3.6 Modelo Digital de Altura 26
3.7 Segmentação 29
3.8 Classes, Tipologia e Amostragem 33
3.9 Classificação 42
3.10 Índices, Razões de Banda e Atributos 43

3.11 Algoritmos de Árvore de Decisão C4.5 / J48 48
3.12 Árvore de Decisão C5.0 / See5 51
3.13 Árvore de Classificação e Regressão - CART 53
3.14 Algoritmos híbridos e otimização de classificação: Boosting, Bagging e Blending 56
3.14.1 Boosting e AdaBoost.M1 57
3.15 Determinação da Acurácia de Produtos Cartográficos 59
3.15.1 Autocorrelação Espacial 60
3.15.2 Amostragem Aleatória Estratificada 61

4 RESULTADOS E DISCUSSÃO 65
4.1 Parâmetros da Segmentação 65
4.2 Comportamento e Escolha dos Atributos pelos Classificadores 68
4.3 Análise Exploratória dos Atributos 72
4.4 Análise de Acurácia da Classificação 81

5 CONCLUSÕES 111

REFERÊNCIAS BIBLIOGRÁFICAS 113

APÊNDICE A - GUIA DE NOMES DOS ATRIBUTOS TRUNCADOS PELO ECOGNITION PARA 12 CARACTERES 127

APÊNDICE B - ÁRVORES DE DECISÃO C4.5, C5.0 E CART COM E SEM MDA PARA AS LEGENDAS DO NÍVEL I E NÍVEL II 129
B.1 Árvore gerada por C4.5 Nível II com MDA 129
B.2 Árvore gerada por C5.0 Nível II com MDA 133
B.3 Árvore gerada por CART Nível II com MDA 138
B.4 Árvore gerada por CART Nível II sem MDA 140
B.5 Árvore gerada por C4.5 Nível II sem MDA 143
B.6 Árvore gerada por C5.0 Nível II sem MDA 147
B.7 Árvore gerada por CART Nível I com MDA 152
B.8 Árvore gerada por CART Nível I sem MDA 154
B.9 Árvore gerada por C4.5 Nível I com MDA 156
B.10 Árvore gerada por C4.5 Nível I sem MDA 160
B.11 Árvore gerada por C5.0 Nível I com MDA 164
B.12 Árvore gerada por C5.0 Nível I sem MDA 167

APÊNDICE C - MATRIZES DE CONFUSÃO C4.5, C5.0 E CART PARA NÍVEL I E NÍVEL II 173

APÊNDICE D - IMAGENS TEMÁTICAS CLASSIFICADAS 181

APÊNDICE E - EXPERIMENTOS REALIZADOS NO LARAD PARA O ESTUDO DO COMPORTAMENTO ESPECTRAL DE ALVOS URBANOS 183

APÊNDICE F - SCRIPT [R] PARA PROCESSAMENTO, CLASSIFICAÇÃO E VALIDAÇÃO DOS SEGMENTOS DA IMAGEM WV-2. 193

1 INTRODUÇÃO

Em nível global, o contingente populacional que reside em áreas urbanas é estimado em 52% pela Divisão de População da Organização das Nações Unidas, (*United Nations Population Division* - UNPD), equivalente a 3,6 bilhões de pessoas (UNPD, 2012). No Brasil, segundo levantamentos de pesquisas realizadas pelo Instituto Brasileiro de Geografia e Estatística (IBGE, 2013a), essa cifra se eleva a 84%.

Considerando que áreas urbanas possuem importância estratégica por abrigarem a imensa maioria da população brasileira, essas regiões, dinâmicas por excelência, demandam contínua atualização cartográfica. Portanto, existe a necessidade dos gestores públicos e tomadores de decisão da esfera privada de deter conhecimento aprofundado acerca das características físicas e de ocupação desses espaços, para fins de planejamento.

Nesse contexto, o sensoriamento remoto (SR) orbital é uma ferramenta chave para o mapeamento e atualização cartográfica do perímetro urbano, oferecendo cobertura sistemática e elevada qualidade geométrica. Além dessas qualidades, citam-se também o imageamento em vários canais multiespectrais e o alto grau de detalhamento dos alvos propiciado pelos sensores de alta resolução espacial (THOMAS et al., 2003).

O incremento das resoluções espacial e espectral são recursos de suma importância no mapeamento de áreas urbanas, reconhecidamente complexas (HEROLD et al., 2002). No final da década de 1990, com o lançamento do primeiro satélite de alta resolução espacial, IKONOS (ou IKONOS-2, que sucedeu o IKONOS-1 após a falha em seu lançamento), surgem novas possibilidades para o mapeamento e classificação de superfícies urbanas, permitindo um detalhamento inusitado na discriminação de seus alvos.

Mas os sensores de alta resolução espacial que sucederam o IKONOS-2 mantiveram, em sua maior parte, as mesmas configurações técnicas, compreendendo quatro bandas multiespectrais nos canais do azul, verde, vermelho e infravermelho próximo. E apesar de esses novos sensores oferecerem elevada resolução temporal e radiométrica, o número de bandas permaneceu limitado em decorrência do *trade-off* entre resolução espacial e espectral (CHEN et al., 2008).

É somente em 2009, com o lançamento do WorldView-2 (WV-2), que se supera esta limitação. Fabricado pela DigitalGlobe, esse sistema sensor inovou ao introduzir quatro novas bandas multiespectrais, além das quatro bandas presentes nos senso-

res de alta resolução espacial precedentes, compreendendo o azul costeiro, amarelo, vermelho-limítrofe e infravermelho próximo-2. Entretanto, por questões judiciais, até fevereiro de 2015, os produtos comercialmente disponíveis eram reamostrados, tendo sua resolução espacial a nadir degradada de 0,46 para 0,50 m na banda pancromática, e de 1,84 para 2,00 m nas multiespectrais (UPDIKE; COMP, 2010).

Kumar e Miller (2006) sugerem que diversas classes presentes nas imagens de SR não apresentam distribuições gaussianas, e as técnicas necessárias para classificar os alvos que apresentam tais distribuições devem ser não-paramétricas, ou seja, devem ser capazes de lidar com distribuições desconhecidas.

No ano de 2008, Wu et al. (2008) incluiram os algoritmos de classificação não paramétrica CART e C4.5 entre os dez melhores algoritmos em mineração de dados. No mesmo ano, Quinlan (2008) criador do C4.5, também declarou a descontinuidade de seu suporte ao C4.5, afirmando que o mesmo fora superado pelo seu sucessor C5.0.

Buscando a otimização dos resultados de classificação, a literatura demonstra também que a utilização de modelos digitais de altura (MDA) pode contribuir diretamente para o discernimento de alvos nas imagens de SR (FAHSI et al., 2000; LEONARDI, 2010; PINHO, 2006a; SESNIE et al., 2008). Mas a geração de um MDA não é tarefa trivial, e sua confecção e validação em campo pode se tornar complexa e custosa (Van Niel et al., 2008).

Neste contexto, este trabalho se propõe a explorar as potencialidades e limitações do sensor WV-2 para a classificação fina de alvos urbanos, com e sem a inclusão de um MDA, utilizando-se de métodos não-paramétricos e tendo como objeto de estudo uma área contendo grande diversidade de materiais de cobertura do solo, localizada em uma área-teste do campus da Universidade Estadual de Campinas (UNICAMP), São Paulo.

1.1 Objetivo Geral

Este trabalho se propõe a avaliar comparativamente classificações de cobertura do solo urbano, utilizando as oito bandas multiespectrais do WV-2 em associação a um modelo digital de altura (MDA) derivado de um estéreo-par desse mesmo sensor. As análises comparativas serão realizadas por meio de três algoritmos não-paramétricos discriminativos (C4.5, C5.0 e CART), complementados por técnicas de *boosting*, visando compreender e quantificar a relação entre o conjunto de dados de entrada nos algoritmos e os resultados das suas classificações, segundo o seu nível de legenda (de-

talhamento). Finalmente, serão avaliados os limites do sensor WV-2 e dos métodos empregados para a discriminação de classes de cobertura do solo urbano.

1.2 Objetivos Específicos

Com a finalidade de alcançar o objetivo geral proposto na Seção 1.1, os seguintes objetivos específicos serão realizados ao longo do trabalho:

a) Explorar diferentes métodos não-paramétricos que correspondem ao estado da arte em classificação de imagens, aplicando-os à discriminação de alvos urbanos em imagem de alta resolução espacial;

b) Avaliar transversalmente o comportamento dos três métodos, identificando as diferenças na seleção de atributos e no desempenho da classificação frente às variações nos dados de entrada;

c) Explorar e discutir as peculiaridades e limitações dos classificadores segundo o discernimento de cada alvo, em função de suas respostas espectrais e com base nos coeficientes de acurácia, provenientes das matrizes de confusão.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 Sensoriamento Remoto Urbano de Alta Resolução Espacial

Para as aplicações urbanas, a literatura reconhece a importância da resolução espacial, em primeira instância, sem contudo desmerecer o papel estratégico da resolução espectral (PINHO et al., 2005; JENSEN, 2009; BLASCHKE et al., 2014). Isto ocorre porque a variedade espectral dos alvos urbanos é muito extensa, e suas propriedades, como textura e cor, tendem a ser marcadamente peculiares, podendo auxiliar ou comprometer seu discernimento.

Em função da melhora na capacidade de armazenamento de dados dos sensores de alta resolução espacial, foi possível incrementar a resolução radiométrica das imagens por eles adquiridas (CHEN et al., 2008). Esta resolução é determinada com base em um *trade-off* entre vários fatores, incluindo a qualidade da imagem, a informação que necessita ser extraída da mesma, além da capacidade de armazenamento do sistema sensor.

Existem restrições nas resoluções do SR orbital que seguem a lei do inverso do quadrado da distância (SLATER, 1980), sendo a principal dentre elas a relação sinal-ruído (do inglês *signal-to-noise ratio* - SNR). Essa relação força a escolha entre uma resolução espacial mais fina em troca da degradação da resolução espectral. Este *trade-off* ocorre porque, objetivando aumentar a resolução espacial de um sensor, o campo de visada instantânea (*Instantaneous Field of View* - IFOV) precisa ser reduzido, mas isso também reduz a quantidade de radiação eletromagnética (REM) que chega aos sensores orbitais em certas faixas do espectro, degradando a qualidade da sua resolução espectral. Algumas das formas de se contornar essas restrições são possíveis pela, fusão entre bandas, redução de altitude das plataformas orbitais ou até mesmo com o aumento da quantidade de fótos detectados pelos sensores através de modificações em seu tamanho físico (CHEN et al., 2008).

Os satélites de alta resolução espacial possuem a vantagem de apresentarem um tempo extremamente curto de revisita, oscilando em torno de 1,1 dias no caso do WV-2. Isto foi possibilitado pela altitude relativamente baixa do mesmo, que é de 770 km, aliada à sua elevada velocidade, sendo capaz de concluir uma órbita completa em torno do globo terrestre em 100 minutos (DIGITALGLOBE, 2009).

Jensen e Cowen (1999) afirmam que, no que tange à resolução temporal, embora haja exceções, grande parte dos usuários de imagens de SR urbano se satisfaz com

informações de cobertura do solo provenientes de revisitas anuais, justificando que não é comum a ocorrência de mudanças significativas na cobertura do solo urbano em períodos de tempo inferiores a este. Já para o caso das resoluções radiométricas, espectrais e espaciais, a habilidade de identificar, discernir e separar os alvos é diretamente influenciada pelo nível de refinamento de tais resoluções.

2.1.1 O Satélite WorldView-2

O satélite comercial de imageamento de alta resolução WV-2 foi lançado em 8 de outubro de 2009 da Base da Força Aérea de *Vandenberg* e foi declarado operacional com compatibilidade máxima de seus instrumentos em 4 de Janeiro de 2010. O satélite possui uma órbita semicircular heliossíncrona e fornece dados com 11 bits de resolução radiométrica em nove bandas espectrais, descritas na Figura 2.1.

O sensor imageador do WV-2 é um instrumento do tipo *push-broom*, sendo a largura da faixa imageadora (*swath*) de 16,4 km. Na data de seu lançamento, o WV-2 inovou ao oferecer uma resolução espacial mais fina que os sistemas disponíveis na época de seu lançamento (Figura 2.2), menor tempo de revisita (1,1 dias) e quatro novas bandas espectrais (azul costeiro [*coastal*], amarelo [*yellow*], vermelho-limítrofe [*red-edge*] e infravermelho próximo-2 [*near infrared-2*]). Posteriormente, estas bandas se mostraram essenciais para estudos em diversas aplicações, como batimetria (BRAMANTE et al., 2013), classificação de espécies de vegetação (LATIF et al., 2012), e no estudo de áreas urbanas (RIBEIRO, 2010; PINHO et al., 2012; HAMEDIANFAR et al., 2014).

Figura 2.1 - Posições das bandas espectrais do sensor WV-2.

Fonte: Adaptado de DigitalGlobe (2009).

Figura 2.2 - Resoluções espectral e temporal do sensor WorldView-2, lançado em 2009, em relação a outros sistemas imageadores.

Fonte: Adaptado de Jensen (2009).

2.2 Processamento de Imagens Digitais

2.2.1 Pré-Processamento

2.2.1.1 Correção Atmosférica

A atmosfera interage com a radiação eletromagnética, provocando significativas alterações no fluxo radiante proveniente do alvo, o que ocasiona mudanças na sua propagação, denominadas espalhamento e absorção. A combinação desses dois efeitos é denominada atenuação atmosférica (SLATER, 1980).

Latorre et al. (2002) citaram as principais consequências desses efeitos sobre a imagem, como a diminuição do brilho da superfície em regiões espectrais específicas e a presença de névoa, com perda de nitidez, nas regiões dos menores comprimentos de onda. Os autores também sugerem que no processo da análise de imagens deve-se considerar as características espectrais do sensor e as condições atmosféricas da data e hora de aquisição desses dados.

Atualmente, há na literatura diversos métodos para correção atmosférica aplicáveis a dados multiespectrais. Para uma melhor compreensão dos mesmos, pode-se dividi-los em métodos alternativos e métodos físicos. Os métodos alternativos são aqueles que utilizam informações intrínsecas das imagens, ou seja, utilizam números digitais (níveis de cinza) existentes em pixels que são representativos de feições específicas. São correções que não dependem de parâmetros atmosféricos ou de dados da superfície, e sim das informações da própria imagem para determinar a contribuição da atmosfera em cada banda espectral. Uma de suas limitações consiste na exigência de alvos específicos nas imagens analisadas (LATORRE et al., 2002). Nos métodos físicos, é fundamental o conhecimento das propriedades ópticas da atmosfera bem como do processo de interação da radiação com a atmosfera e com a superfície. Estes métodos são fundamentados na teoria da transferência radiativa (CHANDRASEKHAR, 1960), a fim de retratar a complexidade da atmosfera, que varia com o tipo de modelo a ser utilizado e o resultado final a que se propõe.

2.2.1.2 Conversão de Nível de Cinza (NC) em Reflectância

A habilidade para se detectar e quantificar mudanças no ambiente terrestre depende dos sensores, que são capazes de fornecer medidas calibradas e consistentes (de precisão e acurácia conhecidas) de aspectos da superfície da Terra ao longo do tempo. A correta interpretação da informação advinda de séries globais e multitemporais de produtos de sensoriamento remoto requer a destreza para se discriminar artefatos

de reais mudanças sob monitoramento na superfície terrena (ROY et al., 2002). A caracterização e calibração radiométricas são um pré-requisito para a criação de dados científicos de alta qualidade, e consequentemente, produtos de melhor nível deles derivados. O cálculo da radiância espectral no sensor é uma etapa fundamental para a conversão de dados de imagem de diferentes sensores em uma escala radiométrica comum com significado físico (CHANDER et al., 2009).

Uma redução na variabilidade entre cenas pode ser obtida pela conversão da radiância espectral no sensor pela reflectância exoatmosférica ou do topo da atmosfera (*Top of Atmosphere* - TOA), também conhecida como albedo planetário. Ao se comparar imagens de diferentes sensores, há três vantagens evidentes no uso da reflectância TOA ao invés da radiância espectral no sensor. Primeiro, ela remove o efeito cosseno de diferentes ângulos zenitais solares devido à diferença de tempo entre as aquisições dos dados. Segundo, a reflectância TOA compensa os diferentes valores de irradiância solar exoatmosférica resultantes de diferenças das bandas espectrais. Terceiro, a reflectância TOA corrige a variação na distância Terra-Sol entre diferentes datas de aquisição dos dados. Essas variações podem ser geográficas e/ou temporalmente significativas. Por fim, é válido salientar a importância da conversão de valores brutos (nível de cinza - NC) em valores físicos antes da determinação de índices de vegetação ou de água, como será o caso da pesquisa a ser desenvolvida. A reflectância TOA é calculada de acordo com a Equação 2.1.

$$\rho_\lambda = \frac{\pi \cdot L_\lambda \cdot d^2}{E_{sun_\lambda} \cdot cos\theta_s}, \tag{2.1}$$

em que ρ_λ = reflectância TOA [adimensional]; π = constante matemática; L_λ = radiância espectral na abertura do sensor $[Wm^{-2}sr^{-1}\mu m^{-1}]$, d = distância Terra-Sol [unidades astronômicas]; E_{sun_λ} = irradiância solar exoatmosférica média $[Wm^{-2}\mu m^{-1}]$; θ_s = ângulo solar zenital [graus]. Cabe observar que o cosseno do ângulo solar zenital é igual ao seno do ângulo de elevação solar (CHANDER et al., 2009).

2.2.1.3 Fusão de Imagens

Dados multiespectrais de sensores comerciais, como as imagens do WV-2, usualmente são fornecidas com a banda pancromática (PAN) de alta resolução espacial e também suas respectivas bandas multiespectrais (MSI ou MULTI). A questão é que algumas feições específicas são visíveis apenas na banda PAN, mas carecem da

informação espectral da MSI. A fim de se obter maior proveito dos dados de sensoriamento remoto, é sugerida a aplicação de técnicas de fusão de imagens (SUN et al., 2014).

Zhang (2002) demonstra que, para evitar distorções e alcançar resultados ótimos, não se deve exceder a razão entre as resoluções espaciais das bandas MSI e PAN de 4:1 (Figura 2.3). De forma similar, Geomatica (2013) recomenda que essa razão não exceda 5:1.

Figura 2.3 - Ilustração conceitual dos limiares de proporção sugeridos para realização de fusão de imagens.

Fonte: Adaptado de Sun et al. (2014).

2.2.2 Geração de Modelos Digitais de Elevação (MDE) a partir de um Estéreo-Par

O Modelo Digital de Superfície (MDS) reproduz o perfil do solo, incluindo todos os objetos situados sobre ele (edificações, vegetação, torres, viadutos etc.). Uma alternativa rápida e de baixo custo financeiro para se gerar um MDS sem a necessidade de missões com sensores aerotransportados é a utilização de dados estereoscópicos digitais adquiridos por meio de sensores orbitais. A estereoscopia é viabilizada pela aquisição de múltiplas imagens sobrepostas em uma mesma linha de voo, que produzem o efeito da paralaxe estereoscópica. A paralaxe pode ser definida como o deslocamento aparente na posição de um objeto com respeito a uma estrutura de referência, causado pela alteração na posição do observador. Tal efeito resulta em

uma condição hiperestereoscópica, que permite a visualização do terreno em três dimensões (RAO et al., 1996; JENSEN, 2009).

Diversos produtos digitais podem ser obtidos através de um MDS, como o Modelo Digital de Terreno (MDT), que representa apenas a superfície do solo sem os objetos que se elevam a partir dele (Figura 2.4), e principalmente o Modelo Digital de Altura (MDA), que é o produto da subtração entre o MDS e o MDT. Assim: $MDS - MDT = MDA$. O MDA pode ser utilizado como um parâmetro de entrada para a classificação de cobertura do solo, aplicação de interesse deste trabalho.

Figura 2.4 - Produtos derivados de modelos digitais de elevação.

Fonte: O autor.

Toutin (2004) sugere que, embora as funções racionais bi e tri-dimensionais sejam majoritariamente utilizadas, sendo aconselhado sempre o uso de modelos físicos determinísticos e tridimensionais como a primeira opção sempre que disponíveis, não há um padrão geral para sua aplicação no SR, sendo inevitável para alguns sensores fazer uso dos modelos empíricos.

2.2.3 Ortorretificação

Dados de imagens adquiridos por sensores orbitais ou aerotransportados são afetados sistematicamente por erros de geometria induzidos pela plataforma do sensor, que introduz distorções do terreno quando o sensor não estiver diretamente a *nadir*. Isso ocorre pois as variações topográficas da superfície da Terra e a inclinação dos satélites ou sensores aerotransportados afetam a distância em que as características dispostas nas imagens aéreas ou de sensores orbitais são apresentadas. Conforme ilustrado na Figura 2.5, quanto maior a variação da topografia do terreno, maior

será a distorção inerente na imagem, sendo possível em alguns casos a ocorrência de distorções de centenas de metros (SIC, 2015).

Modelos matemáticos diversos foram criados nos últimos anos com o intuito de extrair informações tridimensionais, bem como ortorretificar imagens produzidas por sensores orbitais, isto é, eliminar distorções causadas principalmente pela posição do satélite e pela superfície física da Terra (relevo e curvatura), transformando assim a imagem de uma projeção cônica para uma projeção ortogonal, mantendo a constância da escala em toda a ortoimagem (JACOBSEN, 2002).

Figura 2.5 - Ilustração conceitual das distorções nas imagens de sensores orbitais e sua variação após o procedimento de ortorretificação.

Fonte: Adaptado de SIC (2015).

Alguns desses modelos matemáticos baseiam-se nas informações do sensor e em equações de colinearidade, nas quais se inserem parâmetros para a modelagem de erros e orientação interior ou calibração em órbita, ou ainda, adicionando-se informações orbitais, como dados de efemérides (BALTSAVIAS et al., 2001). Outra alternativa de modelagem consiste no uso de Coeficientes Polinomiais Racionais (*Rational Poly-*

nomial Coefficients - RPC), estimados apenas com dados de posição e atitude do sensor (GRODECKI; DIAL, 2001; TAO; HU, 2001), os quais são capazes de resgatar a geometria da imagem no instante de sua captura (PEDRO et al., 2007), através, entre outros, do Modelo de Funções Racionais (*Rational Function Model* - RFM).

Outros modelos, no entanto, não necessitam de informações do sensor ou orbitais. Exemplos neste sentido incluem: **i)** o uso de RPCs estimados por pontos de controle no terreno (*ground control points* - GCP) e não por aproximação do modelo do sensor (modelo físico); **ii)** modelos polinomiais não-racionais, cujos coeficientes também são estimados por GCP, os quais necessitam ser representativos da variabilidade planimétrica e altimétrica da cena; e, por fim, **iii)** a Transformação Linear Direta (*Direct Linear Transformation* - DLT), nas suas diferentes variações, que incluem transformações perspectivas e afins, de diferentes dimensões e que podem ser refinadas por meio da adição de correções de relevo (HATTORI; ONO, 2000; TAO; HU, 2001; VASSILOPOULOU et al., 2002).

2.2.4 Segmentação Multirresolução

A segmentação é objeto de estudo no campo do processamento digital de imagens (PDI) de longa data (BRICE; FENNEMA, 1970; ROBERTS, 1963). Haralick e Shapiro (1985) definem o processo de segmentação como a divisão de uma imagem em diversas partes, sendo condicionado para essa divisão que essas partes compartilhem propriedades homogêneas entre os elementos que a integram. Esses elementos podem ser pixels ou até mesmo outros segmentos, e as propriedades que devem compartilhar similaridades variam conforme a aplicação que se deseja dar para a imagem a ser segmentada.

Dentro das técnicas de SR, o processo de segmentação precede as etapas de classificação que se baseiam em regiões, e dentre os algoritmos de segmentação disponíveis, destaca-se o algoritmo de segmentação desenvolvido por Baatz e Schape (2000), uma vez que este apresenta diversos resultados satisfatórios de aplicação (ADDINK; KLEINHANS, 2008; BENZ et al., 2004; BURNETT; BLASCHKE, 2003; CAMARGO et al., 2011; PINHO et al., 2005; DURIEUX et al., 2008; HAY et al., 2001; HOFMANN et al., 2008; KUX et al., 2008; MALLINIS et al., 2008; WEMMERT et al., 2009). O algoritmo de segmentação proposto por Baatz e Schape (2000) é denominado pelos autores como polivalente em vista da sua multiplicidade de aplicações, e o mesmo se adapta à escala dos objetos presentes nas imagens conforme os parâmetros inseridos pelo usuário, o qual é mundialmente conhecido como segmentação multirresolução ou segmentação multiescala.

A principal vantagem da segmentação multirresolução é a sua capacidade de sintetizar objetos extraídos de diversos níveis de interpretação simultaneamente (Figura 2.6). Blaschke (2010) também sugere que nem sempre se faz necessária a aplicação da segmentação em diversos níveis, mas isto se torna uma necessidade quando os objetos presentes nas imagens se encontram em escalas e formatos muito heterogêneos, o que ocorre de forma muito frequente em imagens de áreas urbanas.

Figura 2.6 - Abstração hierárquica dos alvos na segmentação multirresolução.

Fonte: Adaptado de Blaschke et al. (2014).

2.2.5 Métodos de Classificação de Imagens

2.2.5.1 Sistemas Especialistas e GEOBIA

Um sistema especialista (*Expert System* - ES) pode ser tratado como um arcabouço de normas, regras e dados computacionais agregados de maneira que possam auxiliar em processos de tomada de decisão que, de outra forma, exigiriam a presença de um especialista para ser tomada (JENSEN, 2005). São especialistas, pois atuam em nichos/domínios específicos do conhecimento, como, por exemplo, a reprodução de métodos específicos de interpretação de imagens.

Blaschke et al. (2014) definem a classificação baseada em objetos geográficos (*GEographic Object-Based Image Analysis* - GEOBIA) como um novo paradigma de interpretação de imagens, constituindo-se em um sistema especialista, cuja vantagem principal reside na identificação de propriedades dos objetos presentes no domínio espacial, buscando emular o processo cognitivo humano de interpretar uma cena.

Classificadores paramétricos trabalham com a hipótese de que a função de distribuição de probabilidade estatística em uma imagem assume uma distribuição conhecida *a priori* (pondendo esta ser uma gaussiana ou chi quadrada entre outras). Em contra partida, Kumar e Miller (2006) sugerem que, de forma frequente a natureza das classes presentes em uma imagem de SR não pode ser discretizada de forma paramétrica por se tratarem de distribuições desconhecidas.

Classificadores não-paramétricos são ideais para distribuições desconhecidas porque estão além da noção de distribuições de probabilidade contínuas e métricas, podendo descrever problemas discretos que são resolvidos por métodos de reconhecimento de padrões baseados em regras (DUDA et al., 2000). De forma a identificar esses padrões, os sistemas não-paramétricos precisam ser refinados e treinados, e em geral, isso é feito com grandes volumes de dados que são obtidos via Mineração de Dados.

2.2.5.2 Mineração de Dados

A mineração ou prospecção de dados (do inglês *Data Mining* - DM) pode ser definida como o processo de identificar valores, padrões ou estruturas de interesse em um amplo conjunto de informações heterogêneas, também definido por Fayyad et al. (1996) como uma das subtarefas dentro do paradigma da Descoberta de Conhecimento em Bancos de Dados (DCBD ou do inglês *Knowledge Discovery in Databases* - KDD).

Não somente é possível, como também desejável a aplicação de métodos de KDD em imagens de SR. Atualmente, o volume crescente de dados de SR inviabiliza a mineração manual dos mesmos, tornando crucial o emprego de métodos automatizados para a extração de informações e detecção de padrões de mudança.

A necessidade de automatizar esses processos posiciona a KDD na interseção de áreas distintas, como a Inteligência Computacional (IC), estatística e visualização de dados. Na área de IC, particularmente, existem alguns tipos de técnicas e algoritmos que são utilizados em larga escala na fase de mineração, tais como as Redes Neurais, os Sistemas Especialistas, a Lógica *Fuzzy*, os Sistemas Híbridos e os Algoritmos

Genéticos.

Mais informações sobre os métodos não-paramétricos de mineração e classificação de imagens supracitados serão explicitadas nas Seções 3.11 (C4.5), 3.12 (C5.0) e 3.13 (CART), bem como as plataformas de mineração e classificação que disponibilizam a implementação desses algoritmos, WEKA 3.6 (*Waikato Environment for Knowledge Analysis*) (HALL et al., 2009), eCognition (TRIMBLE, 2011) e [R] (R Core Team, 2016).

2.2.5.3 A Maldição da Dimensionalidade

A origem do termo "Maldição da Dimensionalidade" (do inglês *Curse of Dimensionality*) é atribuída a Bellman (1957) e diz respeito a um fenômeno não exclusivo dos métodos de aprendizado de máquina, mas recorrente em diversas aplicações (análise combinatória, mineração de dados, bancos de dados e áreas afins) que fazem uso de dados multidimensionais.

O fenômeno trata de casos em que o aumento da dimensionalidade gera volumes tão massivos de dados, que a informação útil proveniente das amostras pode se tornar esparsa, ocasionando em aumento da incerteza nas estimativas dos testes estatísticos sobre esses dados. Outro problema decorrente da multidimensionalidade é a falha na identificação de grupos que apresentem similaridades (ou pertençam a mesma classe), visto que em dados multidimensionais os objetos tendem a parecer mais distantes e não correlacionados.

Também denominado de Efeito (HUGHES, 1968) em aprendizado de máquina, esse problema envolve algoritmos com a tarefa de aprender distribuições desconhecidas (não-normais) através de um conjunto limitado de amostras de alta dimensionalidade. Quando o número de amostras de treinamento é fixado em um valor específico, o poder de previsão do algoritmo diminui conforme a dimensionalidade aumenta. Mas essas limitações variam de algoritmo para algoritmo. Alguns algoritmos conseguem lidar com distribuições em espaços não euclidianos (que fogem às três dimensões) melhor do que outros.

2.2.6 Análise da Acurácia da Classificação

Para a validação de mapas temáticos de classificação, Cohen (1960) propõe o coeficiente de concordância *Kappa*, que se baseia na comparação de uma imagem de referência e na análise da acurácia esperada de uma classificação relativa.

Em contrapartida, Pontius e Millones (2011) afirmam que, embora o coeficiente

Kappa seja o método mais amplamente utilizado para os propósitos de validação na literatura de SR, esse coeficiente está equivocado e dá vazão para muitos erros, propondo, então, a utilização de matrizes de confusão simples que considerem a discordância entre as classes e a distribuição das mesmas no espaço.

Mais informações sobre os métodos de validação pelo índice de concordância *Kappa* e as matrizes de confusão serão discutidas na Seção 3.15, referente à determinação da acurácia dos métodos empregados na classificação das imagens.

3 MATERIAL E MÉTODOS

3.1 Área de Estudo

A área-teste deste estudo é um setor do Campus da Universidade Estadual de Campinas (UNICAMP), município de Campinas - São Paulo, localizada nas coordenadas 22°54'3"S, 47°3'26"W (Figura 3.1). Com altitude de 685 m, o município de Campinas conta com uma superfície total estimada em 795 km^2 (IBGE, 2013b), sendo aproximadamente 238 km^2 de cobertura urbana.

Figura 3.1 - WV-2 Composição RGB (5,3,2) da área de estudo em Campinas-SP.

Fonte: Adaptado de Satmap (2014).

Além da disponibilidade de dados, inlcuindo dado de relevo na forma de um MDA cedido por Lacerda (2016), outra característica pertinente para escolha desse local é a heterogeneidade de materiais presentes na região, como: telhados cerâmicos e coberturas metálicas, um grande lago nas proximidades do campus, quadras esportivas de material sintético, piscinas e vegetação diversificada.

3.2 Material

Apresentam-se a seguir o material auxiliar, equipamentos, dados e softwares utilizados neste trabalho:

- FieldSpec PRO Full-Range (FSP-2500P) da fabricante norte-americana *Analytical Spectral Devices* (ASD), para aquisição de dados espectrais em ambiente de campo e laboratório.

- Duas cenas WV-2 datadas de 21/06/2010 e 02/07/2010, com visadas de 19,93°e 6,75°off-nadir, respectivamente, com ângulo de elevação solar de 38,37°e 0% de cobertura de nuvem em ambas as imagens.

- Modelo Digital de Altura ortorretificado e cedido por Lacerda (2016).

- ENVI 5.1 - Para calibração radiométrica dos níveis de cinza e posterior correção atmosférica através do módulo FLAASH.

- WEKA 3-7-13 (*Waikato Environment for Knowledge Analysis*) - Plataforma de distribuição livre e código aberto, concebida para análise de dados através de métodos de aprendizado de máquina.

- eCognition 8.7 - Ambiente de desenvolvimento e análise de imagens baseada em objeto.

- [R] 3.2.4 - Linguagem de programação voltada para análise gráfica e estatística através de métodos computacionais.

- RStudio 0.99.671 - Ambiente integrado de desenvolvimento (IDE) para linguagem [R].

- QGIS 2.8.8 LTR - Sistema de informação geográfica (SIG) de código aberto e distribuição livre.

3.3 Metodologia de Trabalho

Figura 3.2 - Fluxograma de etapas metodológicas do trabalho.

Fonte: Produção do autor.

3.4 Conversão de NC para Reflectância e Correção Atmosférica

O *software* utilizado para aplicação dos procedimentos de correção atmosférica (Seção 2.2.1.1) e fusão de bandas é o ENVI 5.1, por este contar com uma implementação do algoritmo *Fast Line-of-sight Atmospheric Analysis of Hypercubes* - FLAASH para correção atmosférica. Uma característica em comum para a aplicação da correção atmosférica e dos modelos de transferência radiativa citados na Seção 2.2.1.2 é a necessidade dos dados de radiância de superfície como a principal variável de entrada, o que, para o caso do sensor WV-2, é calculado através da Equação 3.1.

$$L_{\lambda_{pixel,banda}} = \frac{K_{banda} \cdot q_{pixel,banda}}{\Delta_{\lambda_{banda}}}, \tag{3.1}$$

em que $L_{\lambda_{pixel,banda}}$ são os pixels de radiância espectral no topo da atmosfera em unidades de $[\frac{W}{m^2 \cdot sr \cdot \mu m}]$, K_{banda} é o fator absoluto de calibração radiométrica de uma dada banda, $q_{pixel,banda}$ são os pixels radiometricamente corrigidos da imagem, e $\Delta_{\lambda_{banda}}$ é a largura efetiva em $[\mu m]$ de uma dada banda (UPDIKE; COMP, 2010).

A mesma equação também é utilizada pela plataforma ENVI 5.1, que dispõe da ferramenta Radometric Calibration (Figura 3.3) para esta finalidade.

Figura 3.3 - ENVI 5.1: Ferramenta para conversão dos níveis de cinza da imagem em valores de radiância (L).

Fonte: Produção do autor.

As configurações da Figura 3.3 estão definidas para uma posterior correção atmosférica, mas o especialista também pode converter a imagem de entrada diretamente para valores não corrigidos de reflectância no topo da atmosfera ρ_λ através da Equação 3.2

$$\rho_\lambda = \frac{\pi \cdot L_\lambda \cdot d^2}{E_{sol_\lambda} \cdot sen\theta}, \tag{3.2}$$

em que L_λ é a radiância em $[\frac{W}{m^2 \cdot sr \cdot \mu m}]$, d é a distância Terra-Sol em unidades astronômicas, E_{sol_λ} é a irradiância solar em unidades de $[\frac{W}{m^2 \cdot \mu m}]$ e θ é a elevação do ângulo solar zenital.

Uma observação pertinente é que a conversão direta para ρ_λ só é possível quando o dado de imagem proveniente do sensor também acompanha o arquivo de metadados

(extensão *.TIL no caso do WV-2) contendo as seguintes variáveis declaradas: *gain*, *offset*, *solar irradiance*, *sun elevation* e *acquisition time*. Mesmo na correção atmosférica através do FLAASH (Figura 3.4), as informações do arquivo de metadados também se fazem necessárias.

Figura 3.4 - ENVI 5.1: interface e parâmetros da ferramenta FLAASH.

Fonte: Produção do autor.

A correção atmosférica por FLAASH também considera parâmetros para estimativa de névoa-seca (*haze*) e densidade de aerossóis em suspensão na data de aquisição da imagem (Figura 3.5). O método utilizado para essa estimativa é chamado de *Kaufman-Tanre* (KAUFMAN et al., 1997) e faz uso das bandas do Azul (Banda 1 do WV-2) e Vermelho (Banda 5) para estimar a influência dos aerossóis em alvos escuros identificados na superfície do solo, utilizando-se das medições da influência de aerossóis nessas bandas, para estimar os valores mais próximos da ρ no nível do solo.

Figura 3.5 - Sub-campos do FLAASH para definição de propriedades multiespectrais do sensor WV-2.

Fonte: Produção do autor.

3.5 Fusão de Imagens

Este trabalho fez uso da combinação de imagens digitais provenientes das bandas MSI e PAN do WV-2 através de método de fusão (pansharpening), utilizando o interpolador por vizinho mais próximo (*Nearest Neighbor* - NN) (SUN et al., 2014). Métodos de fusão como Gram Schmidt são muito robustos, mas o delineamento das bordas da imagem pode sofrer uma suavização indesejada quando o processo de interpolação escolhido pelo usuário é bilinear ou bicúbico (Figura 3.6), levando à mistura espectral dos pixels interpolados. É de interesse que seja mantida a máxima fidelidade espectral dos alvos, dada a necessidade de se enfatizar a separabilidade espectral entre estes.

Figura 3.6 - Interpolação matemática aplicados em fusão de imagens: Convolução Bicúbica (A), Bilinear (B) e Interpolação por Vizinho mais Próximo (C).

Fonte: Adaptado de Kidner et al. (1999).

O procedimento de confecção do Modelo Digital de Altura (Seção 3.6) independe da fusão de imagens, isso ocorre porque apenas as bandas pancromáticas das imagens envolvidas no processo serão utilizadas.

3.6 Modelo Digital de Altura

A metodologia apresentada nesta seção foi originalmente documentada por Lacerda (2016), que cedeu o raster do MDA ortorretificado para este trabalho. Os atributos do Modelo Digital de Superfície (MDS) e o Modelo Digital de Terreno (MDT) estão relacionados ao comportamento do terreno e por isso foram gerados uma única vez. O MDS foi criado antes da ortorretificação na ferramenta OrthoEngine. O MDT, que representa por meio de uma grade regular de cotas altimétricas a superfície do terreno propriamente dita, foi gerado no software QGIS 2.8.3 por meio de uma interpolação entre as coordenadas dos pontos de controle obtidos em campo. O atributo Modelo Digital de Altura (MDA), também chamado por alguns autores de Modelo Digital de Superfície Normalizado (MDSn), é resultante da subtração algébrica entre os modelos digitais de superfície e de terreno (Figura 3.9).

Figura 3.7 - Ilustração conceitual das etapas metodológicas de confecção do MDA.

Fonte: Adaptado de Lacerda (2016).

Para se obter o MDA, é preciso gerar o MDS (Modelo Digital de Superfície), modelo que é obtido por meio de um estéreo-par, duas imagens que façam sobreposição da área de interesse, tomadas a partir de pontos de vista diferentes, que resulta em um modelo que considera tanto as variações do terreno, quanto os elementos presentes sobre ele. Também é necessário gerar o MDT (Modelo Digital do Terreno) por meio da interpolação entre as coordenadas de pontos obtidos em campo, o qual considera apenas as variações do terreno, desconsiderando qualquer elemento que se eleve acima da superfície do terreno. Assim, a subtração entre o MDS e o MDT produziram o MDA (Figura 3.8).

Figura 3.8 - Modelo Digital de Altura proveniente da subtração do MDT pelo MDS.

Fonte: Produção do autor.

3.7 Segmentação

Após o tratamento de todos os dados WV-2, os arquivos de imagem foram importados para a plataforma eCognition 8.7 e empilhados em uma única pseudo-imagem de 17 bandas, em que as oito primeiras bandas são as MSI fusionadas, o MDA representado pela banda 9 e as bandas de 10 até 17 sendo as componentes principais, calculadas na plataforma ENVI 5.1. Inicia-se então, o procedimento de segmentação através da ferramenta de segmentação multirresolução (BAATZ; SCHAPE, 2000) apresentada na Figura 3.9 junto aos seus parâmetros de entrada.

Figura 3.9 - Ferramenta de segmentação e parâmetros de entrada no eCognition.

Fonte: Produção do autor.

Embora a segmentação represente apenas uma fração de todo o trabalho empregado em uma classificação baseada em objeto, ela também exerce grande influência no resultado final, pois os segmentos resultantes servirão de base para as próximas etapas do trabalho.

De forma a primar por resultados ótimos após a classificação, alguns cuidados devem ser tomados. Dentre eles, é digno de nota a necessidade de se isolar as classes de interesse em segmentos distintos, caso contrário, os segmentos resultantes do processo apresentarão uma mistura espectral com as demais classes (Figura 3.10), reduzindo a separabilidade entre as amostras que serão coletadas e propagando erros para os algoritmos responsáveis pela mineração e classificação dos demais segmentos.

Figura 3.10 - Experimento realizado durante o trabalho, que apresenta as quadras de tênis com segmentos bem definidos (A) e segmentos enfatizados em amarelo representando mistura espectral (B).

Fonte: Produção do autor.

Dentro da plataforma eCognition Developer 8.7, o algoritmo de segmentação multirresolução permite a inserção de diversos parâmetros de entrada, como compacidade, suavidade, forma, cor, fator de escala e pesos diferenciados para cada uma das bandas da imagem. Os valores ótimos para cada um desses parâmetros foram estimados heuristicamente, até serem obtidos resultados satisfatórios para os segmentos, apresentados na Seção 4.1. O mesmo se aplica ao parâmetro de peso, que foi inicialmente definido como 1 para todas as bandas, e posteriormente, as possíveis combinações de pesos foram avaliadas conforme sua influência sobre os segmentos mais sensíveis. Cabe observar que imagens de SR de alta resolução em áreas urbanas apresentam alvos de formas complexas e comportamento espectral heterogêneo, para tanto, uma forma de otimizar a segmentação por multirresolução é não restringir os valores de Forma (*Shape*) e Compacidade (*Compactness*), mantendo ambos em 0 (zero). Isso ocorre porque altos valores do parâmetro de Forma constróem segmentos que desconsideram a informação espectral dos pixels, quando, em SR, a informação espectral deveria ser priorizada.

O mesmo vale para o parâmetro de Compacidade, uma vez que áreas urbanas apresentam edificações e demais alvos de formas complexas, restringir os segmentos para formas compactas pode hipersegmentar desnecessariamente a imagem ao invés de unir diversas áreas que pertençam a um mesmo alvo não compacto, como por exemplo, uma rodovia de asfalto ou solo exposto.

É possível, ainda, a utilização de métodos semi-automáticos para ajustes dos parâmetros da segmentação como o *Estimation of Scale Parameter* - ESP ou *Spectral Parameter Tunner* - SPT (ACHANCCARAY et al., 2014).

O resultado final da segmentação é apresentado na Figura 3.11 e os parâmetros utilizados para sua confecção são posteriormente discutidos na Seção 4.1.

Figura 3.11 - Área de estudo segmentada por um fator de escala de 104, gerando um total de 21101 segmentos.

Fonte: Produção do autor.

3.8 Classes, Tipologia e Amostragem

A disposição das classes, apresentada na Tabela 3.1, segue dois níveis de legenda, em que o primeiro nível (Nível I) é uma abstração das classes presentes na imagem, agrupadas em suas respectivas superclasses semânticas, e o segundo nível (Nível II) rompe a abstração dos grupos, apresentando as classes individualmente, conforme a metodologia proposta por Lacerda (2016).

Tabela 3.1 - Disposição hierárquica das legendas de classificação.

Nível I (12 Classes)	Nível II (36 Classes)
Solo Exposto	Solo Exposto
Piscina	Piscina de Azulejo Piscina de Vinil
Lago	Lago
Vegetação Arbórea	Vegetação Arbórea
Vegetação Herbácea	Vegetação Herbácea
Pavimentação Não Viária	Pedra Mineira Carvão Coque
Pavimentação Viária	Asfalto Asfalto Pintado de Branco Asfalto Pintado de Vermelho Bloquete
Materiais Mistos (Pavimentação e Cobertura)	Argamassa de Cimento Argamassa de Cimento Pintado de Vermelho Argamassa de Cimento Pintada de Verde Concreto Pintado de Amarelo Concreto Pintado de Vermelho Concreto Concreto Impermeabilizado
e Coberturas Diversas	Vidro Aramado Fibra de Vidro Pastilha Esmaltada Fibrocimento CRFS Fibrocimento com Amianto Policarbonato Acrílico Plástico Lona
Coberturas Metálicas	Aço Galvanizado Brilhante Aço Galvanizado com Ferrugem Aço Galvanizado Fosco Galvalume A (Branco) Galvalume B Galvalume C
Coberturas Cerâmicas	Coberturas Cerâmicas
Sombra	Sombra

Fonte: Adaptada de Lacerda (2016).

A tipologia de cada classe presente na imagem, disposta na Tabela 3.2 tem a função de identificar os alvos que se tornaram as legendas da Tabela 3.1, não obstante a necessidade de serem identificadas visualmente. Também é de interesse para o SR discernir o seu comportamento espectral, apresentado para todas as trinta e seis classes do segundo nível de legenda, omitindo as legendas do primeiro nível, pois essas são agrupamentos abstratos do Nível II.

Tabela 3.2 - Tipologia e espectro das classes do Nível de Legenda II.

Tipologia / ID	**Classe**	**Reflectância (ρ) nas bandas WV-2**
01	Aço Galvanizado Brilhante	
02	Aço Galvanizado com Ferrugem	
03	Aço Galvanizado Fosco	
04	Acrílico	
05	Argamassa de Cimento	
06	Argamassa de Cimento Pintado de Verde	
07	Argamassa de Cimento Pintado de Vermelho	

(Continua)

Tabela 3.2 - Continuação

Tipologia / ID	**Classe**	**Reflectância (ρ) nas bandas WV-2**
08	Asfalto	
09	Asfalto Pintado de Branco	
10	Asfalto Pintado de Vermelho	
11	Bloquete	
12	Carvão Coque	
13	Cerâmica	
14	Concreto	
15	Concreto Impermeabilizado	
16	Concreto Pintado de Amarelo	

(Continua)

Tabela 3.2 - Continuação

Tipologia / ID	**Classe**	**Reflectância (ρ) nas bandas WV-2**
17	Concreto Pintado de Vermelho	
18	Fibra de Vidro	
19	Fibrocimento com Amianto	
20	Fibrocimento CRFS	
21	Galvalume A	
22	Galvalume B	
23	Galvalume C	
24	Lago	
25	Lona	

(Continua)

Tabela 3.2 - Continuação

Tipologia / ID	**Classe**	**Reflectância (ρ) nas bandas WV-2**
26	Pastilha Esmaltada	
27	Piscina de Azulejo	
28	Piscina de Vinil	
29	Plástico	
30	Policarbonato	
31	Quartzito Amarelo (Pedra Mineira)	
32	Solo Exposto	
33	Sombra	
34	Vegetação Arbórea	

(Continua)

Tabela 3.2 - Continuação

Tipologia / ID	**Classe**	**Reflectância (ρ) nas bandas WV-2**
	Vegetação Herbácea	
	Vidro Aramado	

Fonte: Produção do autor.

Definidas as classes, torna-se possível iniciar o processo de coleta de amostras. Em seu artigo de revisão, Congalton (1991) sugere abordagens para seleção de amostras que visam à viabilidade financeira (custos de campo e coleta) e estatística (quantidade significativa de amostras), tanto para modelos de coleta baseados em pixels, quanto para segmentos. O método aqui empregado carece da sofisticação de Congalton e apresenta algumas restrições, mas toma o artigo como referência em seus princípios, primando por obter uma estatística robusta e representativa. A principal restrição diz respeito à quantidade de amostras idealizadas para cada classe, que foi de aproximadamente 30 segmentos para as classes menores e 75 para as maiores, mas algumas classes não dispunham de tantas amostras, então, materiais mais escassos na imagem como lona ou plásticos tiverem seus segmentos integralmente coletados como amostra para suas classes, conforme ilustra o histograma da Figura 3.12.

Figura 3.12 - Histograma de segmentos amostrais coletados por classe.

Fonte: Produção do autor.

Tipicamente, a coleta de amostras também enfrenta um problema característico do SR, que diz respeito à função de distribuição da reflectância bidirecional (*bi-directional reflectance distribution function* - BRDF) (MILTON, 1987). A radiação eletromagnética (REM) refletida por um alvo varia em zênite e azimute, logo, para medir a sua reflectância (ρ), as medidas precisam ser feitas para todos os ângulos possíveis da fonte (Sol) e do sensor (*fieldspecPRO*). E embora existam dispositivos como o goniômetro, a tarefa de tirar medidas para todos os cenários possíveis de incidência da REM se torna pouco prática em campo e menos ainda para áreas urbanas.

Herold et al. (2004) afirmam que o comportamento não lambertiano de materiais urbanos aumenta conforme as medidas se aproximam do *nadir*. E, ao passo que alguns alvos apresentam comportamentos de espalhamento menos homogêneos conforme o ângulo mensurado (exemplos são concreto ou alumínio), outros alvos tendem a se comportar de forma mais lambertiana, como asfalto e telhados cerâmicos. Em suma, alvos urbanos são reconhecidamente complexos e peculiares em se tratando de seu domínio espectral.

Tais considerações devem ser feitas, pois a coleta de amostras está diretamente relacionada à qualidade espectral de seus pares. Este trabalho não foi exceção, e o esquema de coleta das amostras para a classe Galvalume A, no telhado do galpão do curso de Educação Física da UNICAMP, está esquematizado na Figura 3.13.

Figura 3.13 - Esquema de coleta de amostra para alvos urbanos que apresentam comportamento não-lambertiano (A) e galpão com segmentos coletados como amostras para classe Galvalume A, demarcados em vermelho (B).

Fonte: Produção do autor.

A Figura 3.14 apresenta o conceito da variabilidade espectral dos cinquenta e sete segmentos amostrados para classe Galvalume A para ilustrar a influência da BRDF, presente em diversos alvos sintéticos em ambientes urbanos.

Figura 3.14 - Espectro das amostras coletadas para classe Galvalume A.

Fonte: Produção do autor.

De um total de 21.101 segmentos na imagem, a Figura 3.15 apresenta os 1.796 segmentos coletados como amostra (aprox. 8% da imagem), para as 36 classes, uma média de 40 segmentos por classe, desconsiderando-se 400 amostras da classe Lago.

Figura 3.15 - Totalidade dos segmentos amostrados para as 36 classes do Nível de Legenda de classificação II.

Fonte: Produção do autor.

3.9 Classificação

Após os procedimentos de pré-processamento, segmentação e coleta de amostras, três algoritmos de classificação (C4.5, C5.0 e CART) serão aplicados nos quatro conjuntos de dados de entrada:

- WV-2 com MDA para o Nível I (3 modelos);
- WV-2 sem MDA para o Nível I (3 modelos);
- WV-2 com MDA para o Nível II de Legenda (3 modelos);
- WV-2 sem MDA para o Nível II de Legenda (3 modelos).

O término do processo deve gerar doze modelos de classificação de cobertura do solo. As seções subsequentes abordarão cada um dos métodos não paramétricos de classificação, a seleção dos índices, razões de banda e atributos selecionados para os modelos e também algumas informações referentes às plataformas utilizadas para esta finalidade, sendo o [R] utilizado para confecção do modelo C5.0 e o WEKA (Figura 3.16) utilizado nos modelos C4.5 e CART.

Figura 3.16 - Interface de entrada dos atributos e amostras para os modelos C4.5 e CART na plataforma WEKA, para mineração de dados e confecção das árvores.

Fonte: Produção do autor.

3.10 Índices, Razões de Banda e Atributos

Objetos presentes nas imagens (ou segmentos) apresentam características de forma, espectro, informações estatísticas, e essas propriedades são utilizadas como fonte de informação para definir parâmetros de inclusão ou exclusão utilizados por métodos e algoritmos de classificação de imagens. Para tanto, a Tabela 3.3 apresenta uma compilação dos índices e razões de banda implementados como atributos customizados na plataforma eCognition. Ao todo, são 17 atributos, os quais serão apresentados juntamente com as suas descrições individuais e respectivas referências.

Tabela 3.3 - Atributos customizados não-nativos da plataforma eCognition

Índice	**Descrição**
$BRT = \frac{\sum_{B1}^{B8}}{8}$	***Brightness*** (ou brilho) é dado pela soma ponderada de todos os níveis de cinza de todas as bandas espectrais da imagem segundo as métricas da plataforma eCognition. Seu objetivo principal é enfatizar contraste entre todos os alvos, fazendo os pixels claros ficarem mais claros e os escuros mais escuros (TRIMBLE, 2011)
$\frac{BRT + B2}{B5}$	Soma do brilho médio e a banda do azul (B2) divididos pela banda do vermelho (B5) para destacar telhados cerâmicos e solo exposto (LEONARDI, 2010).
$\frac{BRT + B2}{B6}$	Adaptação de índice (LEONARDI, 2010) que considera a banda do vermelho limítrofe (B6), presente no sensor WV-2.
$BRT - B7$	Subtração simples entre a banda sintética de brilho e a primeira banda do infravermelho próximo do WV-2. O objetivo é enfatizar o brilho de alvos que são ofuscados por materiais de alto brilho que já se destacam naturalmente.
$BRT - B8$	Adaptação da subtração simples entre a banda sintética de brilho e a segunda banda do infravermelho próximo do WV-2.

(Continua)

Tabela 3.3 - Continuação

Índice	**Descrição**
$\frac{B8 - B7}{B8 + B7}$	Índice de Diferença Normalizada para Áreas Construídas - *Normalized Difference Built-Up Index* (**NDBI**) adaptado para WV-2, originalmente concebido para sensores Landsat TM (ZHA et al., 2003).
$\frac{B3 - B8}{B3 + B8}$	Índice Modificado de Diferença Normalizada para Água *(Modified Normalized Difference Water Index* - **MNDWI**), também definido como Índice de Diferença Normalizada para Neve - *Normalized Difference Snow Index* - **NDSI**. Realça feições de água, enquanto suprime ruídos de área construída, vegetação e solo (RIGGS et al., 1994; XU, 2006).
$\frac{B3 - B7}{B3 + B7}$	Adaptação do **NDSI** (ou também **MNDWI**) considerando a banda do infravermelho próximo-1 no WV-2 (B7).
$\frac{B7 - B5}{B7 + B5}$	Índice de Diferença Normalizada para Vegetação - *Normalized Difference Vegetation Index* (**NDVI**). Identifica áreas de vegetação fotossinteticamente ativas e determina sua saúde vegetativa (Rouse Jr. et al., 1973).
$\frac{B8 - B5}{B8 + B5}$	**NDVI** considerando o infravermelho próximo-2 (B8) do WV-2.
$mNIR = \frac{B7 + B8}{2}$	Média simples entre as bandas do infravermelho próximo do WV-2 (B7 e B8).
$\frac{mNIR - B5}{mNIR + B5}$	Adaptação do **NDVI** que considera a média simples das bandas do infravermelho próximo.

(Continua)

Tabela 3.3 - Continuação

Índice	**Descrição**
$mSWIR = \frac{B6 + B7 + B8}{3}$	Média simples entre as bandas do infravermelho próximo 1, 2 e vermelho limítrofe, para incluir todo espectro de 700 nm até 1.043 nm.
$\frac{B3 \cdot B4}{B2 \cdot 1000}$	Novo Índice de Ferro para WorldView - *WorldView New Iron Index* (**WV-NII**). Índice utilizado em aplicações de geologia para identificar feições de óxido ferroso (WOLF, 2010).
$\frac{B1 - B6}{B1 + B6}$	Índice de Área Construída para WorldView - *WorldView Built-Up Index* (**WV-BI**). Índice de diferença normalizada entre as bandas do azul costeiro (B1) e vermelho limítrofe (B6), voltado para identificação de feições de áreas construídas (WOLF, 2010).
$\frac{B6 - B1}{B6 + B1}$	Diferença de Características não Homogêneas para WorldView - *WorldView Non-Homogeneous Feature Difference* (**WV-NHFD**). Índice desenvolvido para as bandas *Coastal* (B1) e *Red Edge* (B6) do sensor WV-2. Enfatiza feições de contraste com alvos naturais como: telhados, veículos e superfícies pavimentadas (WOLF, 2010).
$\frac{B3 - B4}{B3 + B4}$	Índice de Solos para WorldView - *WorldView Soil Index* (**WV-SI**). Índice de diferença normalizada que faz uso das bandas do verde (B3) e do amarelo (B4) para identificação de feições de argila e solos (WOLF, 2010).

Fonte: Produção do autor.

Além dos índices e razões de banda descritos na literatura e apresentados na Tabela 3.3, outros 79 atributos provenientes dos segmentos, como forma, geometria e suas propriedades espectrais e estatísticas, também foram computados e exportados pela plataforma eCognition para um arquivo de texto em formato CSV (Valores separados por vírgula - *Comma Separated Values*), apresentados e descritos individualmente na Tabela 3.4. Para mais detalhes individuais sobre os índices e suas respectivas

equações, é sugerida leitura do manual Trimble (2011).

Tabela 3.4 - Atributos utilizados como descritores relativos à forma, espectro, propriedades estatísticas e classes dos segmentos da imagem.

Atributo	**Descrição**
Assymetry (Assimetria)	Variando de 0 até 1, mede o quão irregular é o polígono em questão. Um índice de valor 1 representa valores ideais, em que a razão entre a largura e comprimento dos vértices principais do segmento são iguais.
Roundness (Circularidade)	Mede o enquadramento da forma do segmento em relação a um circulo traçado em seu redor. Enquanto 0 indica o enquadramento ideal, valores maiores que 0 representam o quanto o segmento se afasta da forma de um círculo. Não há um limite superior para os valores *roundness*.
Compactness (Compacidade)	Variando de 0 até 1, em que 1 é o valor ideal, descreve o quão compacto um objeto é, causando a impressão de uma pequena borda capaz de comportar todo volume do objeto. O índice de compacidade é dado pelo produto da altura pela largura, divididos pelo número de pixels dentro do objeto.
Density (Densidade)	Representa a distribuição dos pixels no espaço dentro do segmento. A forma mais densa é um quadrado, formas afiladas tendem a ser menos densas. A densidade é calculada baseada em uma matriz de covariância, através do número de pixels que formam o objeto dividido pelo raio do objeto.
Elliptic Fit (Enquadramento Eliptico)	Varia de 0 a 1, em que 1 representa um segmento de forma elíptica perfeita e 0 significa que o objeto não apresenta forma elíptica. Este índice é calculado através da projeção de uma elipse sobre o objeto com suas mesmas propriedades de altura e largura, comparando-os com a elipse.

(Continua)

Tabela 3.4 - Continuação

Atributo	**Descrição**
Rectangular Fit (Enquadramento Retangular)	Mede o enquadramento da forma do segmento em relação a um retângulo traçado em seu redor. Enquanto 0 indica não enquadramento, 1 indica o enquadramento ideal.
Border Index (Índice de Borda)	Descreve o quão serrilhada é a forma de um segmento. Quanto maior a irregularidade (serrilhamento), maior será o índice de borda, em que 1 representa valores ideais de uma forma retangular.
Shape Index (Índice de Forma)	Descreve a suavidade da borda de um segmento. Quanto mais suave for a borda deste segmento, MENOR será o seu valor de *Shape Index*. O índice é dado pela propriedade de tamanho da borda, dividido por quatro vezes a raiz quadrada de sua área.
Max pixel (Máximo Pixel[1])	Retorna o valor mais alto dentre todos os pixels dentro de um objeto/segmento.
Mean (Média[1])	Atributos deste grupo representam o primeiro momento estatístico (média) dos pixels dentro de um objeto de imagem (segmento).
ModeMinimun (Menor Moda[1])	Retorna o valor do pixel de maior reincidência, priorizando o valor mais baixo.
Min. pixel (Menor pixel[1])	Retorna o valor mais baixo dentre todos os pixels dentro de um objeto/segmento.
Class Name (Nome de Classe)	Único atributo nominal e mais importante descritor durante a etapa de mineração/classificação. É responsável por distinguir a que classe pertencem as amostras e a quais classes pertencerão os objetos classificados.

Fonte: Produção do autor.

[1]Atributos calculados 17 vezes para cada segmento, sendo oito para as bandas espectrais de B1 até B8, um para o MDA e mais oito para cada banda da componente principal PC1 até PC8.

Quatro principais atributos dizem respeito ao MDA, `modeMinimumda` (moda inferior dos valores dos pixels dentro de um segmento de MDA), `Max_pixel_mda` (dentre todos os pixels, aquele cujo valor é mais alto dentro do segmento), `mean_mda` (média simples entre todos os pixels do segmento) e `Min_pixel_mda` (dentre todos os pixels, aquele cujo valor é mais baixo dentro do segmento). Todos esses atributos derivam do MDA e seus nomes originais foram truncados para 12 caracteres, manifestando-se nas árvores de decisão C4.5, C5.0 e CART como: `modeminimu8`, `max_pixel_2`, `Mean_mda` e `min_pixel_9` respectivamente.

A possibilidade de incluir diversos atributos no conjunto de dados a ser minerado é promissora; as próprias ferramentas que simulam o processo cognitivo humano de identificar padrões em imagens nos ofertam centenas destes atributos (ou *features*), que vão da textura dentro dos objetos da imagem até a informação sobre quem são seus objetos vizinhos.

Mas embora o *big-data*[2] seja uma tendência dentro dos métodos de classificação de imagens em SR (LIU, 2015; MA et al., 2015), cuidados devem ser tomados para se avaliar a necessidade de cada um dos atributos a serem incluídos no conjunto de dados. A decisão sobre a inclusão dos atributos nos mineradores é de suma importância, pois grandes conjuntos de dados que venham a estar repletos de atributos com baixo poder descritivo podem ser responsáveis pela ocorrência de fenômenos como a "Maldição da Dimensionalidade" ou o "Efeito Hughes" (discutidos na Seção 2.2.5.3), nos algoritmos classificadores.

3.11 Algoritmos de Árvore de Decisão C4.5 / J48

Uma árvore de decisão consiste em uma estrutura em forma de fluxograma, em que cada nó interno representa um teste sobre um atributo, cada ramo representa o resultado do teste e cada nó-folha representa o rótulo de uma classe (decisão após o cômputo de todos os atributos). As trajetórias da raiz às folhas representam as regras de classificação (QUINLAN, 1987).

Segundo Quinlan (1993), o algoritmo C4.5 inicia a apuração com um conjunto de casos pertencentes a classes conhecidas. Esses casos são descritos por uma mistura qualquer de propriedades nominais e numéricas, e então, os casos são escrutinizados em busca de padrões que serão posteriormente expressos na forma de uma árvore de

[2] Liu (2015) define *big-data* em SR como uma coleções de dados de grande proporção e complexidade, dificultando o emprego de algoritmos tradicionais de processamento de dados. Os desafios do *big-data* incluem a aquisição, armazenamento, busca, compartilhamento, transferência, análise e visualização dos dados.

decisão ou regras condicionais do tipo `se : então (if : then)`, conforme apresentado no exemplo da Figura 3.17.

Figura 3.17 - Modelo de árvore de decisão concebido por meio de C4.5.

Fonte: Franca et al. (2015).

Algumas premissas do algoritmo devem ser satisfeitas no conjunto inicial dos dados de entrada (*inputs*), para que resultados ótimos sejam obtidos:

- Coerência na descrição dos valores de atributos: Os dados a serem analisados podem ser valores tanto numéricos quanto discretos, mas é imprescindível que estes estejam dispostos em uma estrutura padrão na base de dados.

- Pré-definição das classes: As categorias de classes disponíveis para alocar os dados analisados devem ser definidas antecipadamente, o que torna o C4.5 um algoritmo de classificação supervisionada, em contraste com outros algoritmos que podem criar seus próprios agrupamentos de classes.

- Dados suficientes: Quantidades estatisticamente significativas de casos devem ser fornecidas para alimentar o algoritmo. Valores exatos para uma quantidade ideal são difíceis de definir; a literatura sugere $N = 30$, mas também é pertinente que a quantidade de casos disponibilizados aumente

proporcionalmente à complexidade das classes a serem definidas no conjunto de dados (QUINLAN, 1993).

Após o polimento[3] dos dados, e tomando em conta as sugestões supracitadas, inicia-se o processo de construção da árvore de decisão. Se todos os casos no conjunto de dados pertencerem à mesma classe, a árvore se transforma em uma única folha, alojando todos os casos nessa dada categoria. Se este não for o caso, o processo de construção da árvore segue, calculando para cada atributo o potencial de informação que pode ser extraído em um teste com este atributo, denominado como ***entropia*** e baseado na probabilidade de cada caso ter um valor em particular dentro de um atributo. Também é calculado o ***ganho*** de informação que seria resultante de um teste no atributo (baseado na probabilidade de que cada caso, com um valor específico de um dado atributo, pertença exclusivamente a uma classe). Dependendo do critério de seleção, o algoritmo busca o melhor atributo para dar sequência à ramificação da árvore.

O critério utilizado para selecionar em quais ramificações a árvore irá crescer faz uso da entropia, uma forma de quantificar a desordem no conjunto de dados, calculado pela Equação 3.3,

$$Entropia(\hat{y}) = -\sum_{j-1}^{n} \frac{|y_j|}{|\hat{y}|} \log \frac{|y_j|}{|\hat{y}|}, \tag{3.3}$$

em uma iteração para todos os valores possíveis de $\hat{y}$. A entropia condicional é dada pela Equação 3.4:

$$Entropia(j|\hat{y}) = \frac{|y_j|}{|\hat{y}|} \log \frac{|y_j|}{|\hat{y}|}, \tag{3.4}$$

que por sua vez é utilizada no cálculo do ganho, pela Equação 3.5:

$$Ganho(\hat{y}|j) = Entropia(\hat{y} - Entropia(j|\hat{y})). \tag{3.5}$$

O objetivo é maximizar o ganho através da sua divisão pela entropia geral, considerado o critério de seleção $\hat{y}$ pelo valor j.

[3]Polimento caracteriza a etapa de remover de toda informação que possa ser considerada redundante ou desnecessária ao algoritmo minerador.

Concluída a construção da árvore, ocorre a poda, procedimento que visa ao aumento do poder de generalização da árvore através da redução do ajuste excessivo (*overfitting*), que pode tornar a árvore resultante demasiadamente grande, ao passo que uma árvore pequena é menos sensível aos dados de entrada e pode, por conseguinte, não capturar atributos importantes dos mesmos. O tamanho médio é preferido, uma vez que a árvore resultante neste caso não subestima e tampouco superestima os dados, sendo facilmente interpretada pelo usuário.

No ano de 2008, Wu et al. (2008) publicaram um artigo, o qual nomeou o algoritmo C4.5 como o primeiro colocado na lista de dez melhores algoritmos em mineração de dados, mas no mesmo ano (QUINLAN, 2008) declarou a descontinuidade de seu suporte ao C4.5, afirmando que o mesmo fora superado pelo seu sucessor C5.0, que será apresentado a seguir.

3.12 Árvore de Decisão C5.0 / See5

Denominado de C5.0 para sistemas Unix/Linux e See5 para sistemas Windows, este algoritmo é uma extensão do C4.5 (ou J48 em sua versão escrita em linguagem Java) desenvolvido por Quinlan (1993). Seu principal objetivo é identificar padrões e extrair informação de conjuntos de dados, mas também manter a compreensibilidade das regras utilizadas para este fim. As principais modificações de seu predecessor (C4.5) são:

- Inclusão nativa de técnicas de *boosting* no algoritmo C5.0 permitem a construção de múltiplos classificadores; tais técnicas induzem a um aumento significativo na acurácia preditiva do resultado final.
- Implementação de pesos diferenciados para erros de classificação permitem que algumas falhas sejam identificadas como sendo mais importante do que outras ao longo do processo de decisão.
- "Separar o joio do trigo" (do inglês *winnowing*) é uma técnica aplicada automaticamente nos dados de entrada que descarta os atributos de menor contribuição na etapa de pré-processamento. A função *winnowing* também pode ser desabilitada, caso o usuário tenha preferência pela inclusão de todos os atributos.

Da mesma forma que os demais algoritmos de árvore tratados neste trabalho (C4.5 e CART), o C5.0 também pode operar com decisões baseadas em regras. Segundo

Quinlan (1998), no entanto, uma nova funcionalidade foi adicionada de forma a simplificar a compreensão do processo de classificação, pois as regras poderiam ser examinadas e validadas separadamente, sem a necessidade de se levar em conta todo o contexto de classificação presente em uma árvore.

Para ilustrar o conceito, é apresentada na Tabela 3.5 uma sequência de regras de decisão com o intuito de classificar a renda de indivíduos com receitas acima ou abaixo de U$ 50.000. Os dados utilizados para composição das regras foram extraídos da base de dados do *US Census Bureau* e disponibilizados publicamente através do repositório de dados para ML e KDD da Universidade de Irvine, Califórnia (HETTICH; BAY, 1999).

Tabela 3.5 - Sistema de decisão por regras de classificação para C5.0.

Regra 1: (19425/1, lift 1.1) membros da família abaixo de 18 = ambos os pais presentes -> classe -50000 [1.000]
Regra 2: (1603/8, lift 1.1) educação = 5º ou 6º série perda de capital <= 1876 -> classe -50000 [0.994]
Regra 3: (32, lift 15.7) dividendos em ações > 33000 semanas trabalhadas no ano > 47 -> classe +50000 [0.971]
Regra 4: (132/12, lift 14.6) educação = graduação profissional dividendos em ações > 991 semanas trabalhadas no ano > 47 -> classe +50000 [0.903]

Fonte: Adaptada de Quinlan (1998).

A interpretação das regras é descrita através dos seguintes componentes:

- Um número arbitrário, que serve apenas para identificá-la;
- Estatísticas ***(n, lift x)*** ou ***(n/m, lift x)*** sumarizam o desempenho da regra. De forma similar às folhas de uma árvore de decisão, $\boldsymbol{n}$ é o número de casos do conjunto de treinamento que são contemplados pela regra e $\boldsymbol{m}$, quando presente, representa os casos que não pertencem à classe prevista

pela regra.

- A precisão estimada da regra é calculada pela razão $\frac{(n-m+1)}{(n+2)}$ e ***lift x*** é resultado da divisão da precisão estimada pela frequência relativa da classe predita no conjunto de treinamento.
- Uma ou mais condições que precisam ser satisfeitas para que a regra seja aplicável.
- Apenas uma classe prevista por regra.
- Intervalo de confiança da previsão que varia entre **0,000** e **1,000**.

Embora o algoritmo C5.0 apresente recursos mais modernos, suas origens estruturais são as mesmas de seu predecessor C4.5. Ambos são baseados em abordagens heurísticas sobre a entropia e o ganho através de simples decisões, tomadas por algoritmos gulosos (*greedy*), que tendem a escolher as soluções de maior ganho imediato, ilustrados na Figura 3.18.

Figura 3.18 - Comportamento da seleção de atributos por algoritmos greedy.

Fonte: Produção do autor.

3.13 Árvore de Classificação e Regressão - CART

Desenvolvido por Breiman et al. (1984), CART (do inglês *Classification and Regression Tree*) não foi o primeiro algoritmo de árvore de decisão introduzido nos

métodos de aprendizado de máquina (*Machine Learning* - ML), mas foi o primeiro a ser escrito com rigor analítico embasado em métodos estatísticos sofisticados e teoria probabilística.

Este algoritmo pode ser descrito como um método de aprendizado não paramétrico, com o propósito de prever variáveis dependentes através do processo de particionamento recursivo do espaço de variáveis de transição (Figura 3.19), sendo apresentado na forma de árvores de decisão binárias. Este método é capaz de lidar com variáveis dependentes contínuas (regressão) e categóricas (classificação).

Figura 3.19 - Conceitualização das regras de particionamento dos dados sobre um conjunto de atributos bidimensionais sendo recursivamente definidas.

Fonte: Produção do autor.

Uma grande vantagem do CART em relação aos demais métodos não lineares e não-paramétricos reside na possibilidade de operar tanto de forma supervisionada como não-supervisionada. Assim como o C4.5, as árvores de regressão geradas pelo CART são formadas por um aglomerado de regras que se baseiam diretamente no conjunto de dados de entrada que as modelam, aceitando também condições restritivas impostas de forma a limitar a profundidade alcançada por uma árvore. A profundidade da árvore determina a quantidade de ancestrais que um determinado nó possui, sendo que a raiz é adotada como profundidade zero. A profundidade fornece uma ideia do grau de complexidade que a árvore estimada apresenta.

As principais propriedades de CART são descritas abaixo:

- Atributos menos estáveis (que variam mais) dentro das variáveis tendem a ser selecionados primeiro, pois estes tendem a refinar melhor as regras de divisão (*split*) de uma árvore.
- Quando uma regra é criada, fazendo a divisão de um nó (*node*), este nó (que agora é um nó-filho ou *child node*) herda o conjunto de regras do seu nó superior (*parente node*), recursivamente.
- O processo de divisão cessa, quando o algoritmo CART não identifica mais ganhos na divisão da árvore, ou quando alguma regra pré-determinada força a interrupção da divisão.
- Alternativamente ao processo anterior, a árvore também pode prosseguir com o máximo de divisões possíveis e, posteriormente, sofrer um processo de poda (*prune*).
- Via de regra, cada ramificação (*branch*) da árvore terminará em um nó-folha (*terminal node*), que será sempre composto por um conjunto exclusivo de regras, alojando as classes finais quando um processo de classificação executar essa árvore.

Além das métricas de ganho e entropia já apresentadas nos algoritmos C4.5 e C5.0, e que também se aplicam aqui, a redução da variância é uma propriedade exclusiva do CART, que é sempre empregada na construção das árvores, mas apenas para os casos em que as variáveis de entrada são contínuas.

O algoritmo opera da seguinte forma: a redução da variância de um nó N é definida como a redução total da variância de uma dada variável x por conta de uma divisão nesse mesmo nó, expressa através da Equação 3.6.

$$Rv(N) = \frac{1}{|\mathbb{S}|^2} \sum_{i \in \mathbb{S}} \sum_{i \in \mathbb{S}} \frac{1}{2}(x_i - x_j)^2 - \left(\frac{1}{|\mathbb{S}|^2} \sum_{i \in \mathbb{S}_v} \sum_{j \in \mathbb{S}_v} \frac{1}{2}(x_i - x_j)^2 + \frac{1}{|\mathbb{S}|^2} \sum_{i \in \mathbb{S}_f} \sum_{j \in \mathbb{S}_f} \frac{1}{2}(x_i - x_j)^2 \right), \tag{3.6}$$

em que $\mathbb{S}$ é o conjunto de índices de amostra antes da divisão, $\mathbb{S}_v$ é o conjunto de índices da amostra para a qual o teste de divisão é verdadeiro e $\mathbb{S}_f$ é o conjunto de índices da amostra para a qual o teste de divisão é falso, respectivamente. E todas

as somatórias são estimativas de variância, no entanto, são descritas de forma a não fazer referência a suas respectivas médias.

Reduzir a variância do modelo é uma estratégia eficiente no que tange à redução da árvore de decisão, mas também é importante se ter em mente que até mesmo os casos dentro do primeiro desvio padrão correm o risco de serem excluídos. Uma forma de contornar esse problema é a construção de mais de uma regra com variância reduzida, ao invés de apenas uma regra forte. A construção dessas diversas regras menos robustas para atacar o mesmo problema é material de estudo no campo dos sistemas híbridos, dentro dos métodos de aprendizado de máquina (*Machine Learning* - ML), que será o tópico atacado na próxima seção.

3.14 Algoritmos híbridos e otimização de classificação: Boosting, Bagging e Blending

A habilidade de generalização é o conceito mais central em ML, uma vez que caracteriza o quão bem as informações e resultados aprendidos em uma amostra podem ser aplicados a um novo conjunto de dados. No cenário atual, Wu et al. (2008) demonstram os esforços científicos voltados à pesquisa de sistemas capazes de desenvolver uma habilidade de generalização mais robusta que os métodos tradicionais de aprendizado de máquina.

Choraś et al. (2009) afirmam que o paradigma mais bem-sucedido é o aprendizado híbrido (*ensemble learning*). Os métodos híbridos tentam construir diversas regras simples de separabilidade proveniente de algoritmos de aprendizado de base (*base learners*) e posteriormente acoplá-las a um modelo final mais robusto. Algoritmos de aprendizado de base costumam ser provenientes de dados de treinamento obtidos através de algoritmos que podem ser árvores de decisão, redes neurais ou outras formas de algoritmos de aprendizado de máquina. Como descrito em um jargão utilizado por (WU et al., 2008): "muitas mãos tornam o trabalho mais leve", que seria uma analogia à habilidade de generalização de um modelo híbrido se comparado aos métodos tradicionais. Os três métodos híbridos de otimização de classificação de maior destaque no campo de ML são apresentados:

- ***Bagging*** também chamado de agregação por *bootstrap*, é um método híbrido de otimização de classificação que cria amostras separadas do conjunto de treinamento, gerando também, um classificador para cada um desses novos conjuntos. O resultado desses múltiplos classificadores é então combinado através de recursos como média ponderada ou votos majoritá-

rios. O artifício de maior destaque do método *Bagging* é que o subconjunto amostral utilizado para construção da árvore de decisão é dividido em várias subpartes que são distintas entre si. Isso permite que cada novo classificador construído através desses subconjuntos tenha uma perspectiva diferente para abordar o mesmo problema, uma vez que os classificadores fracos que compõem o ensemble cometem erros em regiões distintas do espaço de atributos.

- ***Boosting*** é um método híbrido que inicia com um classificador sobre os dados de treinamento através de uma função de erro. É feita então, uma segunda classificação que atribui pesos diferenciados para as instâncias erroneamente classificadas pelo primeiro modelo. Esse processo continua até que um limite de modelos de classificação seja atingido ou não haja mais contribuição dos submodelos.

- ***Blending*** se difere dos modelos híbridos anteriores, porque ele agrega não apenas um, mas vários algoritmos para a mesma classificação. O método *Blending* faz uso de um meta-classificador preparado para aprender e manipular as previsões de cada um dos outros algoritmos, fazendo suas próprias previsões com base na observação dos resultados dos outros modelos. O método de Blending também é chamado de Stacking e no ambiente WEKA pode ser encontrado como Stacked Aggregation.

Não é parte do escopo deste trabalho discutir os detalhes sobre os três métodos híbridos previamente citados, e assim sendo, a próxima seção discute apenas o algoritmo AdaBoost.M1, que diz respeito ao método de *Boosting*, utilizado em todos os três classificadores (C4.5, C5.0 e CART) neste trabalho.

3.14.1 Boosting e AdaBoost.M1

Segundo Schapire (2003), *boosting* pode ser descrito como um método genérico aplicável para qualquer algoritmo de aprendizado de máquina que objetiva melhorar a acurácia de um modelo, e ele se baseia na ideia de que é mais fácil encontrar diversas regras "fracas" para separar as classes do que uma única regra "forte" de alta precisão. Para que isso ocorra, diversas "rodadas" (ou iterações) são feitas sobre os algoritmos que interagem com o algoritmo de *boosting*. A cada nova iteração, uma regra fraca é criada, e após diversas iterações, todas as regras são combinadas para criar uma única regra forte que, em teoria, deverá ser mais precisa que as regras utilizadas para sua construção. A Tabela 3.6 apresenta em linguagem conceitual o

desenvolvimento de boosting sendo aplicado em outro algoritmo.

Tabela 3.6 - Descrição em pseudocódigo para um algoritmo conceitual de boosting.

Definição das entradas (*inputs*):
D = Distribuição das amostras;
L = Algoritmos *base learners*; # C4.5, C5.0 ou CART
T = Número de iterações de *boosting*. # 10 iterações, para o caso deste trabalho

Processo:
1. $D_1 = D$ # Inicializa a distribuição
2. $Para\ t = 1, \cdots, T\ faça$:
3. $h_t = L(D_t)$; # Treinar o *base learner* sob a distribuição D_t
4. $\epsilon_t = Pr_{x \sim D_t, y} I[h_t(x) - y]$; # Medir o erro de h_t
5. $D_{t+1} = AjustarDistribuição(D_t, \epsilon_t)$
6. $FimPara$

Saída (*output*):
$H(x) = CombinaResultados(h_t(x))$ # Combinar regras $h_t(x)$ para formar regra forte $H(x)$

Fonte: Adaptado de Wu et al. (2008).

Segundo Wu et al. (2008), o primeiro esboço de um algoritmo capaz de melhorar a precisão (fazer um *boost*) nos resultados de outros algoritmos de ML (*base learners*) foi proposto por Schapire (1990). Mas este ainda apresentava uma falha conceitual por necessitar de informações sobre os erros cometidos pelos *base learners* antes de sua atual execução, o que não acontece na prática, em casos reais de classificação. Com a proposta de resolver essa questão de carência de informação, Freund e Schapire (1997) apresentam então o AdaBoost.

O algoritmo AdaBoost se destacou por treinar sequencialmente uma série de classificadores e depois combiná-los para compor uma previsão, o que foi chamado de *boosting* adaptativo, no qual os erros dos primeiros classificadores recebiam um maior foco dos classificadores que o sucederiam, processo descrito na Equação 3.5 e denominado de combinação aditiva ponderada.

$$H(x) = \sum_{t=1}^{T} \alpha_t h_t(x), \tag{3.7}$$

em que x é o conjunto amostral, α_t são os pesos utilizados na ponderação aditiva e h_t são as hipóteses combinadas para gerar a regra $H(x)$.

Mas o AdaBoost se limita a problemas binários, fazendo modelos de previsão para apenas duas classes, deficiência que abriu caminho para sua extensão, o AdaBoost.M1, algoritmo de *boosting* adaptativo multiclasses, documentado por Freund e Schapire (1997).

A extensão M1 do AdaBoost sofre apenas uma modificação sutil de sua versão original. Sua principal diferença é a substituição do erro para casos binários $|h_t(x_i) - y_i|$ por $[\![h_t(x_i) \neq y_i]\!]$ na linha 4 da Tabela 3.6 e o input de conjuntos de múltiplos casos $D = \langle (x_1, y_1), \cdots, (x_n, y_n) \rangle$ ao invés de $D = (1, \cdots, N)$.

3.15 Determinação da Acurácia de Produtos Cartográficos

O coeficiente de concordância *Kappa* é construído a partir de uma matriz de erros, na qual são expressos os erros de omissão, ou seja, amostras que não foram classificadas de acordo com as classes de referência, e os erros de comissão, correspondentes a amostras de referência classificadas erroneamente como pertencentes a outras classes. A partir das matrizes, são calculados os seguintes índices: **(a)** exatidão global - relação entre o número de amostras classificadas corretamente sobre o número total de amostras de referência; **(b)** acurácia do produtor - relativa aos erros de omissão, a qual representa a relação entre o número de amostras classificadas corretamente da classe k e o número total de amostras de referência da classe k, **(c)** acurácia do usuário - referente aos erros de comissão, a qual corresponde à relação entre o número de amostras classificadas corretamente da classe k e o número total de amostras classificadas da classe k; **(d)** Kappa (K) (Congalton e Green, 2009); e **(e)** Kappa condicional (também conhecido como Kappa-classe).

O índice Kappa é dado pela Equação 3.8.

$$K = \frac{n \sum_{i=1}^{k} n_{ii} - \sum_{i=1}^{k} (n_{i+} n_{+i})}{n^2 - \sum_{i=1}^{k} (n_{i+} n_{+i})}, \tag{3.8}$$

em que n_{ii} - número total de amostras corretamente classificadas da classe k;

n_{i+} - número total de amostras classificadas da classe k;
n_{+i} - número total de amostras coletadas da classe k; e
n - número total de amostras.

Concebido por Rosenfield e Fitzpatrick-Lins (1986) sob a alegação de que os vértices da matriz de confusão estavam sendo subutilizados, o *Kappa* condicional objetiva avaliar o desempenho individual das classes em relação ao total de amostras de referência, através da Equação 3.9.

$$\hat{K} = \frac{(N \cdot x_{ii}) - (x_{i+} \cdot x_{+i})}{(N \cdot x_{i+}) - (x_{i+} \cdot x_{+i})}, \tag{3.9}$$

em que N - soma das amostras de validação coletadas para todas as classes;
x_{ii} - número de elementos corretamente classificados em i (diagonal principal);
x_{i+} - total de elementos classificados para a classe i (soma do eixo y); e
x_{+i} - total de pontos de referência observados para a classe i (soma do eixo x).

Os scripts utilizados para tratamento dos dados de classificação, validação estatística dos mapas e pontos aleatórios são apresentados no APÊNDICE F. Outras questões como a autocorrelação espacial e o método de amostragem, que também podem exercer influência direta sobre o resultado dos testes estatísticos, serão abordados nas Seções 3.15.1 e 3.15.2 a seguir.

3.15.1 Autocorrelação Espacial

A autocorrelação espacial pode ser definida como a susceptibilidade à influência de dada característica sobre as características de suas unidades vizinhas (CONGALTON, 1991). Tal condição é de importância significativa na análise de erros dos segmentos classificados neste trabalho, dado que os mesmos se apresentam em quantidade limitada, possuem alta similaridade com seus vizinhos e não estão distribuídos de forma homogênea pela imagem.

Pelo fato de a área de estudo aqui apresentada se tratar de uma região pequena, a premissa da influência espacial se faz presente pela limitação das amostras disponíveis. Classes como vidro aramado, piscina de vinil ou lona, por exemplo, ocorrem apenas em regiões muito específicas da imagem e em alguns poucos segmentos.

Algumas classes, por apresentarem baixa ocorrência relativa na cena, carecem de uma forma mais honesta de validação, pois o simples sorteio de pontos aleatórios

na imagem não seria capaz de contemplar todas as 36 classes do trabalho. Classes que apresentam uma área de maior ocorrência, como vegetação rasteira e herbácea, naturalmente receberiam mais pontos do que as classes que se manifestam em áreas menores, como lona ou fibra de vidro.

3.15.2 Amostragem Aleatória Estratificada

Segundo Congalton (1991), o método a ser utilizado durante a etapa de amostragem é de demasiada importância na avaliação de acurácia em uma dada classificação, e isso ocorre porque a seleção de um método de amostragem influencia diretamente na confecção da matriz de confusão, que é a representação da totalidade da imagem classificada. A má decisão referente ao método de amostragem pode incutir um viés na matriz de confusão, possivelmente subestimando ou sobrestimando a acurácia (representada através do coeficiente *Kappa*).

Apesar da menor complexidade acoplada ao se trabalhar com as propriedades estatísticas do método de amostragem aleatória simples, este método de amostragem não é a melhor alternativa para todos os cenários. A amostragem aleatória simples tende a negligenciar as áreas pequenas importantes para classificação, a não ser para os casos em que o número de amostras é aumentado de forma significativa. Por esta razão, a amostragem aleatória estratificada deve ser favorecida quando houver a necessidade de um valor mínimo de amostras para cada classe.

Congalton e Green (2002) sugerem que o método de amostragem estratificada também pode falhar no quesito praticidade, quando a geração e coleta de pontos aleatórios de informação no solo ocorrer em áreas inacessíveis e também pela questão restritiva de que os pontos só podem ser gerados após a classificação da imagem. Isso limita a coleta de dados de validação para o final do projeto ao invés de executar a atividade em conjunto com a coleta de amostras de treinamento.

A diferença de área disponível para cada uma das classes do trabalho sugere que se adote a distribuição de pontos de validação estratificados por classe, conforme sugerido por Congalton (1991). Tal abordagem usa como base um mapa temático de referência. E a estratificação é feita através de uma rotina escrita em [R], presente no Apêndice F, que faz uma álgebra de mapa simples e segue as seguintes etapas:

(I) O usuário define alguns parâmetros e dentre eles a quantidade mínima de pontos que devem ser distribuídos para cada uma das classes do mapa vetorial selecionado para estratificação.

(II) Iniciado o algoritmo, calcula-se o percentual de área total para cada uma das classes presentes neste mapa.

(III) Multiplica-se o percentual de área da classe N pela soma total de todos os pontos gerados para validação, onde N = número da classe.

(IV) O resultado de (III) é individualmente armazenado, representando a quantidade de pontos que devem ser gerados para a classe N.

(V) O algoritmo faz i iterações, gerando uma quantidade de pontos específica para cada classe, conforme armazenado em (IV), onde i representa o número de classes presentes na referência.

(VI) Os pontos gerados após todas as iterações em (V) são unidos de forma a criar uma única malha de pontos, encerrando o *script* e gerando um novo arquivo vetorial em formato *shapefile*.

De forma a buscar significância estatística, foi escolhida uma média de 30 pontos por classe, para um total de 36 classes no Nível de Legenda II, totalizando 1.080 pontos. Em seguida, os pontos do Nível de Legenda II tiveram suas classes remapeadas para o Nível I, uma vez que as duas apresentam relação hierárquica. O mapa de pontos aleatórios gerados pelo algoritmo de estratificação é apresentado na Figura 3.20 e posteriormente discutido no Capítulo 4.

Figura 3.20 - Resultado do sorteio de pontos de validação estratificados por classe.

Fonte: Produção do autor.

4 RESULTADOS E DISCUSSÃO

Neste capítulo, serão apresentados os resultados de quatro dos doze cenários de classificação através dos três algoritmos propostos, C4.5, C5.0 e CART, para os Níveis I e II de Legenda de classificação, como também a influência do MDA em cada resultado. E de forma a tornar a leitura menos exaustiva, as oito demais classificações e suas respectivas matrizes de confusão foram redirecionadas para o Apêndice C (matrizes de confusão) e Apêndice D (imagens classificadas).

4.1 Parâmetros da Segmentação

Primando pela separabilidade entre as diversas classes, os especialistas sempre buscam uma suposta segmentação ideal, mas a definição do que é uma segmentação ideal é vaga; então, de forma a amenizar a arbitrariedade deste processo e em favor da reprodutibilidade deste experimento, a Tabela 4.1 documenta os procedimentos empíricos utilizados neste trabalho durante a tomada de decisão quanto aos valores escolhidos para segmentação.

A segmentação da imagem se iniciou no nível de escala 60 e seguiu até o nível 120, com um crescimento de 2 em 2 no fator de escala. Um dos cuidados tomados com esse método foi o foco em alvos de interesse que são frágeis e precisavam se manter distintos após a segmentação (piscinas, lona, vidro aramado, concreto pintado de vermelho presente no heliporto e demais classes que apresentam poucas amostras devido a sua baixa ocorrência na imagem), cujos exemplos são apresentados na Figura 4.1. Portanto, toda vez que os segmentos destes alvos se mesclavam a seus vizinhos, a segmentação era fixada na escala vigente e ajustes eram feitos no peso das bandas até que o segmento se tornasse distinto, permitindo sequência no procedimento.

Tabela 4.1 - Histórico sequencial dos parâmetros utilizados na segmentação.

Escala	Segmentos	Peso da Banda								Observações
		B1	B2	B3	B4	B5	B6	B7	B8	
60	45259	1	1	1	1	1	1	1	1	MDA e componentes principais (PC1:PC8) recebem peso = 0.
62	42636	1	1	1	1	1	1	1	1	
64	39860	1	1	1	1	1	1	1	1	
66	37793	1	1	1	1	1	1	1	1	
68	35713	1	1	1	1	1	1	1	1	
70	33609	1	1	1	1	1	1	1	1	Parte vermelha da quadra de tênis se mistura com a parte verde.
70	37830	1	1	1	1	1	1	1	1	PC1 recebe peso = 1. Aumento de contraste adiciona 4.331 novos segmentos.
70	34120	1	1	1	1	1	2	1	1	PC1 retorna para peso = 0 e B6 recebe peso = 2.
72	32493	1	1	1	1	1	2	1	1	
80	26790	1	1	1	1	1	2	1	1	Segmento de vidro aramado permanece distinto. Lona e demais polímeros perdem definição.
80	29830	1	1	1	1	1	2	1	1	Segmento de sombra da lona separa, mas parte vermelha da quadra de tênis se mistura com a verde. PC1 novamente recebe peso = 1.
80	27798	1	1	1	1	1	2	1	1	PC2 recebe peso = 1. Quadra de tênis se separa, mas lona permanece misturada com sombra.
80	26548	1	1	1	1	1	2	1	1	B4 (Amarelo) recebe peso = 2. Segmentos do heliporto, faixas de sinalização no asfalto e solo exposto refinados.
80	30010	1	1	1	1	1	2	1	1	PC2 recebe peso = 0. Mistura no segmento de lona permanece.
80	...	.	.	.	.	.	.	.	.	Valor de escala para segmentação fixada em 80: Início dos experimentos apenas com peso das bandas.
82	28813	1	1	1	2	3	2	2	4	Peso das bandas separam satisfatoriamente segmentos de interesse permitindo aumento do fator de escala.
84	27560	1	1	1	2	3	2	2	4	
86	26509	1	1	1	2	3	2	2	4	
88	25530	1	1	1	2	3	2	2	4	
90	24587	1	1	1	2	3	2	2	4	
92	23737	1	1	1	2	3	2	2	4	
94	22913	1	1	1	2	3	2	2	4	PC1:PC3 recebem peso = 1
96	22154	1	1	1	2	3	2	2	4	
98	22073	1	1	1	2	3	2	2	4	
100	21795	1	1	1	2	3	2	2	4	
110	17361	1	1	1	2	3	2	2	4	Vidro aramado se mistura com sombra.
120	15064	1	1	1	2	3	2	2	4	Lona se mistura com materiais adjacentes.
104	**21101**	**1**	**1**	**1**	**2**	**3**	**2**	**2**	**4**	Fator de escala retrocede de 120 para 104, pois é atingido o limite de escala capaz de separar os segmentos de vidro aramado. **Fim da segmentação**.

Fonte: Produção do autor.

Figura 4.1 - Alvos de interesse sensíveis ao processo de segmentação.

Fonte: Produção do autor.

A Figura 4.1a apresenta duas das três quadras de tênis presentes na área de estudo, sendo estas as únicas representantes da classe *Argamassa de Cimento Pintado de Vermelho*. Seus segmentos permitiriam uma escala maior do que os segmentos da classe Concreto Pintado de Vermelho (Figura 4.1d), localizada no heliporto do hospital, mas os segmentos de ambas as classes precisam se manter distintos na imagem, principalmente por apresentarem uma quantidade limitada de amostras, conforme já foi apresentado no histograma de amostras coletadas por classe (Figura 3.12). Outras classes são ainda mais sensíveis ao tamanho dos segmentos, como, por exemplo, as amostras de *Aço Galvanizado com Ferrugem* (Figura 4.1c), ou ainda as classes *Vidro Aramado* e *Acrílico*, ambas no telhado do hospital (Figura 4.1e). As classes de *Cerâmica* (Figura 4.1b) ou as ciclofaixas que alocam a classe *Asfalto Pintado de Vermelho* (Figura 4.1f), tendem a ser menos sensíveis ao processo de segmentação por suas propriedades espectrais distintas no vermelho, mas ainda é importante valorizar o peso da B5 (Vermelho) no processo de segmentação, para que esta não deturpe os segmentos, já limitados para essas classes.

4.2 Comportamento e Escolha dos Atributos pelos Classificadores

As análises descritas neste capítulo dizem respeito apenas às três primeiras árvores apresentadas no APÊNDICE B, que são as árvores dos classificadores C4.5, C5.0 e CART, respectivamente, com MDA e referentes ao Nível II. Esta análise não leva em consideração as subárvores provenientes do método de *boosting*, responsável por construir outras nove árvores para cada classificador, sendo 10 árvores por algoritmo com e sem o MDA, que juntos totalizam um conjunto de 60 árvores apenas para o Nível II.

Para auxiliar a compreensão de como os algoritmos C4.5, C5.0 e CART se comportam com e sem o MDA, a Tabela 4.2 identifica quais atributos foram mais frequentes durante a etapa de classificação dos segmentos. Grifados em cinza são os atributos que dizem respeito ao MDA, sendo estes `max_pixel_2` (valor do maior píxel de MDA), `min_pixel_9` (valor do menor píxel de MDA), `mean_mda` (um dos poucos atributos com nome intuitivo, representa média simples do MDA dentro do segmento analisado) e `modeMinimu8` (moda do píxel de MDA), que foi utilizado, mas não aparece na listagem por não estar entre os dez mais utilizados.

É importante notar que o nome dos atributos foi truncado para 12 caracteres pelo eCognition no momento de exportar os atributos para um arquivo de texto. E para plena compreensão da Tabela 4.2, é sugerida a leitura das legendas originais discriminadas no Apêndice , que apresenta a listagem dos nomes dos atributos antes e depois de serem truncados pela plataforma.

Ao todo, os classificadores receberam 94 atributos por segmento, de um total de 21.101 segmentos, gerando pouco mais de $1,9 \cdot 10^6$ observações em uma tabela em formato CSV com informações detalhadas sobre o recorte da cena. O arquivo de texto CSV tem tamanho aproximado de 16 megabytes, e o processo de classificação durou aproximadamente duas horas para cada um dos classificadores em todos os cenários (com e sem a presença do MDA), com exceção do C5.0, que terminou a varredura e classificação do CSV em doze segundos.

Dos 94 atributos, 82 foram utilizados quando somadas todas as ocorrências das três árvores, e os 12 atributos que não se fizeram presentes em nenhuma das árvores são apresentados na Tabela 4.3. A Tabela 4.4 apresenta os 10 atributos mais utilizados, e por fim, a Tabela 4.5 apresenta os 10 atributos menos utilizados pelos algoritmos de construção das árvores.

Tabela 4.2 - Ranking dos 10 atributos mais utilizados por cada classificador no Nível de Legenda II, com e sem o MDA.

	Ordem	Termo Chave	Ocorrências		Ordem	Termo Chave	Ocorrências
CART	1.	ndsi1	8	CART+MDA	1.	ndbi	10
	2.	mean_pc7	8		2.	brtmn1	8
	3.	mean_pc8	6		3.	max_pixel_2	8
	4.	asymmetry	6		4.	wvbi	6
	5.	wvbi	4		5.	min_pixel_9	6
	6.	ndbi	4		6.	mean_pc7	6
	7.	mean_layer5	4		7.	max_pixel_6	6
	8.	min_pixel_6	4		8.	mean_layer4	4
	9.	min_pixel_3	4		9.	min_pixel_12	4
	10.	brtmn1	4		10.	min_pixel_2	4
C45	1.	border_ind	18	C45+MDA	1.	border_ind	20
	2.	brtpbor	12		2.	asymmetry	16
	3.	mean_pc4	10		3.	wvsi	8
	4.	ndvi1	8		4.	brtpbor	6
	5.	brtmn1	8		5.	ndsi1	6
	6.	mean_pc7	8		6.	compactnes	6
	7.	brt	8		7.	min_pixel_6	6
	8.	max_pixel_3	8		8.	min_pixel_3	6
	9.	asymmetry	8		9.	mean_pc4	6
	10.	ndsi2	6		10.	mean_pc6	6
C50	1.	border_ind	12	C50+MDA	1.	mean_mda	16
	2.	ndbi	12		2.	min_pixel_9	12
	3.	mean_pc6	10		3.	max_pixel_2	10
	4.	mean_pc7	10		4.	border_ind	8
	5.	brtpbor	8		5.	wvsi	8
	6.	ndsi1	8		6.	min_pixel_3	8
	7.	min_pixel_3	8		7.	mean_pc3	8
	8.	mean_pc8	8		8.	wvbi	7
	9.	max_pixel_3	8		9.	ndsi2	6
	10.	ndsi2	6		10.	ndsi1	6

*Grifados em cinza são os atributos que dizem respeito ao MDA.

Fonte: Produção do autor.

Tabela 4.3 - Atributos descartados pelos classificadores nas seis primeiras árvores dos algoritmos C4.5, C5.0 e CART com e sem atributos do MDA.

Ordem	Atributo
1.	Max_pixel_13
2.	Mean_Layer
3.	Min_pixel_15
4.	mNIR
5.	modeMinimu
6.	modeMinimu3
7.	modeMinimu4
8.	modeMinimu5
9.	modeMinimu6
10.	modeMinimu7
11.	modeMinimu10
12.	WVNHFD

Fonte: Produção do autor.

As árvores construídas para todas as classificações são apresentadas no APÊNDICE B e referem-se ao Nível de Legenda I e II, ou seja, 12 e 36 classes, respectivamente. As árvores de classificação do Nível de Legenda I são menores por contarem com apenas 12 classes e supostamente devem apresentar um maior poder de generalização, por consequência levando à utilização de menos atributos.

Tabela 4.4 - Atributos mais utilizados pelas árvores de decisão para as classificações do Nível de Legenda II.

TOP 10		
Ordem	Termo Chave	Ocorrências
1.	border_ind	60
2.	ndbi	40
3.	asymmetry	37
4.	mean_pc7	36
5.	brtpbor	34
6.	ndsi1	34
7.	min_pixel_3	32
8.	mean_pc8	32
9.	mean_pc4	30
10.	wvbi	29

Fonte: Produção do autor.

Tabela 4.5 - Atributos menos utilizados pelas árvores de decisão para as classificações do Nível de Legenda II.

BOTTOM 10		
Ordem	Termo Chave	Ocorrências
72.	modeminimu2	4
73.	modeminimu8	4
74.	modeminimu11	4
75.	max_pixel_5	4
76.	mswir	4
77.	elliptic_f	4
78.	max_pixel_14	4
79.	mean_layer1	2
80.	min_pixel_10	2
81.	max_pixel_15	2

Fonte: Produção do autor.

4.3 Análise Exploratória dos Atributos

A seção anterior apresentou os atributos mais utilizados pelas árvores de decisão para buscar a separação ideal entre as classes, restando agora compreender como essas classes se comportam dentro de cada atributo. Para tal, foram utilizados diagramas de espalhamento, gerados em linguagem [R], com as distribuições das amostras para alguns dos atributos que foram apresentados na Seção 4.2. Todos os diagramas de espalhamento (*scatterplots*) foram confeccionados de forma a seguir o esquema de cores do mapa temático, apresentado de antemão na Figura 4.2.

Figura 4.2 - Legenda das classes temáticas no Nível de Legenda II para os diagramas de espalhamento.

Fonte: Produção do autor.

Também é feita uma ressalva para apresentação dos *scatterplots* no que diz respeito

à remoção de valores anômalos ou *outliers* (setas azuis) e *zoom* (seta vermelha) no gráfico conceitual ilustrado na Figura 4.3. Alguns segmentos apresentam pontos espúrios recorrentes, que dificultam a visualização dos alvos de interesse e precisam ser removidos do diagrama para que seja possível a interpretação dos resultados.

É imprescindível notar que a presença de *outliers* nas amostras pode influenciar, mas não compromete o trabalho dos algoritmos de classificação, uma vez que estes apresentam mecanismos de tratamento mais ou menos robustos para esta finalidade, variando conforme o algoritmo.

Figura 4.3 - Comportamento das amostras para valores de NDVI, apresentando *outliers* em azul e um recorte sugerindo *zoom* para a área de interesse em vermelho.

Fonte: Produção do autor.

Pode ser observado na Figura 4.3 que a visualização do resultado final é comprometida por amostras, principalmente da classe sombra, que fogem aos limiares dos

índices de NDVI. Tal efeito possivelmente ocorre por conta da correção atmosférica por FLAASH, e em segunda instância, por uma possível distorção das grandezas radiométricas nos pixels de alguns alvos da imagem quando estas foram inseridas no eCognition, uma vez que a plataforma faz uma normalização em todos os dados de entrada. É de extrema importância primar pela integridade espectral dos dados durante seu processamento, mas tais fenômenos de distorção não chegam a ser de todo um problema, uma vez que os algoritmos se preocupam apenas com a separabilidade entre as amostras, abstraindo os conceitos físicos incutidos em seus níveis de cinza (NC).

A mesma distribuição de amostras pontilhadas no quadrado vermelho da Figura 4.3 é apresentada na Figura 4.4. Essa informação é relevante, porque tais imposições de recortes nos eixos X e Y também serão aplicadas nos diagramas subsequentes, primando pela melhor distinção dos resultados.

Figura 4.4 - Valores das amostras para os atributos de NDVI 1 e 2.

Fonte: Produção do autor.

Aplicados os devidos filtros, torna-se possível observar uma alta correlação positiva, que é natural, considerando que ambos os eixos são variações das bandas do vermelho (B5) e vermelho limítrofe (B6) para o mesmo atributo de NDVI. A separabilidade entre as amostras se torna distinta para cada classe, e embora o NDVI1 tenha sido utilizado apenas para ilustrar a necessidade de *zoom* e recorte de *outliers* nos diagramas de espalhamento, ele também foi o 4°atributo mais utilizado pelo modelo C4.5 sem MDA. O atributo de NDVI também se demonstrou consistente em separar alvos fotossinteticamente ativos e mais próximos de 1 em ambos os eixos, incluindo a classe Lago, sugerindo maiores quantidades de matéria orgânica em suspensão. Outros atributos também apresentam grandes contribuições no que diz respeito à separabilidade entre as classes, como, por exemplo o atributo de Brilho (*Brightness* ou BRT), apresentado junto ao `max_pixel_2` (nome truncado do atributo `max_pixel_mda`), ilustrado na Figura 4.5.

Figura 4.5 - Diagrama de espalhamento dos atributos Max_Pixel_MDA x BRT.

Fonte: Produção do autor.

É interessante notar a coerência presente nas classes referentes às edificações mais altas da imagem, que são também as classes de maior brilho, chegando até a saturar esse atributo. A classe Acrílico tampouco é exceção, visto que grande parte das amostras disponíveis para essa classe foram coletadas no telhado central do hospital, definida como a edificação mais alta da imagem. Alvos de baixa reflectância, que absorvem mais REM, estão presentes na parte inferior do eixo Y, apresentando brilho inferior à zero por conta de seus valores negativos na banda do Azul Costeiro (B1).

O atributo de brilho foi recorrente nas análises aqui apresentadas, por ser um excelente separador para as amostras, e o índice proposto por Leonardi (2010) também traz grandes contribuições para a separabilidade espectral do conjunto de dados, explorando o potencial do atributo de brilho, a banda do Azul e as bandas do Vermelho, conforme apresentado na Figura 4.6.

Figura 4.6 - Contraste dos alvos de alta reflectância no Azul (B2) e Vermelho (B5), explorados pelo índice de Leonardi (2010) adaptado para Vermelho Limítrofe (B6).

Fonte: Produção do autor.

Ambos os índices enfatizam alvos de maior reflectância nos comprimentos de onda do Azul (B2), tipicamente representados pelas amostras da classe Lago, piscinas e segmentos de sombra das vegetações que margeiam o lago. Os demais alvos que apresentam alta reflectância na B2 e também nas demais bandas, como as ligas metálicas de Galvalume A, B e C, tendem a se concentrar no ponto médio da correlação entre o índice de Leonardi (2010) e sua variação com a B6, neutralizados pela divisão da B2 pela B5 e B6, separados apenas pelo seu valor de brilho. Mas o diferencial desses atributos está nas classes que refletem mais nos vermelhos da B5 e B6, otimizando a separabilidade entre os alvos de cerâmica, solo exposto, concreto pintado de amarelo e asfalto pintado de vermelho.

Outros atributos dignos de destaque foram as componentes principais, que apresentaram um potencial peculiar de separabilidade entre as classes. A Figura 4.7 apresenta a PC8, mas outras componentes também se fizeram presentes no *ranking* de atributos mais utilizados pelos classificadores.

Figura 4.7 - Atributo proveniente da oitava componente principal versus o atributo de brilho, que juntos apresentam potencial de separabilidade entre algumas classes.

Fonte: Produção do autor.

Os experimentos demonstraram que as componentes principais apresentaram diversas formas de contribuição. A primeira componente se comportou como uma espécie de contraste sintético, o que sugere que ela possa ser utilizada como um atributo complementar ou um possível substituto para o atributo de brilho, mas cabe notar que ela também apresenta os níveis de cinza invertidos. Tal contribuição também era esperada, uma vez que as transformações ortogonais provenientes de uma análise por componentes principais intenciona gerar a primeira variável de saída da análise, como o atributo de maior contribuição para a variância do conjunto. A Tabela 4.6 é apresentada de forma a ilustrar a contribuição de cada banda espectral na composição das principais componentes.

Tabela 4.6 - Percentual de contribuição das bandas espectrais do WV-2 na composição da Análise por Principais Componentes.

	B1	B2	B3	B4	B5	B6	B7	B8
PC1	12,05%	12,20%	12,33%	12,39%	12,38%	12,72%	12,86%	13,06%
PC2	17,39%	11,02%	4,89%	3,24%	2,89%	6,11%	21,53%	32,91%
PC3	34,76%	5,17%	0,51%	14,93%	22,82%	11,30%	5,29%	5,21%
PC4	1,64%	0,35%	2,42%	2,19%	0,20%	0,31%	50,01%	42,87%
PC5	1,52%	5,54%	7,75%	0,14%	45,31%	32,76%	6,40%	0,57%
PC6	24,73%	15,81%	28,42%	9,30%	0,94%	15,40%	0,35%	5,04%
PC7	0,12%	12,17%	11,27%	41,23%	13,84%	18,78%	2,23%	0,36%
PC8	7,78%	37,73%	32,39%	16,57%	1,61%	2,62%	1,30%	0,00%

Fonte: Produção do autor.

Também foram observadas contribuições da segunda componente durante a etapa de segmentação e escolha do peso das bandas; isso ocorreu por conta de contrastes pontuais nos pixels da cena, que se manifestaram nas bordas dos objetos da imagem. E, de uma forma mais frequente, esses contrastes surgiram sistematicamente nas transições entre as classes que apresentam maiores contrastes entre si, como alvos orgânicos (vegetação) versus alvos minerais e/ou sintéticos (asfaltos e concretos) .

A terceira componente apresentou contribuições sutis na separabilidade dos alvos de maior reflectância no vermelho, mas também enfatizou confusões sobre os segmentos da classe Lona, impedindo que um maior peso lhe fosse atribuído durante o processo

de segmentação. E nas componentes de cinco a sete, foram observadas oscilações entre contribuição e ruído para diversas classes, principalmente sobre os alvos que apresentam maior reflectância nos comprimentos de onda entre 620 nm e 900 nm, que compreende as bandas do vermelho, vermelho limítrofe e infravermelho próximo-1 do sensor WV-2.

De forma a apresentar a variabilidade no comportamento de cada componente, é apresentada a Figura 4.8, que ilustra o resultado da análise por componentes principais.

Figura 4.8 - Atributos resultantes da análise por componentes principais sobre as oito bandas multiespectrais do sensor WV-2.

Fonte: Produção do autor.

4.4 Análise de Acurácia da Classificação

Nesta seção, são apresentados os resultados gerais dos modelos de classificação e discutidos os resultados dos melhores modelos, quando utilizados todos os 94 atributos (com MDA) versus 90 atributos (sem MDA). Também são apresentados os resultados para o primeiro e segundo níveis de legenda de classificação, dispostos na Tabela 4.7.

Tabela 4.7 - Resultados da avaliação dos coeficientes de concordância *Kappa*.

Nível de Legenda	Situação	Classificador	Coeficiente Kappa	σ^2 do *Kappa*
I	com MDA	C5.0	0,7338436	0,000193455
		C4.5	**0,741837**	0,000189977
		CART	0,7298324	0,000195151
	sem MDA	C5.0	0,7336602	0,000193052
		C4.5	0,735121	0,000193764
		CART	**0,7450042**	0,000188382
II	com MDA	C5.0	0,6858477	0,000185674
		C4.5	0,7092944	0,000176655
		CART	**0,7247778**	0,000172092
	sem MDA	C5.0	0,6965478	0,000182773
		C4.5	**0,7168415**	0,000175587
		CART	0,7145529	0,000177949

Fonte: Produção do autor.

Os algoritmos de classificação que apresentaram o maior coeficiente de concordância *Kappa* estão marcados em negrito, sendo estes: C4.5 ($K = 0,74$) com o melhor resultado estatístico de classificação fazendo uso do MDA, e o CART ($K = 0,74$) com o melhor resultado sem MDA, respectivamente, ambos para o Nível de Legenda I, que totaliza 12 classes (*Cerâmica, Coberturas Diversas, Coberturas Metálicas, Lago, Pavimentação Não-Viária, Pavimentação Viária, Piscina, Solo Exposto, Sombra, Vegetação Herbácea, Vegetação Rasteira e Materiais Mistos*).

Para o Nível de Legenda II, CART ($K = 0,72$) é a melhor classificação com o uso do MDA, sendo que sem o uso do MDA o maior coeficiente Kappa é do algoritmo C4.5 ($K = 0,71$), para um total de 36 classes (*Aço Galvanizado Brilhante,*

Aço Galvanizado com Ferrugem, Aço Galvanizado Fosco, Acrílico, Argamassa de Cimento, Argamassa de Cimento Pintado de Verde, Argamassa de Cimento Pintado de Vermelho, Asfalto, Asfalto Pintado de Branco, Asfalto Pintado de Vermelho, Bloquete, Carvão Coque, Coberturas Cerâmicas, Concreto, Concreto Impermeabilizado, Concreto Pintado de Amarelo, Concreto Pintado de Vermelho, Fibra de Vidro, Fibrocimento Amianto, Fibrocimento CRFS, Galvalume A, Galvalume B, Galvalume C, Lago, Lona, Pastilha Esmaltada, Piscina de Azulejo, Piscina de Vinil, Plástico, Policarbonato, Quartzito Amarelo, Solo Exposto, Sombra, Vegetação Arbórea, Vegetação Herbácea e Vidro Aramado).

Os erros, acertos e concentração de pontos de validação específicos para cada classe foram computados através do mapa de referência (C4.5 Nível II com MDA), sendo este também utilizado na estratificação das amostras de validação e apresentado na Figura 4.9. Tais informações sobre os erros e acertos serão discutidas em maior detalhe posteriormente e podem ser observadas nas matrizes de confusão do Nível de Legenda I, sendo apresentadas apenas as matrizes com maiores valores de Kappa com MDA (Tabela 4.9) e sem MDA (Tabela 4.10).

O mesmo vale para as matrizes de confusão do segundo nível, mas apenas as matrizes dos modelos classificadores estatisticamente mais distintos (maior σ^2) são apresentadas, independente da qualidade do Kappa, sendo estes o CART com MDA (Tabela 4.12) e C5.0 com MDA (Tabela 4.13).

Figura 4.9 - Acertos (a), erros (b) e concentração por área(c), observados para os pontos de validação, em relação à classificação C4.5 Nível II com MDA.

Fonte: Produção do autor.

Foi possível validar 1.218 pontos que se mantiveram constantes e foram utilizados para os 12 cenários de classificação. Tais pontos foram distribuídos aleatoriamente e estratificados por classe (Tabela 4.8) através de linguagem [R] (APÊNDICE - F).

Tabela 4.8 - Distribuição por classe dos pontos de validação.

NÍVEL I	PONTOS	NÍVEL II	PONTOS
Solo Exposto	41	Solo Exposto	41
Piscina	36	Piscina de Azulejo	30
		Piscina de Vinil	6
Lago	58	Lago	58
Vegetação Rasteira	94	Vegetação Herbácea	94
Vegetação Arbórea	72	Vegetação Arbórea	72
Pavimentação Não Viária	46	Pedra Mineira	26
		Carvão Coque	20
Pavimentação Viária	142	Asfalto	84
		Asfalto Pintado de Vermelho	30
		Asfalto Pintado de Branco	22
		Bloquete	6
Materiais Mistos (Pavimentação e Cobertura)	256	Arg. de Cimento	22
		Arg. de Cimento Pint. de Vermelho	26
		Arg. de Cimento Pintada de Verde	27
		Concreto Pintado de Amarelo	6
		Concreto Pintado de Vermelho	25
		Concreto	111
		Concreto Impermeabilizado	39
e Coberturas Diversas	172	Vidro Aramado	22
		Fibra de Vidro	5
		Pastilha Esmaltada	9
		Fibrocimento CRFS	19
		Fibrocimento com Amianto	81
		Policarbonato	5
		Acrílico	18
		Plástico	9
		Lona	4
Coberturas Metálicas	210	Aço Galvanizado Brilhante	24
		Aço Galvanizado com Ferrugem	11
		Aço Galvanizado Fosco	57
		Galvalume A	62
		Galvalume B	31
		Galvalume C	25
Coberturas Cerâmicas	17	Coberturas Cerâmicas	17
Sombra	74	Sombra	74
TOTAL	**1218**		**1218**

Fonte: Produção do autor.

As principais concentrações referentes aos pontos de validação são apresentadas nas manchas vermelhas da Figura 4.9c, estando essas associadas às classes de baixa ocorrência na imagem. As concentrações ocorrem, pois só é permitido ao *script* a geração de pontos aleatórios sobre os segmentos de determinadas classes, daí a estratificação (CONGALTON; GREEN, 2009). Como as classes de baixa ocorrência costumam dispor de poucos segmentos, muitos pontos se concentram na pouca área disponível dentro dos segmentos. Os pontos nunca caem no mesmo pixel, mas podem cair mais de uma vez no mesmo segmento. Tal fenômeno pode exercer contribuição significativa para o aumento ou diminuição dos índices de validação dos mapas, e é importante ter essa informação em mente durante o processo de análise das matrizes de confusão.

É possível observar concentrações de acertos (Figura 4.9a) nas classes de *Piscina de Azulejo, Piscina de Vinil, Concreto Pintado de Vermelho* (heliporto do hospital) e *Argamassa de Cimento Pintado de Verde/Vermelho* (quadras de tênis). Os erros da Figura 4.9b dizem respeito à classificação C4.5 Nível II com MDA e estão principalmente concentrados nas diversas classes de telhados, concretos diversos e solo exposto/cerâmica.

Tabela 4.9 - Matriz de confusão do algoritmo C4.5 fazendo uso de MDA, modelo que apresentou o maior valor de *Kappa* (0,74) para o Nível de Legenda I.

Fonte: Produção do autor.

Tabela 4.10 - Matriz de confusão CART, modelo que apresentou o maior *Kappa* (0,74) para o Nível de Legenda I sem a necessidade de um MDA.

Fonte: Produção do autor.

Nas matrizes, a escala de cores da legenda representa o coeficiente de concordância do *Kappa* condicional (*Kappa*-classe). Na diagonal principal, as células em tons de vermelho intenso representam maior acurácia, enquanto que os tons mais claros representam maior discordância para esta classe. Também cabe notar que mesmo classes com grandes quantidades de acertos na diagonal principal não implicam um alto valor K-classe, uma vez que a sua equação não considera apenas a quantidade total de acertos, mas também a quantidade total de amostras coletadas, acurácia do usuário e acurácia do produtor. A classe *Cerâmica* - C2 pode ilustrar esse conceito, pois ela apresenta 82% de acerto nas observações, mas seu K-classe = 0,63 (valores individuais das demais classes são apresentados na Tabela 4.11).

Para a classe de Pavimentação Viária, a maior fonte de confusão de suas observações se concentra em Materiais Mistos e Coberturas Diversas. Quando somadas, essas três classes compreendem 20 subclasses do Nível II, o que significa 55% de todas as classes propostas para este nível. Tal relação de confusão na matriz não surpreende quando tomada em consideração a diversidade espectral dos alvos que a integram: *Asfalto, Asfalto Pintado de Vermelho, Asfalto Pintado de Branco* e *Bloquete*.

Mei et al. (2012) argumentam que o asfalto é caracterizado por um comportamento espectral pouco variado e algumas de suas feições também podem ser prejudicadas por ocorrerem nas janelas de atenuação atmosférica do vapor d´água (H_2O), suprimindo feições de betumes e reduzindo a separabilidade desta classe, conforme medido in situ e ilustrado na Figura 4.10.

Tabela 4.11 - Acurácia do Produtor, Usuário e Kappa-classe individual para todos os cenários de classificação do Nível de Legenda I.

	CART+MDA			CART			C4.5+MDA			C4.5			C5.0+MDA			C5.0		
	K-classe	Prd.	Usr.	*K*-classe	Prd.	Usr.	*K*-classe	Prd.	Usr.	*K*-classe	Prd.	Usr.	*K*-classe	Prd.	Usr.	*K*-classe	Prd.	Usr.
C12 - Vegetação Herbácea	0,84	0,62	0,83	0,83	0,84	0,63	0,82	0,65	0,81	0,89	0,89	0,63	0,84	0,86	0,63	0,87	0,88	0,64
C11 - Solo Exposto	0,59	**0,39**	0,58	0,62	0,63	0,46	0,57	**0,39**	0,56	0,53	0,55	0,54	0,44	0,45	0,49	0,46	0,47	0,44
C10 - Vegetação Arbórea	0,85	0,86	0,84	0,83	0,84	0,79	0,86	0,83	0,85	0,80	0,81	0,83	0,83	0,84	0,86	0,75	0,77	0,86
C9 - Pavimentação não-viária	0,88	0,76	0,87	0,92	0,92	0,78	0,86	0,83	0,86	0,83	0,83	0,87	0,78	0,79	0,83	0,80	0,80	0,89
C8 - Sombra	0,82	0,80	0,81	0,81	0,82	0,86	0,78	0,82	0,77	0,81	0,82	0,82	0,76	0,78	0,81	0,83	0,84	0,80
C7 - Coberturas Diversas	0,66	0,87	0,60	0,59	0,65	0,83	0,62	0,88	0,56	0,54	0,60	0,83	0,58	0,64	0,82	0,54	0,61	0,77
C6 - Materiais Mistos	0,76	0,67	0,70	0,70	0,76	0,75	0,75	0,66	0,68	0,66	0,73	0,72	0,70	0,76	0,67	0,74	0,79	0,73
C5 - Coberturas Metálicas	0,83	0,76	0,79	0,86	0,88	0,79	0,89	0,77	0,87	0,88	0,90	0,76	0,83	0,86	0,78	0,84	0,87	0,80
C4 - Piscina	0,75	**1,00**	0,74	**1,00**	**1,00**	0,97	0,95	**1,00**	0,95	**1,00**	**1,00**	**1,00**	0,95	0,95	**1,00**	0,97	0,97	0,97
C3 - Lago	**1,00**	**1,00**	**1,00**	**1,00**	**1,00**	**1,00**	**1,00**	**1,00**	**1,00**	**1,00**	**1,00**	**1,00**	**1,00**	**1,00**	0,98	0,98	0,98	**1,00**
C2 - Cerâmica	0,48	0,82	0,48	0,55	0,56	0,82	0,64	0,82	0,63	0,49	0,50	0,76	0,60	0,61	0,82	0,49	0,50	0,82
C1 - Pavimentação Viária	0,71	0,77	0,67	0,63	0,67	0,73	0,76	0,82	0,73	0,73	0,76	0,70	0,71	0,74	0,78	0,67	0,71	0,70

* Valores ≤ 0,40 estão grifados em vermelho e valores = 1, em verde.

Fonte: Produção do autor.

Figura 4.10 - Reflectância para as classes de Asfalto e Concreto e interferência das janelas de absorção atmosférica por vapor d´água (H_2O).

Fonte: Adaptado de Jensen (2009).

Ainda na classe de Pavimentação Viária, as matrizes do primeiro nível também apresentam a influência do MDA. De um total de 142 observações, o CART Nível I sem MDA ($K = 0,63$) acertou 104 observações (73%), contra um aumento de 9% com a inclusão dos atributos de MDA para a classificação por C4.5 ($K = 0,76$), que acertou 82% para esta mesma classe, representando uma diferença de 0,13 no K-classe. De forma peculiar, a classe de *Cerâmica*, que se manifesta exclusivamente no telhado das edificações, parece não sofrer influência do MDA. Sua classificação foi estável, com poucos erros, apresentando apenas uma confusão característica com a classe de solo exposto, uma vez que ambas compartilham de propriedades químicas similares originárias da argila (PINHO, 2006b; RIBEIRO, 2010).

A classe *Lago* apresentou grande potencial de separabilidade, com um total de 58 observações, sua acurácia foi de 100% em todos os cenários e K-classe = 1. Mas também cabe ressaltar que ao contrário de outras classes, a classe *Lago* dispunha de

uma vasta área para coleta de amostras de treinamento, permitindo previsões mais robustas nos modelos de classificação.

Assim, como a classe *Lago*, a classe de *Piscina* tampouco apresentou problemas de separabilidade espectral e foi classificada com $K = 1$. A literatura sugere que tamanha precisão pode ser fruto de sua cor cyan que facilita sua separabilidade dentre outros alvos urbanos (HEROLD et al., 2004; PINHO, 2006b). Cabe notar que a classe *Piscina* apresenta apenas 12 segmentos em toda imagem, isso significa que essas amostras apresentam uma alta correlação espacial e seus 30 pontos de validação são muito repetitivos porque mesmo que se assumisse uma nova distribuição aleatória, a chance de um novo ponto cair dentro dos 12 segmentos é alta, conforme previamente apresentado na Figura 4.9c, correspondente ao mapa de concentração dos pontos de validação. Essa informação sugere que embora a classificação para a classe Piscina tenha sido ótima, os indicadores estatísticos para essa classe podem não condizer com o real poder de predição dos modelos para esta legenda.

A classe *Coberturas Metálicas* ($K = 0,89$) é uma abstração de subclasses compostas por metais diversos, estes materiais sintéticos tendem a apresentar alto brilho em cena. Em experimentos realizados no Laboratório de Radiometria do DPA/INPE (LARAD), demonstraram-se similaridades espectrais em sua composição. Tal comportamento justificaria a confusão com outros alvos de origem sintética, como alguns subitens da classe *Coberturas Diversas* e *Materiais Mistos* (para detalhes das bibliotecas espectrais de alvos urbanos realizados em laboratório, vide APÊNDICE E).

Materiais Mistos ($K = 0,75$) e *Coberturas Diversas* ($K = 0,62$) apresentaram confusão mútua por se tratar de superclasses compostas por alvos espectralmente muito heterogêneos, mas o algoritmo CART, sem apoio dos atributos de MDA, apresentou 8% a mais de precisão para *Materiais Mistos* em relação ao C4.5 com MDA. Isso representa um acerto de 192 observações do CART contra 170 do C4.5 com MDA, de um total de 256 pontos observados.

A classe *Sombra* ($K = 0,81$) é composta por segmentos que pertencem a outras classes, que por sua vez são sobrepostas pelas projeções de objetos que carecem de iluminação, e essa característica introduz diversidade espectral na parca reflectância remanescente do comportamento espectral dessa classe, que não possui validação de campo.

Pavimentação Não-Viária ($K = 0,92$) é composta por materiais de calçamento que são utilizados exclusivamente por pedestres. A peculiaridade espectral de suas

subclasses (*Carvão Coque* e *Quartzito Amarelo* no Nível II), permitiu boa distinção de suas 46 observações, ocorrendo apenas seis omissões em *Materiais Mistos* e duas em *Pavimentação Viária*.

As classes de *Vegetação Arbórea* ($K = 0,83$), *Solo Exposto* ($K = 0,62$) e *Vegetação Herbácea* ($K = 0,83$) não apresentam subclasses, sendo suas legendas repetidas para o Nível de Legenda II, fator que motiva o adiamento de sua discussão para as matrizes de confusão das legendas do Nível II, apresentadas nas Tabela 4.12 e Tabela 4.13.

Tabela 4.12 - Matriz de confusão para classificação por CART com MDA para o Nível de Legenda II.

Fonte: Produção do autor.

Tabela 4.13 - Matriz de confusão para classificação por C5.0 com MDA para o Nível de Legenda II.

Fonte: Produção do autor.

As matrizes de confusão do Nível II permitem inferir com maior precisão quais alvos desse nível foram responsáveis pelos erros de omissão e comissão herdados pelas classes abstratas do Nível I. Algumas considerações também serão feitas a respeito das regras responsáveis por identificar os atributos de melhor separabilidade para as classes do Nível II. O resumo do desempenho individual de cada caso foi disposto na Tabela 4.14.

A classe *Aço Galvanizado Brilhante* ($K = 0,38$ para o cenário CART+MDA e $K = 0,43$ para C5.0+MDA) foi uma classe de de baixa omissão, mas alta comissão. Com 21 acertos de 24 observações, sua acurácia do produtor foi computada em 0,88. Mas outras 11 classes cometeram erros de comissão para Aço Galvanizado Brilhante, justificando seu baixo K-classe. Na árvore de decisão gerada pelo algoritmo C50+MDA, o principal atributo utilizado para separar seus pares foi `Mean_pc5 > -51.92517` (média na quinta componente principal), e tal regra enquadrou 70 dos 92 segmentos amostrados para essa classe. Outros atributos de menor contribuição, mas que também foram utilizados para alocar as demais amostras de treinamento, incluem: média no MDA e NDBI (Índice de Diferença Normalizada para Áreas construídas), o que é coerente, haja vista que a classe *Aço Galvanizado Brilhante* se manifesta apenas no telhado de edificações e pertence à superclasse *Coberturas Metálicas*.

Para *Aço Galvanizado com Ferrugem*, o comportamento da classificação em relação à inclusão do modelo digital de altura foi negativa. No melhor cenário de classificação, CART sem MDA, seu K atingiu 0,80, e no pior cenário (CART+MDA), o valor de K foi de 0,35. A influência negativa do MDA para esta classe pode haver ocorrido por uma má decisão do algoritmo, uma vez que, na árvore proposta por C5.0 com MDA, todas as 7 amostras foram alocadas pela regra `Max_pixel_MDA > 29.76`. Mas para os modelos onde o MDA não pode ser utilizado, forçar os algoritmos a recorreram a outros atributos acaba por se mostrar uma abordagem espectral mais razoável, utilizando por exemplo a regra `NDBI > 0.032`. O NDBI faz uso das bandas NIR1 e NIR2 do WV-2, que compreendem o espectro de 772-954 nm, favorecendo a identificação de materiais que apresentam em sua composição minerais ferrosos (DRURY, 1993).

Tabela 4.14 - Acurácia do Produtor, Usuário e Kappa-classe individual para todos os cenários de classificação do Nível de Legenda II.

	CART+MDA			CART			C4.5+MDA			C4.5			C5.0+MDA			C5.0		
	K-classe	Prd.	Usr.	*K*-classe	Prd.	Usr.	*K*-classe	Prd.	Usr.	*K*-classe	Prd.	Usr.	*K*-classe	Prd.	Usr.	*K*-classe	Prd.	Usr.
C36 - Vidro Aramado	0,95	0,95	0,82	0,89	0,89	0,77	0,69	0,70	0,95	**1,00**	**1,00**	0,73	0,84	0,84	0,95	0,67	0,68	0,77
C35 - Vegetação Herbácea	0,86	0,87	0,70	0,79	0,81	0,66	0,89	0,90	0,67	0,85	0,86	0,68	0,89	0,89	0,63	0,90	0,91	0,65
C34 - Vegetação Arbórea	0,86	0,87	0,83	0,83	0,84	0,79	0,91	0,91	0,89	0,80	0,81	0,85	0,83	0,84	0,89	0,88	0,88	0,85
C33 - Sombra	0,73	0,75	0,85	0,81	0,82	0,88	0,92	0,93	0,85	0,82	0,83	0,85	0,72	0,74	0,85	0,73	0,75	0,85
C32 - Solo Exposto	0,63	0,64	0,56	0,47	0,49	0,46	0,60	0,61	0,54	0,64	0,66	0,46	0,50	0,51	0,44	0,46	0,47	0,44
C31 - Quartzito Amarelo	**1,00**	**1,00**	0,77	0,80	0,80	0,77	0,80	0,81	0,81	0,64	0,65	0,77	0,87	0,88	0,81	0,58	0,59	0,77
C30 - Policarbonato	0,50	0,50	0,80	0,75	0,75	0,60	**0,17**	**0,17**	0,80	**0,40**	0,40	0,80	**0,25**	**0,25**	0,80	**0,23**	**0,24**	0,80
C29 - Plástico	0,60	0,60	**1,00**	0,75	0,75	**1,00**	0,50	0,50	**1,00**	0,47	0,47	**1,00**	0,69	0,69	**1,00**	0,56	0,56	**1,00**
C28 - Piscina de Vinil	**0,38**	**0,38**	0,83	0,67	0,67	**1,00**	**0,28**	**0,29**	**1,00**	0,55	0,56	0,83	**0,37**	**0,38**	**1,00**	0,67	0,67	**1,00**
C27 - Piscina de Azulejo	0,97	0,97	**1,00**	0,94	0,94	**1,00**	**1,00**	**1,00**	**1,00**	**1,00**	**1,00**	**1,00**	**1,00**	**1,00**	**1,00**	0,97	0,97	**1,00**
C26 - Pastilha Esmaltada	0,56	0,56	**1,00**	0,60	0,60	**1,00**	**0,31**	**0,31**	**1,00**	0,64	0,64	**1,00**	0,40	0,41	**1,00**	0,69	0,69	**1,00**
C25 - Lona	0,50	0,50	**1,00**	0,50	0,50	**1,00**	**0,19**	**0,19**	**1,00**	**0,37**	**0,38**	0,75	**0,17**	**0,18**	0,75	**0,37**	**0,38**	0,75
C24 - Lago	**1,00**	**1,00**	**1,00**	**1,00**	**1,00**	**1,00**	**1,00**	**1,00**	**1,00**	**1,00**	**1,00**	**1,00**	**1,00**	**1,00**	0,98	0,98	0,98	0,95
C23 - Galvalume C	0,67	0,68	0,68	0,72	0,72	0,84	0,61	0,62	0,72	0,74	0,75	0,84	0,57	0,58	0,72	0,66	0,67	0,80
C22 - Galvalume B	0,94	0,94	0,97	**1,00**	**1,00**	0,97	**1,00**	**1,00**	0,97	0,91	0,91	0,97	**1,00**	**1,00**	0,97	0,94	0,94	**1,00**
C21 - Galvalume A	**1,00**	**1,00**	0,65	0,90	0,91	0,63	0,96	0,97	0,47	0,98	0,98	0,69	0,82	0,83	0,48	**1,00**	**1,00**	0,61
C20 - Fibrocimento CRFS	**0,40**	0,41	0,68	0,43	0,44	0,74	0,49	0,50	0,68	**0,37**	**0,38**	0,79	0,42	0,43	0,63	0,40	0,41	0,84
C19 - Fibrocimento Amianto	0,84	0,86	0,80	0,78	0,79	0,75	0,89	0,90	0,75	0,84	0,85	0,79	0,88	0,89	0,70	0,89	0,89	0,73
C18 - Fibra de Vidro	0,41	0,42	**1,00**	**1,00**	**1,00**	0,60	**0,29**	**0,29**	**1,00**	**0,20**	**0,20**	0,60	**0,29**	**0,29**	**1,00**	**0,37**	**0,38**	0,60
C17 - Concreto Pintado de Vermelho	**1,00**	**1,00**	**1,00**	0,92	0,93	**1,00**	0,89	0,89	**1,00**	0,89	0,89	**1,00**	0,96	0,96	**1,00**	0,96	0,96	**1,00**
C16 - Concreto Pintado de Amarelo	0,50	0,50	**1,00**	0,54	0,55	**1,00**	0,43	0,43	**1,00**	0,54	0,55	**1,00**	**0,37**	**0,38**	**1,00**	**0,33**	**0,33**	**1,00**
C15 - Concreto Impermeabilizado	0,82	0,83	0,62	0,69	0,70	0,59	0,76	0,77	0,51	0,95	0,95	0,51	0,75	0,76	0,49	0,86	0,86	0,49
C14 - Concreto	0,40	0,46	**0,19**	0,42	0,48	**0,34**	0,47	0,52	**0,23**	0,50	0,54	**0,29**	0,47	0,52	**0,23**	0,47	0,52	**0,30**
C13 - Cerâmica	0,53	0,54	0,88	0,49	0,50	0,88	0,51	0,52	0,88	0,53	0,54	0,82	**0,40**	0,41	0,76	0,43	0,43	0,76
C12 - Carvão Coque	0,95	0,95	**1,00**	0,87	0,87	**1,00**	0,66	0,67	**1,00**	0,77	0,77	**1,00**	0,74	0,74	**1,00**	0,87	0,87	**1,00**
C11 - Bloquete	**0,20**	**0,20**	0,67	0,57	0,57	0,67	**0,24**	**0,24**	**1,00**	0,46	0,46	**1,00**	**0,19**	**0,19**	0,67	**0,31**	**0,31**	0,83
C10 - Asfalto Pintado de Vermelho	0,92	0,93	0,83	0,93	0,93	0,87	0,83	0,83	0,83	0,82	0,83	0,80	0,91	0,92	0,73	0,82	0,82	0,77
C9 - Asfalto Pintado de Branco	**0,29**	**0,31**	0,68	**0,18**	**0,20**	0,41	0,56	0,57	0,73	**0,24**	**0,25**	0,55	**0,30**	**0,32**	0,55	**0,21**	**0,22**	0,45
C8 - Asfalto	0,88	0,88	0,82	0,82	0,83	0,81	0,92	0,93	0,76	0,79	0,80	0,77	0,87	0,88	0,77	0,84	0,86	0,77
C7 - Argamassa de Cimento Pintado de Vermelho	**1,00**	**1,00**	0,96	0,92	0,93	0,96	0,92	0,93	0,96	0,96	0,96	0,96	0,92	0,93	0,96	0,89	0,89	0,96
C6 - Argamassa de Cimento Pintado de Verde	**1,00**	**1,00**	**1,00**	**1,00**	**1,00**	**1,00**	0,93	0,93	**1,00**	**1,00**	**1,00**	**1,00**	0,96	0,96	**1,00**	**1,00**	**1,00**	**1,00**
C5 - Argamassa de Cimento	0,44	0,45	0,64	**0,36**	**0,38**	0,55	**0,32**	**0,33**	0,45	**0,35**	**0,36**	0,59	**0,36**	**0,37**	0,50	**0,37**	**0,38**	0,45
C4 - Acrílico	0,77	0,78	0,78	0,78	0,79	0,61	0,75	0,75	**1,00**	0,84	0,85	0,61	0,77	0,78	0,78	0,62	0,63	0,56
C3 - Aço Galvanizado Fosco	0,70	0,71	**0,39**	0,69	0,70	**0,37**	0,72	0,73	**0,33**	0,68	0,70	0,40	0,55	0,57	**0,37**	0,69	0,70	**0,37**
C2 - Aço Galvanizado com Ferrugem	**0,35**	**0,36**	0,91	0,80	0,80	0,73	**0,38**	**0,38**	0,91	0,55	0,56	0,91	**0,39**	0,40	0,91	0,55	0,56	0,91
C1 - Aço Galvanizado Brilhante	**0,38**	**0,40**	0,88	**0,37**	**0,38**	0,83	0,54	0,55	0,67	0,42	0,44	0,71	0,43	0,44	0,63	**0,40**	0,41	0,75

Fonte: Produção do autor.

Dentre as classes de cobertura de Aço, *Aço Galvanizado Fosco* foi a classe com o maior erro de omissão e que apresentou o melhor K-classe ($K = 0,55$ ~$0,72$). Hipoteticamente, isto pode ser devido ao fato de que seus segmentos apresentam pigmentação majoritariamente azul, acarretando maior reflectância na B2 (Azul) do WV-2. Os índices favorecidos pelos classificadores para essa classe foram quase que exclusivamente espectrais, com exceção dos cenários que permitiram MDA. Os índices favoritos foram WVBI, WVSI, (BRT-B2)/B5 e NDSI.

O favorecimento dos índices e razões de banda em detrimento do MDA sugere que a classe *Aço Galvanizado Fosco* apresenta maior poder de generalização pelos classificadores do que as outras classes de cobertura de aço galvanizado com ferrugem e brilhante. Mas ela também apresenta confusões severas com materiais que fazem uso da B2 para sua separabilidade, como: *Argamassa de Cimento, Concreto Impermeabilizado, Fibrocimento CRFS, Plástico* e *Policarbonato.*

A classe *Acrílico* apresentou comportamento satisfatório no que diz respeito à acurácia do produtor, acurácia do usuário e K-classe ($K = 0,62$ no pior cenário: C5.0 sem MDA, e $K = 0,84$ no melhor: C4.5 também sem MDA). A classe *Acrílico* permite observar que a presença do MDA nos atributos também pode prejudicar a precisão dos classificadores e levar a ajustes excessivos do modelo em algumas regras (*overfit*). Isso ocorre porque a área de estudo apresenta um padrão de elevação natural, permitindo quase que um valor de MDA exclusivo para cada segmento. Esse padrão de elevação é observado na linha vermelha, resultante de uma regressão simples dos níveis do cinza do MDA (Figura 4.11b), extraídos de um transecto arbitrariamente traçado no meio do recorte da imagem da área de estudo (Figura 4.11a).

Figura 4.11 - Transecto perfilado sobre a área de estudo (a), descrevendo os níveis de cinza do MDA em uma curva de tendência em vermelho (b).

Fonte: Produção do autor.

Um total de 10 classes apresentou diversos problemas de separabilidade, sendo estas: *Argamassa de Cimento, Aço Galvanizado Brilhante, Aço Galvanizado com Ferrugem, Asfalto Pintado de Branco, Bloquete, Concreto Pintado de Amarelo, Lona, Pastilha Esmaltada, Piscina de Vinil* e *Policarbonato*. Os três principais problemas para essas classes foram identificados nas árvores de decisão como sendo:

- Similaridade espectral, impedindo o classificador de agrupar uma quantidade significativa de amostras em uma ou duas regras fortes. Um exemplo é a classe Argamassa de Cimento que apresenta apenas 11 segmentos (de 22) resguardados por quatro regras. Essas regras disputam espaço na árvore com diversas outras classes dentro de sua ramificação e, dentre outros, favorecem os atributos `Asymmetry`, `Border_Index`, `Média_no_MDA` e `Mínimo_valor_de_píxel_na_PC5` para classificar respectivamente 3,3,3 e 2 segmentos;

- Mistura espectral: segmentos de classes como *Pastilha Esmaltada* e as coberturas de Aço são muito finos e muitas vezes ocorrem nas extremidades das edificações, isso faz com que os segmentos se contaminem espectralmente com as classes adjacentes e reduzam o poder de generalização das regras do classificador;

- Déficit de amostras: quantidades insuficientes de segmentos para fornecer informação confiável ao classificador durante o processo de construção das árvores. As classes mais afetadas por isso são: *Policarbonato*, *Concreto Pintado de Amarelo, Fibra de Vidro, Lona, Aço Galvanizado com Ferrugem, Argamassa de Cimento, Piscina de Vinil, Pastilha Esmaltada* e *Bloquete*. Todos com $K \leq 0,40$.

Também apresentam escassez de amostras: *Quartzito Amarelo, Acrílico, Argamassa de Cimento Pintado de Vermelho, Plástico, Carvão Coque, Argamassa de Cimento Pintado de Verde, Vidro Aramado, Piscina de Azulejo* e *Concreto Pintado de Vermelho*. A deficiência de amostras não permite isentar essas classes de possíveis *overfit* ou *underfit*, mas seus indicadores de precisão foram satisfatórios ($K \geq 0,50$), possivelmente por sua exclusividade espectral ou valores de elevação do MDA, apresentando comportamentos distintos nesses atributos que facilitam sua separabilidade, conforme pode ser observado em suas respectivas tipologias na Tabela 3.2.

Para os casos das classes *Argamassa de Cimento Pintado de Verde* e *Argamassa de Cimento Pintado de Vermelho*, seu bom comportamento não surpreende. Essas clas-

ses são referentes às quadras esportivas de tênis, e suas propriedades espectrais são distintas o suficiente para que regras como `média_na_B5` (vermelho) sejam capazes de enquadrar todos os casos da classe Argamassa de Cimento Pintado de Vermelho na árvore de decisão proposta pelo algoritmo C5.0 com MDA.

As classes de *Asfalto* e *Asfalto Pintado de Vermelho* apresentaram boa separabilidade. Sua baixa reflectância na B1 (Azul Costeiro) e alto brilho no vermelho (por conta da pintura das ciclo-faixas) permitiram regras com relativo poder de generalização e discriminação de suas amostras. Mas em simultâneo, o principal motivo de erros para a classe Asfalto foi justamente a sua baixa reflectância, que gerou confusões sutis com as classes *Carvão Coque* (material de composição da pista de corrida) e *Sombra*.

O número de regras utilizadas pelos algoritmos diverge sobremaneira para diversas classes, mas ressalva especial deve ser feita para o CART com e sem a utilização do MDA, especificamente para a classe *Cerâmica*. Enquanto os algoritmos C4.5 e C5.0 (com e sem MDA) utilizam de 5 a 7 regras "fracas" para separar esta classe, o algoritmo CART fez uso de apenas 2 regras "fortes", em que "forte" e "fraco" são adjetivos que expressam a quantidade de segmentos cobertos pelas regras. Uma possível justificativa para tal comportamento do CART poderia ser seu mecanismo de redução da variância, que não existe no C4.5 e C5.0. Com a variância reduzida, as amostras poderiam ser separadas através de regras mais generalistas, capazes de enquadrar mais casos de classes. Regras "fracas" dizem respeito à utilização de diversas regras que cobrem poucos casos na árvore, ao passo que regras fortes enquadram grandes quantidades ou até mesmo todos os segmentos de amostras utilizadas para o treinamento do classificador. Cabe observar que não se pretende aqui fazer juízo de qual abordagem é melhor, principalmente porque ambas abordagens são necessárias na construção das árvores e algumas classes são espectralmente mais complexas, exigindo mais regras. Tais observações são feitas apenas por mérito de compreender diferenças pontuais entre o comportamento dos algoritmos C4.5, C5.0 e CART.

Ainda em relação à classe *Cerâmica*, esta se comportou de forma regular com apenas alguns casos de comissão pela classe *Solo Exposto*. Esse fenômeno também é recorrente na literatura e ocorre porque ambos materiais compartilham de diversas propriedades físico-químicas (HEROLD et al., 2004; RIBEIRO, 2010), acarretando um comportamento espectral muito similar, observado nas amostras ilustradas na Figura 4.12.

Figura 4.12 - Reflectância espectral mensurada no eixo Y para as amostras de Cerâmica (a) e Solo Exposto (b) com base nas oito bandas do WV-2.

Fonte: Produção do autor.

A Figura 4.13 também apresenta a comparação das medidas realizadas em laboratório (LARAD) sobre uma telha portuguesa, representante da classe *Cerâmica* (Figura 4.13a) e sua contraparte mensurada na imagem WV-2 e suas respectivas bandas após a correção atmosférica (Figura 4.13b), para este mesmo alvo. Cabe notar que o gráfico da Figura 4.13a foi ajustado para máxima reflectância em 0,8, porque este apresenta medidas de reflectância inferiores aos valores mensurados na Figura 4.13b, possivelmente porque, de forma diversa da imagem orbital, o LARAD não permite a interferência de REM proveniente de fontes externas no sensor.

Figura 4.13 - Reflectância mensurada em laboratório (a) versus a reflectância normalizada para os valores da imagem WV-2 (b), ambas para a classe Cerâmica.

Fonte: Produção do autor.

A classe *Vegetação Arbórea* também se comportou de maneira satisfatória, apresentando poucas confusões com as classes *Vegetação Herbácea* e *Solo Exposto*. Acredita-se que as confusões com *Vegetação Herbácea* sejam provenientes de dois principais fatores. O primeiro, ilustrado na Figura 4.14, seria a inexistência de uma terceira classe intermediária para categorizar vegetações arbustivas (A), que são maiores que a vegetação rasteira (B), mas que ainda são menos densas do que uma vegetação arbórea madura (C).

Figura 4.14 - Representantes da classe de Vegetação Arbórea que potencialmente contribuem para confusão com as classes de Vegetação Herbácea e Solo Exposto.

Fonte: Produção do autor.

Não é pretensão deste trabalho discriminar as variadas categorias de vegetação, mas apenas compreender aspectos que poderiam ser considerados responsáveis pela confusão no resultado das matrizes. O segundo fator de influência também poderia ser um aspecto da Vegetação Arbórea, quando a ramificação das árvores se desenvolve de forma esparsa e o dossel não apresenta densidade suficiente para cobrir o solo (palmeiras ou outras plantas ornamentais). Isso quer dizer que, quando observado a *nadir*, a vegetação das copas apresenta espaços que permitem ver o solo abaixo, conforme ilustrado por (D) na Figura 4.14. E Como se trata de um sensor de alta resolução espacial (0,50 m na banda pancromática), esses espaços no dossel contaminam os segmentos de *Vegetação Arbórea* com atributos e propriedades espectrais de *Vegetação Herbácea* e até mesmo *Solo Exposto*. Mas além do resultado das matrizes, experimentos específicos são sugeridos para testar esse fenômeno.

A classe *Solo Exposto* também apresentou confusão com *Vegetação Herbácea*, e esse fenômeno pode ser claramente observado na pigmentação da grama dentro do quadrado (B) na Figura 4.14. Tal semelhança espectral é resultado de uma estiagem severa no ano de 2010 (CEPAGRI, 2015; G1, 2014), mesmo ano de aquisição da imagem WV-2. Como consequência desse evento, a literatura também sugere que pode haver contribuição severa para similaridade espectral entre essas classes (CLARK, 1999), apresentadas na Figura 4.15 e corroborados pelas amostras apresentadas na Figura 4.16.

Figura 4.15 - Similaridade espectral entre alvos de vegetação seca e solo exposto.

Fonte: Adaptado de Clark (1999).

Figura 4.16 - Comparação da reflectância (eixo Y) para as amostras de Solo (A) e Vegetação Herbácea (B).

Fonte: Produção do autor.

Concreto foi a classe de maior confusão no trabalho em todos os cenários de classificação, o que não significa que esta classe não seja passiva de discernimento, conforme demonstrado por Ribeiro (2010), que separou os aspectos de intemperismo desse material em subcategorias (concreto claro, médio e escuro), melhorando sobremaneira os resultados de sua classificação. Herold et al. (2004) justificam parte dessa confusão pelo comportamento não lambertiano de retroespalhamento da REM por materiais sintéticos de composições mineralógicas e rugosidade afins, como coberturas baseadas em asfalto, concretos e materiais aluminizados. Tal resultado sugere que a classe Concreto não é espectralmente robusta como *Cerâmica* para suportar generalizações intraclasse que permitam a inclusão de diferentes intemperismos no mesmo conjunto amostral de treinamento dos algoritmos.

Em contrapartida, a classe *Concreto Impermeabilizado* apresentou bons resultados de separabilidade e menos confusões nas matrizes. Experimentos com esse material no LARAD sugerem que a inclusão de uma camada de resina para impermeabilização nas telhas feitas com esse material não modificam seu comportamento espectral (APÊNDICE E), e então, a disparidade entre os resultados positivos de *Concreto Impermeabilizado* contra as confusões na matriz para classe *Concreto* foram atribuídas ao MDA, uma vez que os segmentos de *Concreto Impermeabilizado* ocorrem exclusivamente acima das edificações.

Fibrocimento com Amianto e *Fibrocimento CRFS* apresentaram bons resultados na matriz, e a confusão mútua é quase inexistente para todos os cenários de classificação. O principal motivo para essa distinção pode ser atribuído ao intemperismo das telhas de amianto. Uma peculiaridade desse material foi o fato de ter sua venda proibida, e quando elas se quebram, os técnicos da manutenção da UNICAMP precisam substituir as telhas de amianto por telhas de fibrocimento CRFS, que apresentam condições muito superiores de conservação, conforme ilustrado na Figura 4.17 (a parte branca representa o espaço fora da área de estudo).

Figura 4.17 - Contraste entre as velhas telhas de fibrocimento amianto e as novas de CRFS, dispostos em um recorte do transecto na área de estudo.

Fonte: Produção do autor.

O algoritmo C5.0 também fez uma escolha peculiar para separar os segmentos de *Fibrocimento CRFS* ao optar pelo atributo `Assymetry`, uma vez que este índice no eConition é calculado através da razão simples entre a largura e o comprimento de um dado segmento. E os segmentos dessa classe, em sua maioria, embora não exclusivamente, apresentam perfis alongados na vertical e estreitos na horizontal, que podem ser observados nos polígonos em laranja da Figura 4.17.

Embora compartilhem de propriedades físico-químicas similares, as classes de *Galvalume A, B* e *C* ainda são espectralmente distintas. Todas apresentaram bons resultados de classificação, mesmo quando não utilizados os atributos do MDA. O modelo gerado pelo algoritmo C5.0 sem MDA favoreceu o atributo `mínimo_valor_pixel_PC4` para separar a classe *Galvalume C*, a razão (BRT+B2)/B5 denominada BRTpBoR para separar *Galvalume A*, e o NDSI1 para

Galvalume B. A escolha desses atributos foi repetida também nos modelos C4.5 e CART sem MDA, com exceção do atributo `máximo_valor_pixel_B7` (NIR1 do WV-2), utilizado pelo CART sem MDA para separar os segmentos de *Galvalume A*, que também é quase equivalente à utilização da quarta componente principal, uma vez que essa recebe mais de 92% de contribuição das bandas 7 e 8 (NIR1 e NIR2) em sua composição, conforme previamente ilustrado na Tabela 4.6.

Por fim, é encerrada a discussão das matrizes de confusão, sendo apresentados os resultados para o primeiro e segundo nível de legenda de classificação, nas Figura 4.18 e Figura 4.19, respectivamente.

Na sequência, também é apresentada na Figura 4.20 uma imagem que ilustra a discordância entre os classificadores, de forma a auxiliar o leitor na identificação de diferenças entre os resultados dos mapas temáticos.

Figura 4.18 - Classificações que apresentaram os maiores coeficientes Kappa com e sem o MDA para o Nível de Legenda I.

Fonte: Produção do autor.

Figura 4.19 - Classificações acusadas como mais distintas através do teste Z com $p < 0,05$ para o Nível de Legenda II.

Fonte: Produção do autor.

Figura 4.20 - Mapa de divergência entre os classificadores dos Níveis I e II.

Fonte: Produção do autor.

5 CONCLUSÕES

Neste trabalho, foram explorados métodos de mineração de dados, refinados pelo algoritmo híbrido de boosting (AdaBoost.M1) e conjugados a análise de imagem baseada em objeto, para fins de classificação de cobertura do solo urbano em um transecto do Campus da UNICAMP (SP) a partir de imagens provenientes do sensor WV-2. Foram analisados três algoritmos de árvore de decisão C4.5 (J48), C5.0 (See5) e CART em dois níveis de legenda de classificação, com e sem o auxílio de um modelo digital de altura (MDA), gerando como resultado 12 cenários de classificação para uma mesma imagem da área de estudo.

O sensor WV-2 se mostrou eficiente na classificação muito fina de imagens de SR urbano, mas a carência de amostras para classes como Pastilha Esmaltada, Lona, Bloquete e Piscina de Vinil sugerem que não seja recomendada a classificação das mesmas até que problemas como a mistura espectral dentro desses segmentos seja sanada, possivelmente por novos sensores que apresentem uma resolução espacial mais fina, como o WV-3 e, muito em breve, o WV-4 e o GeoEye-2.

O algoritmo de boosting se mostrou eficiente nos 12 experimentos de classificação, auxiliando os classificadores a atingirem valores de exatidão temática nas árvores de decisão de até 0,74 para 12 classes de legenda (Nível I) e 0,72 para as 36 classes do Nível II. A influência do AdaBoost.M1 sobre os algoritmos J48, See5 e CART só pode ser observada através da comparação de resultados com e sem a aplicação do AdaBoost.M1 para um mesmo classificador, sendo esta uma sugestão para trabalhos futuros.

Sobre a influência do MDA nas classificações, a baixa variação nos índices de concordância Kappa para as legendas dos Níveis I e II, em todos os cenários de classificação (C4.5, C5.0 e CART, com e sem MDA), sugerem que a presença dos atributos derivados do MDA, apesar de em sua maioria exercerem influência positiva nas classificações, não produzem incremento sensível nos índices Kappa. Cabe notar, por outro lado, que as classes que ocorrem no topo das edificações, como Vidro Aramado, Cerâmica e Fibrocimento CRFS, se beneficiam de forma significativa dos atributos provenientes do modelo digital de altura.

No que diz respeito à análise comparativa, é possível concluir que o modelo CART com MDA gerou a árvore mais compacta. Com 62 nós-folha e tamanho total da árvore de 123 nós, CART com MDA fez uso aproximado de duas regras por classe para as 36 classes do Nível de Legenda II, contra uma média de três a quatro regras

por classe dos modelos gerados por C4.5 (229 nós com MDA e 251 sem MDA) e C5.0 (233 nós com MDA e 237 sem).

Uma surpresa positiva foi o tempo de processamento dos modelos gerados pelo C5.0, sucessor do C4.5. Ele apresentou tempo médio de 20 segundos para classificar todos os segmentos da imagem após a confecção do modelo, duas ordens de magnitude mais rápido que os algoritmos C4.5 e CART, que levaram pouco mais de duas horas para cada rodada dos modelos de classificação das legendas do Nível II. Este fenômeno pode ser atribuído a forma como o algoritmo C5.0 foi implementado, permitindo Hyper-Threading (hiperprocessamento) com diversos núcleos do processador.

A utilização de uma área de estudo restrita permitiu a redução de custos e complexidade do trabalho, mas também é sugerida a experimentação dos modelos em outras áreas da mesma cena. Uma vez que os classificadores se mostraram robustos, mas algumas classes apresentaram uma baixa quantidade de amostras, testar os modelos aqui apresentados em uma outra área da imagem seria uma forma de verificar a presença de classes que poderiam apresentar ajustes excessivos aos modelos de clasificação.

Por fim, conclui-se que o objetivo geral e os objetivos específicos propostos foram atingidos de forma satisfatória. Algumas sugestões para correção e otimização dos resultados deste e de outros trabalhos que façam uso de metodologias similares são:

a) Testar a resiliência a outliers dos algoritmos classificadores aqui utilizados, embora todos apresentem mecanismos contra valores espúrios;

b) Avaliar a qualidade cartográfica do MDS por meio de testes de tendência (ou deslocamento sistemático) e precisão;

c) Verificar a influência da correção atmosférica, correção topográfica, pureza das amostras na acurácia das classificações e testar também os métodos híbridos de otimização e classificação, respectivamente bagging e blending.

REFERÊNCIAS BIBLIOGRÁFICAS

ACHANCCARAY, P.; AYMA, V. A.; JIMENEZ, L.; GARCIA, S.; HAPP, P.; FEITOSA, R. Q.; PLAZA, A. A free software tool for Automatic Tuning of Segmentation Parameters. **South-Eastern European Journal of Earth Observation and Geomatics**, v. 3, p. 707–712, 2014. 31

ADDINK, E. A.; KLEINHANS, M. G. Recognizing meanders to reconstruct river dynamics of the Ganges. In: **GEOBIA 2008 - Pixels, Objects, Intelligence: GEOgraphic Object Based Image Analysis for the 21st Century**. [S.l.: s.n.], 2008. University. 13

BAATZ, M.; SCHAPE, A. Multiresolution Segmentation : an optimization approach for high quality multi-scale image segmentation. **Journal of Photogrammetry and Remote Sensing**, v. 58, p. 12–23, 2000. 13, 29

BALTSAVIAS, E. P.; GRUEN, A.; VANGOOL, L. **Automatic Extraction of Man-made Objects from Aerial and Satellite Images III**. Taylor & Francis, 2001. (Monte Verit{à}, v. 3). ISBN 9789058092526. Disponível em: <https://books.google.com.br/books?id=kjI4k58eOl4C>. 12

BELLMAN, R. **Dynamic Programming**. [s.n.], 1957. 140–153 p. ISSN 01628828. ISBN 0486428095. Disponível em: <http://scholar.google.com/scholar?hl=en{&}btnG=Search{&}q=intitle:Adaptive+Dynamic+programming{#}0>. 16

BENZ, U. C.; HOFMANN, P.; WILLHAUCK, G.; LINGENFELDER, I.; HEYNEN, M. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 58, n. 3-4, p. 239–258, jan 2004. ISSN 09242716. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0924271603000601>. 13

BLASCHKE, T. Object based image analysis for remote sensing. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 65, n. 1, p. 2–16, jan 2010. ISSN 09242716. Disponível em: <http://linkinghub.elsevier.com/retrieve/pii/S0924271609000884>. 14

BLASCHKE, T.; HAY, G. J.; KELLY, M.; LANG, S.; HOFMANN, P.; ADDINK, E.; Queiroz Feitosa, R.; MEER, F. van der; WERFF, H. van der; COILLIE, F.

van; TIEDE, D. Geographic Object-Based Image Analysis - Towards a new paradigm. **ISPRS journal of photogrammetry and remote sensing : official publication of the International Society for Photogrammetry and Remote Sensing (ISPRS)**, v. 87, n. 100, p. 180–191, jan 2014. ISSN 0924-2716. Disponível em: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3945831{&}tool=pmcentrez{&}rendertype=abstract>. 5, 14, 15

BRAMANTE, J. F.; RAJU, D. K.; SIN, T. M. Multispectral derivation of bathymetry in Singapore's shallow, turbid waters. **International Journal of Remote Sensing**, Taylor & Francis, v. 34, n. 6, p. 2070–2088, mar 2013. ISSN 0143-1161. Disponível em: <http://www.tandfonline.com/doi/full/10.1080/01431161.2012.734934{#}.VNWD2nXN-00>. 6

BREIMAN, L.; FRIEDMAN, J. H.; OLSHEN, R. A.; STONE, C. J. **Classification and Regression Trees**. [s.n.], 1984. 368 p. ISBN 0412048418. Disponível em: <http://scholar.google.com/scholar?hl=en{&}btnG=Search{&}q=intitle:Classification+and+regression+trees{#}0>. 53

BRICE, C. R.; FENNEMA, C. L. Scene analysis using regions. **Artificial Intelligence**, v. 1, n. 3-4, p. 205–226, jan 1970. ISSN 00043702. Disponível em: <http://www.sciencedirect.com/science/article/pii/0004370270900081>. 13

BURNETT, C.; BLASCHKE, T. A multi-scale segmentation/object relationship modelling methodology for landscape analysis. In: **Ecological Modelling**. [S.l.: s.n.], 2003. v. 168, p. 233–249. ISBN 0304-3800. ISSN 03043800. 13

CAMARGO, F. F.; ALMEIDA, C. M.; FLORENZANO, T. G.; HEIPKE, C.; FEITOSA, R. Q.; COSTA, G. Aster/Terra imagery and a multilevel semantic network for semi-automated classification of landforms in a subtropical area. **Photogrammetric Engineering and Remote Sensing**, v. 77, p. 619–629, 2011. ISSN 00991112 (ISSN). Disponível em: <http://www.scopus.com/inward/record.url?eid=2-s2.0-79960384003{&}partnerID=40{&}md5=b18093e718a18b171134631111761d96>. 13

CEPAGRI. **Centro de Pesquisas Meteorologicas e Climaticas Aplicadas a Agricultura**. 2015. Disponível em: <http://www.cepagri.unicamp.br/>. 102

CHANDER, G.; MARKHAM, B. L.; HELDER, D. L. Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. **Remote Sensing of Environment**, v. 113, p. 893–903, 2009. ISSN 00344257. 9

CHANDRASEKHAR, S. **Radiative Transfer**. [S.l.: s.n.], 1960. 393 p. ISBN 0486605906. 8

CHEN, S.; SU, H.; ZHANG, R.; TIAN, J.; YANG, L. The Tradeoff Analysis for Remote Sensing Image Fusion Using Expanded Spectral Angle Mapper. **Sensors**, v. 8, n. 1, p. 520–528, jan 2008. ISSN 1424-8220. Disponível em: <http://www.mdpi.com/1424-8220/8/1/520/>. 1, 5

CHORAŚ, M.; BHANU, B.; CHEN, H.; CHAMPOD, C.; KOMATSU, N.; NAKANO, M.; CAVOUKIAN, A.; STOIANOV, A.; ZHOU, Z.-H.; KUKULA, E. P.; ELLIOTT, S. J.; SARKAR, S.; LIU, Z. Ensemble Learning. In: LI, S. Z.; JAIN, A. (Ed.). **Encyclopedia of Biometrics**. Boston, MA: Springer US, 2009. cap. Ensemble L, p. 270–273. ISBN 978-0-387-73003-5. Disponível em: <http://dx.doi.org/10.1007/978-0-387-73003-5{_}293http://www.springerlink.com/index/10.1007/978-0-387-73003-5{_}293>. 56

CLARK, R. N. Spectroscopy of rocks and minerals, and principles of spectroscopy. **Remote sensing for the earth sciences: Manual of remote sensing**, v. 3, p. 3–58, 1999. 102, 103

COHEN, J. A coefficient of agreement of nominal scales. **Educational and Psychological Measurement**, v. 20, p. 37–46, 1960. ISSN 0013-1644. Disponível em: <http://epm.sagepub.com/cgi/doi/10.1177/001316446002000104>. 16

CONGALTON, R. G. **A review of assessing the accuracy of classifications of remotely sensed data**. jul 1991. 35–46 p. Disponível em: <http://www.sciencedirect.com/science/article/pii/003442579190048B>. 38, 60, 61

CONGALTON, R. G.; GREEN, K. **Assessing the Accuracy of Remotely Sensed Data: Principles and Practices**. [s.n.], 2002. ISBN 1420048562. Disponível em: <http://books.google.com.br/books/about/Assessing{_}the{_}Accuracy{_}of{_}Remotely{_}Sense.html?id=THm9VJ4O-iQC{&}pgis=1>. 61

_____. _____. CRC Press/Taylor & Francis, 2009. (Mapping Science Series). ISBN 9781420055122. Disponível em: <http://books.google.com.br/books?id=FKssAQAAMAAJ>. 85

DIGITALGLOBE. **WorldView-2 Data Sheet**. 2009. 2 p. Disponível em: <https://www.digitalglobe.com/sites/default/files/DG{_}WorldView2{_}DS{_}PROD.pdf>. 5, 6

DRURY, S. A. **Image Interpretation in Geology**. [S.l.: s.n.], 1993. p. 283 p. ISSN 10106049. ISBN 0412488809. 94

DUDA, R. O.; HART, P. E.; STORK, D. G. **Pattern Classification**. [S.l.: s.n.], 2000. 654 p. ISSN 1740634X. ISBN 978-0-471-05669-0. 15

DURIEUX, L.; LAGABRIELLE, E.; NELSON, A. A method for monitoring building construction in urban sprawl areas using object-based analysis of Spot 5 images and existing GIS data. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 63, p. 399–408, 2008. ISSN 09242716. 13

FAHSI, A.; TSEGAYE, T.; TADESSE, W.; COLEMAN, T. Incorporation of digital elevation models with Landsat-TM data to improve land cover classification accuracy. **Forest Ecology and Management**, v. 128, n. 1-2, p. 57–64, 2000. ISSN 03781127. 2

FAYYAD, U.; PIATETSKY-SHAPIRO, G.; SMYTH, P. From data mining to knowledge discovery in databases. **AI magazine**, v. 17, p. 37–53, 1996. ISSN 0738-4602. 15

FRANCA, D. G. M.; LOTTE, R. G.; ALMEIDA, C. M. D.; SIANI, S. M. O.; KORTING, T. S.; FONSECA, L. M. G.; SILVA, L. T. da. Object-based image analysis for urban land cover classification in the city of Campinas - SP, Brazil. In: **JURSE - Joint Urban Remote Sensing Event**. Lausanne, Switzerland: [s.n.], 2015. p. 4. 49

FREUND, Y.; SCHAPIRE, R. E. A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. **Journal of Computer and System Sciences**, v. 55, n. 1, p. 119–139, aug 1997. ISSN 00220000. 58, 59

G1. **Ano de 2013 foi o mais seco desde 1989 em Campinas**. 2014. Disponível em: <http://g1.globo.com/sp/campinas-regiao/noticia/2014/01/ano-de-2013-foi-o-mais-seco-desde-1989-em-campinas-diz-cepagri.html>. 102

GEOMATICA. **PAN Sharpening**. 2013. 2 p. Disponível em: <http://www.pcigeomatics.com/pdf/PanSharp.pdf>. 10

GRODECKI, J.; DIAL, G. IKONOS geometric accuracy. In: **Proceedings of joint workshop of ISPRS working groups I/2, I/5 and IV/7 on high resolution mapping from space**. [s.n.], 2001. p. 19–21. Disponível em: <http://satimagingcorp.s3.amazonaws.com/site/pdf/IKONOSGeometricAccuracy-ISPRS202001.pdf>. 13

HALL, M.; FRANK, E.; HOLMES, G.; PFAHRINGER, B.; REUTEMANN, P.; WITTEN, I. H. **The WEKA data mining software**. 2009. 10 p. Disponível em: <http://portal.acm.org/citation.cfm?doid=1656274.1656278$\delimiter"026E30F$npapers2://publication/doi/10.1145/1656274.1656278>. 16

HAMEDIANFAR, A.; SHAFRI, H. Z. M.; MANSOR, S.; AHMAD, N. Improving detailed rule-based feature extraction of urban areas from WorldView-2 image and lidar data. **International Journal of Remote Sensing**, v. 35, p. 1876–1899, 2014. ISSN 0143-1161. Disponível em: <http://dx.doi.org/10.1080/01431161.2013.879350>. 6

HARALICK, R. M.; SHAPIRO, L. G. Image segmentation techniques. **Computer Vision, Graphics, and Image Processing**, v. 29, n. 1, p. 100–132, jan 1985. ISSN 0734189X. Disponível em: <http://linkinghub.elsevier.com/retrieve/pii/S0734189X85901537>. 13

HATTORI, S.; ONO, T. Orientation of high-resolution satellite images based on affine projection. **International Archives of Photogrametry and Remote Sensing**, XXXIII, 2000. Disponível em: <http://www.isprs.org/proceedings/XXXIII/congress/part3/359_XXXIII-part3.pdf>. 13

HAY, G. J.; MARCEAU, D. J.; DUBÉ, P.; BOUCHARD, A. A multiscale framework for landscape analysis: Object-specific analysis and upscaling. **Landscape Ecology**, v. 16, p. 471–490, 2001. ISSN 09212973. 13

HEROLD, M.; GARDNER, M.; HADLEY, B.; ROBERTS, D. The spectral dimension in urban land cover mapping from high-resolution optical remote sensing data. In: **Symposium A Quarterly Journal In Modern Foreign Literatures**. [s.n.], 2002. v. 6, p. 8 pp. Disponível em: <http://www.geogr.uni-jena.de/{~}c5hema/pub/istanb{_}herold{_}gardn{_}hadl{_}roberts.pdf>. 1

HEROLD, M.; ROBERTS, D. a.; GARDNER, M. E.; DENNISON, P. E. Spectrometry for urban area remote sensing - Development and analysis of a spectral library from 350 to 2400 nm. **Remote Sensing of Environment**, v. 91, n. 3-4, p. 304–319, jun 2004. ISSN 00344257. Disponível em: <http://linkinghub.elsevier.com/retrieve/pii/S0034425704000768>. 39, 90, 99, 104

HETTICH, S.; BAY, S. D. **The UCI KDD Archive**. 1999. Disponível em: <http://kdd.ics.uci.edu>. 52

HOFMANN, P.; STROBL, J.; BLASCHKE, T.; KUX, H. Detecting informal settlements from Quickbird data in Rio de Janeiro using an object based approach. In: **Object-Based Image Analysis**. [S.l.: s.n.], 2008. p. 531–553. ISBN 3540770577. 13

HUGHES, G. On the mean accuracy of statistical pattern recognizers. **IEEE Transactions on Information Theory**, v. 14, n. 1, p. 55–63, jan 1968. ISSN 0018-9448. Disponível em: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1054102>. 16

IBGE. **2010 census**. 2013. 1 p. Disponível em: <http://censo2010.ibge.gov.br/en/sobre-censo>. 1

_____. **Área Territorial - Consulta por Município**. 2013. Disponível em: <http://www.ibge.gov.br/home/geociencias/areaterritorial/area.shtm>. 19

JACOBSEN, K. Calibration aspects in direct georeferencing of frame imagery. **INTERNATIONAL ARCHIVES OF PHOTOGRAMMETRY REMOTE SENSING AND SPATIAL INFORMATION SCIENCES**, v. 34, p. 82–88, 2002. ISSN 1682-1750. 12

JENSEN, J. R. **Introductory Digital Image Processing: A Remote Sensing Perspective**. Prentice Hall, 2005. (Prentice Hall series in geographic information science). ISBN 9780131453616. Disponível em: <http://books.google.com.br/books?id=F84PAQAAIAAJ>. 14

_____. **Sensoriamento remoto do ambiente: uma perspectiva em recursos terrestres**. [s.n.], 2009. 587 p. ISBN 9788560507061. Disponível em: <http://www.parentese.com.br/pdf/jensen.pdf>. 5, 7, 11, 89

JENSEN, J. R.; COWEN, D. C. Remote sensing of urban/suburb an infrastructure and socio-economic attributes. **Photogrammetric Engineering and Remote Sensing**, v. 65, p. 611–622, 1999. ISSN 00991112. Disponível em: <HardCopy$\delimiter"026E30F$nhttp://www.scopus.com/scopus/inward/record.url?eid=2-s2.0-0032986310{&}partner=40{&}rel=R4.5.0>. 5

KAUFMAN, Y.; WALD, A.; REMER, L.; FLYNN, L. The MODIS 2.1-μm channel-correlation with visible reflectance for use in remote sensing of aerosol. **IEEE Transactions on Geoscience and Remote Sensing**, v. 35, n. 5, p. 1286–1298, 1997. ISSN 01962892. Disponível em: <http://adsabs.harvard.edu/abs/1997ITGRS..35.1286K>. 24

KIDNER, D.; DOREY, M.; SMITH, D. What's the point? Interpolation and extrapolation with a regular grid DEM. In: **IV International Conference on GeoComputation**. Fredericksburg, Virginia: [s.n.], 1999. Disponível em: <http://www.geovista.psu.edu/sites/geocomp99/Gc99/082/gc{_}082.htm>. 26

KUMAR, M.; MILLER, D. A. A NON-PARAMETRIC CLASSIFICATION STRATEGY FOR REMOTELY SENSED IMAGES USING BOTH SPECTRALAND TEXTURAL INFORMATION. 2006. Disponível em: <http://people.duke.edu/~mk176/publications/Kumar&Miller_Classification_ISPRS.pdf>. 2, 15

KUX, H. J. H.; ARAÚJO, E. H. G.; CARTWRIGHT, W.; GARTNER, G.; MENG, L.; PETERSON, M. P. Object-based Image Analysis using QuickBird satellite images and GIS data, case study Belo Horizonte (Brazil) Object-Based Image Analysis. In: **Object-Based Image Analysis**. Springer, 2008. p. 571–588. ISBN 978-3-540-77058-9. Disponível em: <http://dx.doi.org/10.1007/978-3-540-77058-9{_}31>. 13

LACERDA, C. S. d. A. **Classificação de áreas urbanas com imagens multiespectrais e hiperespectrais utilizando métodos não-paramétricos**. 380 p. Tese (Doutorado) — Instituto Nacional de Pesquisas Espaciais (INPE), 2016. 20, 26, 27, 33

LATIF, Z. A.; ZAMRI, I.; OMAR, H. Determination of tree species using Worldview-2 data. In: **2012 IEEE 8th International Colloquium on Signal Processing and its Applications**. IEEE, 2012. p. 383–387. ISBN 978-1-4673-0961-5. Disponível em: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6194754>. 6

LATORRE, M.; CARVALHO, A. O. J. D.; CARVALHO, A. P. F.; SHIMABUKURO, Y. E. CORREÇÃO ATMOSFÉRICA: CONCEITOS E FUNDAMENTOS. **Espaço & Geografia**, v. 5, p. 153–178, 2002. ISSN 1516-9375. 8

LEONARDI, F. **Abordagens cognitivas e mineração de dados aplicadas a dados ópticos orbitais e de laser para classificação de cobertura do solo urbano**. Dissertação (Mestrado) — Instituto Nacional de Pesquisas Espaciais, 2010. Disponível em: <http://urlib.net/8JMKD3MGP7W/376H8M2>. xvi, 2, 43, 76, 77

LIU, P. A Survey of Remote-sensing Big Data. **Frontiers in Environmental Science**, v. 3, n. June, p. 1–6, 2015. ISSN 2296-665X. Disponível em: <http://journal.frontiersin.org/article/10.3389/fenvs.2015.00045/abstract>. 48

MA, Y.; WU, H.; WANG, L.; HUANG, B.; RANJAN, R.; ZOMAYA, A.; JIE, W. Remote sensing big data computing: Challenges and opportunities. **Future Generation Computer Systems**, v. 51, p. 47–60, oct 2015. ISSN 0167739X. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0167739X14002234>. 48

MALLINIS, G.; KOUTSIAS, N.; TSAKIRI-STRATI, M.; KARTERIS, M. Object-based classification using Quickbird imagery for delineating forest vegetation polygons in a Mediterranean test site. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 63, p. 237–250, 2008. ISSN 09242716. 13

MEI, A.; FIORE, N.; SALVATORI, R.; D'ANDREA, A.; FONTANA, M. Spectroradiometric Laboratory Measures on Asphalt Concrete: Preliminary Results. **Procedia - Social and Behavioral Sciences**, v. 53, p. 514–523, 2012. ISSN 18770428. Disponível em: <http://www.sciencedirect.com/science/article/pii/S1877042812043649>. 87

MILTON, E. J. Review Article Principles of field spectroscopy. **International Journal of Remote Sensing**, v. 8, n. 12, p. 1807–1827, dec 1987. ISSN 0143-1161. Disponível em: <http://www.tandfonline.com/doi/abs/10.1080/01431168708954818>. 39

PEDRO, P. d. C.; ANTUNES, A. F. B.; MITISHITA, E. A. **ORTORETIFICAÇÃO DE IMAGENS DE ALTA RESOLUÇÃO**

UTILIZANDO OS MODELOS APM (AFFINE PROJECTION MODEL) E RPC (RATIONAL POLYNOMIAL COEFFYCIENT). jul 2007. Disponível em: <http://ojs.c3sl.ufpr.br/ojs2/index.php/bcg/article/view/8245>. 13

PINHO, C. M. D. d. **Análise orientada a objetos de imagens de satélites de alta resolução espacial aplicada à classificação de cobertura do solo no espaço intra-urbano: o caso de São José dos Campos-SP**. 178 p. Dissertação (Mestrado) — Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2005-09-23 2006. Disponível em: <http://urlib.net/sid.inpe.br/MTC-m13@80/2005/11.23.13.40>. Acesso em: 21 set. 2016. 2

PINHO, C. M. D. d. **Análise orientada a objetos de imagens de satélites de alta resolução espacial aplicada à classificação de cobertura do solo no espaço intra-urbano: o caso de São José dos Campos-SP**. 178 p. Dissertação (Mestrado) — Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2005-09-23 2006. Disponível em: <http://urlib.net/sid.inpe.br/MTC-m13@80/2005/11.23.13.40>. Acesso em: 04 out. 2016. 89, 90

PINHO, C. M. D. de; FONSECA, L. M. G.; KORTING, T. S.; ALMEIDA, C. M. de; KUX, H. J. H. Land-cover classification of an intra-urban environment using high-resolution images and object-based image analysis. **International Journal of Remote Sensing**, Taylor & Francis, v. 33, n. 19, p. 5973–5995, oct 2012. ISSN 0143-1161. Disponível em: <http://www.tandfonline.com/doi/abs/10.1080/01431161.2012.675451{#}.VNWHkHXN-00http://www.tandfonline.com/doi/abs/10.1080/01431161.2012.675451>. 6

PINHO, C. M. D. de; RENNÓ, C. D.; KUX, H. J. H. Avaliação de técnicas de fusão aplicadas à imagens Quickbird. In: **XII Simpósio Brasileiro de Sensoriamento Remoto**. Goiânia, Brasil: [s.n.], 2005. p. p. 4225–4232. Disponível em: <http://marte.sid.inpe.br/col/ltid.inpe.br/sbsr/2004/11.22.20.14/doc/4225.pdf>. 5, 13

PONTIUS, R. G.; MILLONES, M. **Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment**. 2011. 4407–4429 p. 16

QUINLAN, J. R. **Simplifying decision trees**. 1987. 221–234 p. 48

_____. **C4.5: Programs for Machine Learning**. [s.n.], 1993. 302 p. ISSN 08856125. ISBN 1558602380. Disponível em: <http://portal.acm.org/citation.cfm?id=152181>. 48, 50, 51

_____. **Página pessoal de Jhon Ross Quinlan**. 2008. Disponível em: <http://www.rulequest.com/Personal/>. 2, 51

QUINLAN, R. **C5.0: An Informal Tutorial**. 1998. Disponível em: <http://www.rulequest.com/see5-unix.html>. 52

R Core Team. **R: A Language and Environment for Statistical Computing**. Vienna, Austria: [s.n.], 2016. Disponível em: <https://www.r-project.org>. 16

RAO, T. C. M.; RAO, K. V.; KUMAR, A. R.; RAO, D. P.; DEEKSHATULA, B. L. Digital terrain model (DTM) from Indian remote sensing (IRS) satellite data from the overlap area of two adjacent paths using digital photogrammetric techniques. **Photogrammetric Engineering and Remote Sensing**, v. 62, p. 727–731, 1996. ISSN 00991112. Disponível em: <{\T1\textless}GotoISI{\T1\textgreater}://A1996UN90800018>. 11

RIBEIRO, B. M. G. **AVALIAÇÃO DAS IMAGENS WORLDVIEW-II PARA O MAPEAMENTO DA COBERTURA DO SOLO URBANO UTILIZANDO O SISTEMA INTERIMAGE**. 179 p. Tese (Doutorado) — INPE, 2010. 6, 89, 99, 104

RIGGS, G. A.; HALL, D. K.; SALOMONSON, V. V. Snow index for the Landsat Thematic Mapper and moderate resolution imaging spectroradiometer. In: **International Geoscience and Remote Sensing Symposium (IGARSS)**. [S.l.: s.n.], 1994. v. 4, p. 1942–1944. 44

ROBERTS, L. G. **Machine perception of three-dimensional solids**. 82 p. Tese (Doutorado) — Massachusetts Institute of Technology, 1963. Disponível em: <http://hdl.handle.net/1721.1/11589>. 13

ROSENFIELD, G. H.; FITZPATRICK-LINS, K. A coefficient of agreement as a measure of thematic classification accuracy. **Photogrammetric Engineering & Remote Sensing**, v. 52, n. 1979, p. 223–227, 1986. ISSN 00991112. 60

Rouse Jr., J. W.; HAAS, R. H.; SCHELL, J. A.; DEERING, D. W. Monitoring Vegetation Systems in the Great Plains with ERTS. In: **Third ERTS Symposium**. [S.l.]: NASA, 1973. 44

ROY, D. P.; BORAK, J. S.; DEVADIGA, S.; WOLFE, R. E.; ZHENG, M.; DESCLOITRES, J. The MODIS Land product quality assessment approach. **Remote Sensing of Environment**, v. 83, p. 62–76, 2002. ISSN 00344257. 9

SATMAP. **Satmap**. 2014. Disponível em: <http://www.satmap.com.br/>. 19

SCHAPIRE, R. E. The Strength of Weak Learnability. **Machine Learning**, v. 5, n. 2, p. 197–227, 1990. ISSN 15730565. 58

_____. The boosting approach to machine learning: an overview. **Nonlinear Estimation and Classification**, v. 171, p. 149–171, 2003. ISSN 09300325. Disponível em: <http://www.ams.org/mathscinet/search/publications.html?pg1=MR{&}s1=MR2005788$\delimiter"026E30F$npapers2://publication/uuid/13F19159-186B-4FAA-9433-AE2BE3D153D8>. 57

SESNIE, S. E.; GESSLER, P. E.; FINEGAN, B.; THESSLER, S. Integrating Landsat TM and SRTM-DEM derived variables with decision trees for habitat classification and change detection in complex neotropical environments. **Remote Sensing of Environment**, v. 112, n. 5, p. 2145–2159, 2008. ISSN 00344257. 2

SIC. **Satellite Imaging Corporation**. 2015. Disponível em: <http://www.satimagingcorp.com/services/orthorectification/>. 12

SLATER, P. N. **Remote sensing, optics and optical systems**. Addison-Wesley Pub. Co., 1980. 575p p. (Remote sensing). ISBN 9780201072501. Disponível em: <http://books.google.com.br/books?id=k4JBAQAAIAAJ>. 5, 8

SUN, W.; CHEN, B.; MESSINGER, D. W. Nearest-neighbor diffusion-based pan-sharpening algorithm for spectral images. **Optical Engineering**, International Society for Optics and Photonics, v. 53, n. 1, p. 013107, jan 2014. ISSN 0091-3286. Disponível em: <http://opticalengineering.spiedigitallibrary.org/article.aspx?articleid=1819137http://opticalengineering.spiedigitallibrary.org/article.aspx?doi=10.1117/1.OE.53.1.013107>. 10, 25

TAO, C.; HU, Y. Use of the Rational Function Model for Image Rectification. **Canadian Journal of Remote Sensing**, Taylor & Francis, v. 27, n. 6, p. 593–602, dec 2001. ISSN 0703-8992. Disponível em: <http://www.tandfonline.com/doi/abs/10.1080/07038992.2001.10854900>. 13

THOMAS, N.; HENDRIX, C.; CONGALTON, R. G. A comparison of urban mapping methods using high-resolution digital imagery. **Photogrammetric Engineering and Remote Sensing**, v. 69, p. 963–972, 2003. ISSN 00991112. Disponível em: <http://www.asprs.org/publications/pers/2003journal/september/2003{_}sep{_}963-972.pdf>. 1

TOUTIN, T. Review article: Geometric processing of remote sensing images: models, algorithms and methods. **International Journal of Remote Sensing**, Taylor & Francis, v. 25, n. 10, p. 1893–1924, 2004. 11

TRIMBLE. **eCognition Developer 8.7 User Guide**. Munich, Germany: [s.n.], 2011. 258 p. Disponível em: <http://www.ecognition.com/>. 16, 43, 46

UNPD. **World Urbanization Prospects: The 2011 Revision**. [s.n.], 2012. 318 p. Disponível em: <http://esa.un.org/unpd/wpp/ppt/CSIS/WUP_2011_CSIS_4.pdf>. 1

UPDIKE, T.; COMP, C. Radiometric use of WorldView-2 imagery. **Digitalglobe**, p. 1–16, 2010. Disponível em: <http://global.digitalglobe.com/sites/default/files/Radiometric_Use_of_WorldView-2_Imagery%20(1).pdf>. 2, 22

Van Niel, T. G.; MCVICAR, T. R.; LI, L.; GALLANT, J. C.; YANG, Q. The impact of misregistration on SRTM and DEM image differences. **Remote Sensing of Environment**, v. 112, n. 5, p. 2430–2442, 2008. ISSN 00344257. 2

VASSILOPOULOU, S.; HURNI, L.; DIETRICH, V.; BALTSAVIAS, E.; PATERAKI, M.; LAGIOS, E.; PARCHARIDIS, I. Orthophoto generation using IKONOS imagery and high-resolution DEM: A case study on volcanic hazard monitoring of Nisyros Island (Greece). In: **ISPRS Journal of Photogrammetry and Remote Sensing**. [S.l.: s.n.], 2002. v. 57, p. 24–38. ISSN 09242716. 13

WEMMERT, C.; PUISSANT, A.; FORESTIER, G.; GANCARSKI, P. Multiresolution Remote Sensing Image Clustering. **IEEE Geoscience and Remote Sensing Letters**, v. 6, 2009. ISSN 1545-598X. 13

WOLF, A. F. **Using WorldView 2 Vis-NIR MSI Imagery to Support Land Mapping and Feature Extraction Using Normalized Difference Index Ratios**. Longmont, Colorado: [s.n.], 2010. 13 p. Disponível em: <http://harrisgeospatial.com/portals/0/pdfs/envi/8{_}bands{_}Antonio{_}Wolf.pdf>. 45

WU, X.; KUMAR, V.; Ross Quinlan, J.; GHOSH, J.; YANG, Q.; MOTODA, H.; MCLACHLAN, G. J.; NG, A.; LIU, B.; YU, P. S.; ZHOU, Z.-H.; STEINBACH, M.; HAND, D. J.; STEINBERG, D. Top 10 algorithms in data mining. **Knowledge and Information Systems**, v. 14, n. 1, p. 1–37, jan 2008. ISSN 0219-1377. Disponível em: <http://link.springer.com/10.1007/s10115-007-0114-2>. 2, 51, 56, 58

XU, H. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. **International Journal of Remote Sensing**, Taylor & Francis, v. 27, n. 14, p. 3025–3033, jul 2006. ISSN 0143-1161. Disponível em: <http://www.tandfonline.com/doi/abs/10.1080/01431160600589179{#}.VvCt5uIrK70http://www.tandfonline.com/doi/abs/10.1080/01431160600589179>. 44

ZHA, Y.; GAO, J.; NI, S. Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. **International Journal of Remote Sensing**, v. 24, n. 3, p. 583–594, jan 2003. ISSN 0143-1161. Disponível em: <http://www.tandfonline.com/doi/abs/10.1080/01431160304987>. 44

ZHANG, Y. PROBLEMS IN THE FUSION OF COMMERCIAL HIGH-RESOLUTION SATELITTE IMAGES AS WELL AS LANDSAT 7 IMAGES AND INITIAL SOLUTIONS. **ISPRS Commission IV, Symposium 2002 Geospatial Theory, Processing and Applications**, XXXIV, p. 6, 2002. Disponível em: <http://www.isprs.org/proceedings/XXXIV/part4/pdfpapers/220.pdf>. 10

APÊNDICE A - GUIA DE NOMES DOS ATRIBUTOS TRUNCADOS PELO ECOGNITION PARA 12 CARACTERES

#	Output	Original
1	asymmetry	Asymmetry
2	border_ind	Border_ind
3	brt	BRT
4	brtmn1	BRTmN1
5	brtmn2	BRTmN2
6	brtpbor	BRTpBoR
7	brtpbore	BRTpBoRE
8	**class_name**	**Class_name**
9	compactnes	Compactnes
10	density	Density
11	elliptic_f	Elliptic_F
12	max_pixel_	Max_pixel_7
13	max_pixel_1	Max_pixel_8
14	**max_pixel_2**	**Max_pixel_MDA**
15	max_pixel_3	Max_pixel_pc1
16	max_pixel_4	Max_pixel_pc2
17	max_pixel_5	Max_pixel_B5
18	max_pixel_6	Max_pixel_pc4
19	max_pixel_7	Max_pixel_pc5
20	max_pixel_8	Max_pixel_pc6
21	max_pixel_9	Max_pixel_pc7
22	max_pixel_10	Max_pixel_pc8
23	max_pixel_11	Max_pixel_B4
24	max_pixel_12	Max_pixel_B1
25	max_pixel_13	Max_pixel_B3
26	max_pixel_14	Max_pixel_B2
27	max_pixel_15	Max_pixel_B6
28	max_pixel_16	Max_pixel_pc3
29	mean_layer	Mean_Layer_B7
30	mean_layer1	Mean_Layer_B8
31	mean_layer2	Mean_Layer_B1
32	mean_layer3	Mean_Layer_B3
33	mean_layer4	Mean_Layer_B4
34	mean_layer5	Mean_Layer_B5
35	mean_layer6	Mean_Layer_B2
36	mean_layer7	Mean_Layer_B6
37	**mean_mda**	**Mean_MDA**
38	mean_pc1	Mean_pc1
39	mean_pc2	Mean_pc2
40	mean_pc3	Mean_pc3
41	mean_pc4	Mean_pc4
42	mean_pc5	Mean_pc5
43	mean_pc6	Mean_pc6
44	mean_pc7	Mean_pc7
45	mean_pc8	Mean_pc8
46	min_pixel_	Min_pixel_B5
47	min_pixel_1	Min_pixel_pc8
48	min_pixel_2	Min_pixel_pc7
49	min_pixel_3	Min_pixel_pc6
50	min_pixel_4	Min_pixel_pc5
51	min_pixel_5	Min_pixel_pc4
52	min_pixel_6	Min_pixel_pc3
53	min_pixel_7	Min_pixel_pc2
54	min_pixel_8	Min_pixel_pc1
55	**min_pixel_9**	**Min_pixel_MDA**
56	min_pixel_10	Min_pixel_B8
57	min_pixel_11	Min_pixel_B7
58	min_pixel_12	Min_pixel_B1
59	min_pixel_13	Min_pixel_B2
60	min_pixel_14	Min_pixel_B3
61	min_pixel_15	Min_pixel_B4
62	min_pixel_16	Min_pixel_B6
63	mnir	mNIR
64	modeminimu	modeMinimupc8
65	modeminimu1	modeMinimupc7
66	modeminimu2	modeMinimupc6
67	modeminimu3	modeMinimupc5
68	modeminimu4	modeMinimupc4
69	modeminimu5	modeMinimupc3
70	modeminimu6	modeMinimupc2
71	modeminimu7	modeMinimupc1
72	**modeminimu8**	**modeMinimuMDA**
73	modeminimu9	modeMinimuB8
74	modeminimu10	modeMinimuB7
75	modeminimu11	modeMinimuB6
76	modeminimu12	modeMinimuB5
77	modeminimu13	modeMinimuB4
78	modeminimu14	modeMinimuB2
79	modeminimu15	modeMinimuB3
80	modeminimu16	modeMinimuB1
81	mswir	mSWIR
82	ndbi	NDBI
83	ndsi1	NDSI1
84	ndsi2	NDSI2
85	ndvi1	NDVI1
86	ndvi2	NDVI2
87	ndvi3	NDVI3
88	rectangula	Rectangula
89	roundness	Roundness
90	shape_inde	Shape_inde
91	wvbi	WVBI
92	wvnhfd	WVNHFD
93	wvnii	WVNII
94	wvsi	WVSI

APÊNDICE B - ÁRVORES DE DECISÃO C4.5, C5.0 E CART COM E SEM MDA PARA AS LEGENDAS DO NÍVEL I E NÍVEL II

B.1 Árvore gerada por C4.5 Nível II com MDA

```
 J48 pruned tree
------------------

Max_pixel_.4 <= -1591.778076
|   NDVI1 <= 0.511419
|   |   Max_pixel_.1 <= -121: sombra (16.0)
|   |   Max_pixel_.1 > -121
|   |   |   Border_ind <= 1.823529: fibrocimento_amianto (3.0/1.0)
|   |   |   Border_ind > 1.823529: piscina_azulejo (5.0)
|   NDVI1 > 0.511419: lago (442.0)
Max_pixel_.4 > -1591.778076
|   Mean_Layer.5 <= 1164.12
|   |   modeMinimu.15 <= -6
|   |   |   Mean_pc3 <= 328.335594
|   |   |   |   Mean_pc5 <= -106.656215
|   |   |   |   |   BRTpBoR <= 0.908808: ceramica (7.0)
|   |   |   |   |   BRTpBoR > 0.908808: concreto (3.0)
|   |   |   |   Mean_pc5 > -106.656215
|   |   |   |   |   Min_pixel_.14 <= -149
|   |   |   |   |   |   modeMinimu.9 <= -60
|   |   |   |   |   |   |   Max_pixel_.3 <= -59347.671875: asfalto (2.0)
|   |   |   |   |   |   |   Max_pixel_.3 > -59347.671875
|   |   |   |   |   |   |   |   BRTmN1 <= 59.86284
|   |   |   |   |   |   |   |   |   Min_pixel_.5 <= -366.194794: sombra (24.0/1.0)
|   |   |   |   |   |   |   |   |   Min_pixel_.5 > -366.194794
|   |   |   |   |   |   |   |   |   |   modeMinimu.8 <= 22.897768: sombra (8.0/1.0)
|   |   |   |   |   |   |   |   |   |   modeMinimu.8 > 22.897768
|   |   |   |   |   |   |   |   |   |   |   Min_pixel_.3 <= -155.748718: sombra (2.0)
|   |   |   |   |   |   |   |   |   |   |   Min_pixel_.3 > -155.748718: vegetacao_arborea (6.0)
|   |   |   |   |   |   |   |   BRTmN1 > 59.86284
|   |   |   |   |   |   |   |   |   Mean_pc4 <= -70.999516
|   |   |   |   |   |   |   |   |   |   Mean_pc4 <= -73.396498: sombra (10.0)
|   |   |   |   |   |   |   |   |   |   Mean_pc4 > -73.396498: concreto (2.0)
|   |   |   |   |   |   |   |   |   Mean_pc4 > -70.999516: sombra (159.0)
|   |   |   |   |   |   modeMinimu.9 > -60
|   |   |   |   |   |   |   Mean_Layer.2 <= -185.176471
|   |   |   |   |   |   |   |   Max_pixel_.2 <= 26.805666: sombra (2.0)
|   |   |   |   |   |   |   |   Max_pixel_.2 > 26.805666: vegetacao_arborea (3.0)
|   |   |   |   |   |   |   Mean_Layer.2 > -185.176471
|   |   |   |   |   |   |   |   modeMinimu.16 <= -191: policarbonato (2.0)
|   |   |   |   |   |   |   |   modeMinimu.16 > -191: fibrocimento_amianto (3.0)
|   |   |   |   |   Min_pixel_.14 > -149
|   |   |   |   |   |   Max_pixel_.10 <= 114.59613
|   |   |   |   |   |   |   Asymmetry <= 0.977404: aco_brilhante (3.0/1.0)
|   |   |   |   |   |   |   Asymmetry > 0.977404: fibrocimento_amianto (2.0)
|   |   |   |   |   |   Max_pixel_.10 > 114.59613
|   |   |   |   |   |   |   Compactnes <= 2.16: concreto (3.0)
|   |   |   |   |   |   |   Compactnes > 2.16: sombra (2.0/1.0)
|   |   |   Mean_pc3 > 328.335594
|   |   |   |   Min_pixel_ <= -593
```

```
|   |   |   |   |   Max_pixel_.1 <= 1587
|   |   |   |   |   |   Border_ind <= 1.84375: sombra (2.0)
|   |   |   |   |   |   Border_ind > 1.84375: vegetacao_arborea (3.0)
|   |   |   |   |   Max_pixel_.1 > 1587
|   |   |   |   |   |   Min_pixel_.9 <= 32.885559: vegetacao_arborea (128.0)
|   |   |   |   |   |   Min_pixel_.9 > 32.885559
|   |   |   |   |   |   |   BRT <= 222.488073: vegetacao_arborea (4.0)
|   |   |   |   |   |   |   BRT > 222.488073: vegetacao_rasteira (2.0)
|   |   |   |   Min_pixel_ > -593: vegetacao_rasteira (3.0/1.0)
|   |   modeMinimu.15 > -6
|   |   |   NDVI1 <= 0.15524
|   |   |   |   NDVI2 <= 0.078043
|   |   |   |   |   Min_pixel_.9 <= 25.302431
|   |   |   |   |   |   WVSI <= -0.203472
|   |   |   |   |   |   |   Min_pixel_.8 <= -63285.695313
|   |   |   |   |   |   |   |   Border_ind <= 1.214286: asfalto (3.0)
|   |   |   |   |   |   |   |   Border_ind > 1.214286
|   |   |   |   |   |   |   |   |   Asymmetry <= 0.694804: argamassa_cimento_pintado_vermelho (2.0/1.0)
|   |   |   |   |   |   |   |   |   Asymmetry > 0.694804: vegetacao_rasteira (2.0/1.0)
|   |   |   |   |   |   |   Min_pixel_.8 > -63285.695313: carvao_coque (9.0)
|   |   |   |   |   |   WVSI > -0.203472
|   |   |   |   |   |   |   Border_ind <= 4.709091
|   |   |   |   |   |   |   |   Mean_pc6 <= -44.294682: fibrocimento_amianto (3.0/1.0)
|   |   |   |   |   |   |   |   Mean_pc6 > -44.294682: asfalto (109.0/1.0)
|   |   |   |   |   |   |   Border_ind > 4.709091: asfalto_pintado_branco (5.0/1.0)
|   |   |   |   |   Min_pixel_.9 > 25.302431
|   |   |   |   |   |   WVSI <= -0.055943
|   |   |   |   |   |   |   BRT <= 360.367347: vidro_aramado (6.0)
|   |   |   |   |   |   |   BRT > 360.367347: asfalto (2.0)
|   |   |   |   |   |   WVSI > -0.055943
|   |   |   |   |   |   |   Max_pixel_.2 <= 32.737488: policarbonato (4.0/1.0)
|   |   |   |   |   |   |   Max_pixel_.2 > 32.737488: aco_fosco (4.0)
|   |   |   |   NDVI2 > 0.078043
|   |   |   |   |   Min_pixel_.3 <= -123.692596
|   |   |   |   |   |   modeMinimu.2 <= -939.210632: concreto (3.0)
|   |   |   |   |   |   modeMinimu.2 > -939.210632
|   |   |   |   |   |   |   NDBI <= 0.103527: concreto (3.0)
|   |   |   |   |   |   |   NDBI > 0.103527
|   |   |   |   |   |   |   |   Max_pixel_.16 <= 503.276459: fibrocimento_amianto (58.0)
|   |   |   |   |   |   |   |   Max_pixel_.16 > 503.276459
|   |   |   |   |   |   |   |   |   Max_pixel_.16 <= 546.665344: concreto (4.0/1.0)
|   |   |   |   |   |   |   |   |   Max_pixel_.16 > 546.665344: fibrocimento_amianto (11.0)
|   |   |   |   |   Min_pixel_.3 > -123.692596
|   |   |   |   |   |   Density <= 1.299588: asfalto (3.0)
|   |   |   |   |   |   Density > 1.299588: asfalto_pintado_branco (2.0/1.0)
|   |   |   NDVI1 > 0.15524
|   |   |   |   NDSI1 <= -0.393552
|   |   |   |   |   Min_pixel_.16 <= -388
|   |   |   |   |   |   BRT <= 86.180346: sombra (3.0/1.0)
|   |   |   |   |   |   BRT > 86.180346: fibrocimento_amianto (3.0/1.0)
|   |   |   |   |   Min_pixel_.16 > -388
|   |   |   |   |   |   Mean_Layer.7 <= 1600.552632
|   |   |   |   |   |   |   NDVI3 <= 0.352876: concreto (10.0/1.0)
|   |   |   |   |   |   |   NDVI3 > 0.352876
|   |   |   |   |   |   |   |   modeMinimu.12 <= 668
|   |   |   |   |   |   |   |   |   modeMinimu.12 <= -411: vegetacao_arborea (2.0/1.0)
|   |   |   |   |   |   |   |   |   modeMinimu.12 > -411
```

```
|   |   |   |   |   |   |   |   |   |   Mean_pc1 <= -62255.825103: solo_exposto (2.0/1.0)
|   |   |   |   |   |   |   |   |   |   Mean_pc1 > -62255.825103: vegetacao_rasteira (48.0/3.0)
|   |   |   |   |   |   |   |   modeMinimu.12 > 668: solo_exposto (3.0)
|   |   |   |   |   |   Mean_Layer.7 > 1600.552632: solo_exposto (7.0)
|   |   |   |   NDSI1 > -0.393552
|   |   |   |   |   Mean_pc8 <= -68.662459: argamassa_cimento_pintado_verde (7.0)
|   |   |   |   |   Mean_pc8 > -68.662459
|   |   |   |   |   |   Compactnes <= 1.865341: policarbonato (3.0/1.0)
|   |   |   |   |   |   Compactnes > 1.865341: plastico (4.0)
|   Mean_Layer.5 > 1164.12
|   |   NDSI2 <= -0.420767
|   |   |   modeMinimu.1 <= -279.613403
|   |   |   |   Min_pixel_.4 <= -386.32843
|   |   |   |   |   WVBI <= -0.845891: ceramica (17.0)
|   |   |   |   |   WVBI > -0.845891
|   |   |   |   |   |   Mean_Layer.3 <= 1576.847826
|   |   |   |   |   |   |   Min_pixel_ <= 606: ceramica (2.0)
|   |   |   |   |   |   |   Min_pixel_ > 606: asfalto_pintado_vermelho (38.0/1.0)
|   |   |   |   |   |   Mean_Layer.3 > 1576.847826
|   |   |   |   |   |   |   Compactnes <= 1.316239: concreto_pintado_vermelho (2.0)
|   |   |   |   |   |   |   Compactnes > 1.316239: ceramica (5.0)
|   |   |   |   Min_pixel_.4 > -386.32843: argamassa_cimento_pintado_vermelho (10.0)
|   |   |   modeMinimu.1 > -279.613403
|   |   |   |   NDSI1 <= -0.454906
|   |   |   |   |   Min_pixel_.3 <= 51.887234: solo_exposto (25.0/1.0)
|   |   |   |   |   Min_pixel_.3 > 51.887234: ceramica (3.0)
|   |   |   |   NDSI1 > -0.454906
|   |   |   |   |   Mean_mda <= 21.712175: bloquete (4.0/1.0)
|   |   |   |   |   Mean_mda > 21.712175
|   |   |   |   |   |   Max_pixel_.8 <= 606.387146: concreto (8.0)
|   |   |   |   |   |   Max_pixel_.8 > 606.387146: asfalto_pintado_vermelho (3.0/1.0)
|   |   NDSI2 > -0.420767
|   |   |   Mean_mda <= 33.200557
|   |   |   |   Max_pixel_.2 <= 27.4527
|   |   |   |   |   Mean_pc3 <= -990.864507
|   |   |   |   |   |   BRTpBoR <= 1.490375
|   |   |   |   |   |   |   Mean_pc6 <= 218.766316
|   |   |   |   |   |   |   |   Max_pixel_.9 <= 217.851303: argamassa_cimento (3.0/1.0)
|   |   |   |   |   |   |   |   Max_pixel_.9 > 217.851303: concreto_pintado_amarelo (9.0)
|   |   |   |   |   |   |   Mean_pc6 > 218.766316: quartzito_amarelo (14.0)
|   |   |   |   |   |   BRTpBoR > 1.490375
|   |   |   |   |   |   |   Asymmetry <= 0.804238
|   |   |   |   |   |   |   |   Border_ind <= 1.153846: piscina_vinil (5.0)
|   |   |   |   |   |   |   |   Border_ind > 1.153846: concreto_impermeabilizado (5.0)
|   |   |   |   |   |   |   Asymmetry > 0.804238: galvalume_a (8.0)
|   |   |   |   |   Mean_pc3 > -990.864507
|   |   |   |   |   |   NDSI2 <= -0.128786
|   |   |   |   |   |   |   Mean_pc1 <= -68145.187812
|   |   |   |   |   |   |   |   Border_ind <= 1.1: fibra_vidro (4.0)
|   |   |   |   |   |   |   |   Border_ind > 1.1: concreto_impermeabilizado (18.0)
|   |   |   |   |   |   |   Mean_pc1 > -68145.187812
|   |   |   |   |   |   |   |   Min_pixel_.6 <= -339.564789
|   |   |   |   |   |   |   |   |   Asymmetry <= 0.949928
|   |   |   |   |   |   |   |   |   |   Mean_pc7 <= 31.914955: asfalto_pintado_branco (33.0/2.0)
|   |   |   |   |   |   |   |   |   |   Mean_pc7 > 31.914955
|   |   |   |   |   |   |   |   |   |   |   Min_pixel_.8 <= -69639.03125: concreto_impermeabilizado (7.0/1.0)
|   |   |   |   |   |   |   |   |   |   |   Min_pixel_.8 > -69639.03125
```

```
|   |   |   |   |   |   |   |   |   |   |   |   Mean_pc6 <= 41.921515: concreto (17.0)
|   |   |   |   |   |   |   |   |   |   |   |   Mean_pc6 > 41.921515: asfalto_pintado_branco (5.0/1.0)
|   |   |   |   |   |   |   |   |   Asymmetry > 0.949928: argamassa_cimento (5.0/1.0)
|   |   |   |   |   |   |   |   Min_pixel_.6 > -339.564789: plastico (3.0/1.0)
|   |   |   |   |   |   NDSI2 > -0.128786
|   |   |   |   |   |   |   Max_pixel_.12 <= 2879: aco_fosco (12.0)
|   |   |   |   |   |   |   Max_pixel_.12 > 2879: piscina_azulejo (3.0/1.0)
|   |   |   |   Max_pixel_.2 > 27.4527
|   |   |   |   |   Max_pixel_.12 <= 4205
|   |   |   |   |   |   WVBI <= -0.251603
|   |   |   |   |   |   |   Min_pixel_.1 <= -236.700043
|   |   |   |   |   |   |   |   Max_pixel_.3 <= -61684.292969: plastico (3.0/2.0)
|   |   |   |   |   |   |   |   Max_pixel_.3 > -61684.292969: concreto (5.0)
|   |   |   |   |   |   |   Min_pixel_.1 > -236.700043
|   |   |   |   |   |   |   |   Max_pixel_.14 <= 1542
|   |   |   |   |   |   |   |   |   Border_ind <= 2.115385: fibrocimento_amianto (2.0/1.0)
|   |   |   |   |   |   |   |   |   Border_ind > 2.115385: concreto (2.0)
|   |   |   |   |   |   |   |   Max_pixel_.14 > 1542
|   |   |   |   |   |   |   |   |   Roundness <= 0.122017
|   |   |   |   |   |   |   |   |   |   Asymmetry <= 0.528523: lona (2.0)
|   |   |   |   |   |   |   |   |   |   Asymmetry > 0.528523: fibrocimento_crfs (9.0)
|   |   |   |   |   |   |   |   |   Roundness > 0.122017
|   |   |   |   |   |   |   |   |   |   WVSI <= -0.079601
|   |   |   |   |   |   |   |   |   |   |   Border_ind <= 1.4: aco_brilhante (4.0)
|   |   |   |   |   |   |   |   |   |   |   Border_ind > 1.4: aco_ferrugem (7.0)
|   |   |   |   |   |   |   |   |   |   WVSI > -0.079601
|   |   |   |   |   |   |   |   |   |   |   Mean_pc5 <= 200.965151
|   |   |   |   |   |   |   |   |   |   |   |   WVSI <= -0.009132
|   |   |   |   |   |   |   |   |   |   |   |   |   Mean_pc8 <= 20.608874
|   |   |   |   |   |   |   |   |   |   |   |   |   |   Border_ind <= 1.25: fibrocimento_crfs (3.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   |   Border_ind > 1.25
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   Max_pixel_.10 <= 234.989471: aco_brilhante (79.0/2.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   Max_pixel_.10 > 234.989471: fibrocimento_crfs (2.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   Mean_pc8 > 20.608874
|   |   |   |   |   |   |   |   |   |   |   |   |   |   Elliptic_F <= 0.70297: fibrocimento_crfs (11.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   |   Elliptic_F > 0.70297: aco_brilhante (5.0/1.0)
|   |   |   |   |   |   |   |   |   |   |   |   WVSI > -0.009132: aco_fosco (2.0/1.0)
|   |   |   |   |   |   |   |   |   |   |   Mean_pc5 > 200.965151: galvalume_a (4.0)
|   |   |   |   |   |   WVBI > -0.251603: aco_fosco (16.0/1.0)
|   |   |   |   |   Max_pixel_.12 > 4205
|   |   |   |   |   |   Shape_inde <= 1.360147: lona (4.0)
|   |   |   |   |   |   Shape_inde > 1.360147
|   |   |   |   |   |   |   NDBI <= 0.042605: galvalume_a (43.0/1.0)
|   |   |   |   |   |   |   NDBI > 0.042605: aco_brilhante (3.0)
|   |   |   Mean_mda > 33.200557
|   |   |   |   Min_pixel_.6 <= -1581.449463
|   |   |   |   |   Max_pixel_.4 <= 3498.86499
|   |   |   |   |   |   Mean_pc4 <= 54.557117: galvalume_c (3.0)
|   |   |   |   |   |   Mean_pc4 > 54.557117: galvalume_b (36.0)
|   |   |   |   |   Max_pixel_.4 > 3498.86499
|   |   |   |   |   |   Min_pixel_.6 <= -1916.166626: galvalume_c (68.0)
|   |   |   |   |   |   Min_pixel_.6 > -1916.166626
|   |   |   |   |   |   |   Asymmetry <= 0.670679: acrilico (2.0)
|   |   |   |   |   |   |   Asymmetry > 0.670679: galvalume_a (2.0)
|   |   |   |   Min_pixel_.6 > -1581.449463
|   |   |   |   |   Border_ind <= 1.243902
|   |   |   |   |   |   Asymmetry <= 0.567878: acrilico (10.0)
```

```
|   |   |   |   |   |   Asymmetry > 0.567878: sombra (2.0/1.0)
|   |   |   |   |   Border_ind > 1.243902
|   |   |   |   |   |   NDSI1 <= -0.117383
|   |   |   |   |   |   |   BRTpBoR <= 1.580593
|   |   |   |   |   |   |   |   Asymmetry <= 0.926325: concreto (3.0)
|   |   |   |   |   |   |   |   Asymmetry > 0.926325: pastilha_esmaltada (2.0)
|   |   |   |   |   |   |   BRTpBoR > 1.580593: fibra_vidro (5.0/1.0)
|   |   |   |   |   |   NDSI1 > -0.117383
|   |   |   |   |   |   |   Max_pixel_.11 <= 4670: aco_fosco (11.0)
|   |   |   |   |   |   |   Max_pixel_.11 > 4670: concreto_impermeabilizado (2.0)

Number of Leaves  :  115

Size of the tree :  229
```

B.2 Árvore gerada por C5.0 Nível II com MDA

```
 C5.0 [Release 2.07 GPL Edition]
-------------------------------

Class specified by attribute 'outcome'

Read 1796 cases (94 attributes) from undefined.data

Decision tree:

Max_pixel_.4 <= -1591.778:
:...NDSI2 <= -0.725163: lago (442)
:   NDSI2 > -0.725163:
:   :...Max_pixel_.16 > 1115.027: piscina_azulejo (5)
:       Max_pixel_.16 <= 1115.027:
:       :...NDVI3 <= 0.221289: fibrocimento_amianto (3/1)
:           NDVI3 > 0.221289: sombra (16)
Max_pixel_.4 > -1591.778:
:...Mean_Layer.5 <= 1164.12:
    :...modeMinimu.15 > -6:
    :   :...NDVI1 <= 0.15524:
    :   :   :...NDVI2 <= 0.078043:
    :   :   :   :...Min_pixel_.9 <= 25.30243:
    :   :   :   :   :...WVSI <= -0.203472:
    :   :   :   :   :   :...Min_pixel_.12 > -39: asfalto (4/1)
    :   :   :   :   :   :   Min_pixel_.12 <= -39:
    :   :   :   :   :   :   :...Min_pixel_.3 <= -142.9033: concreto (2/1)
    :   :   :   :   :   :       Min_pixel_.3 > -142.9033: carvao_coque (10/1)
    :   :   :   :   :   WVSI > -0.203472:
    :   :   :   :   :   :...Max_pixel_.11 > 2165: asfalto_pintado_branco (3/1)
    :   :   :   :   :       Max_pixel_.11 <= 2165:
    :   :   :   :   :       :...Border_ind <= 5.033898: asfalto (111/2)
    :   :   :   :   :           Border_ind > 5.033898: asfalto_pintado_branco (3/1)
    :   :   :   :   Min_pixel_.9 > 25.30243:
    :   :   :   :   :...Min_pixel_.9 > 41.34003: vidro_aramado (6)
    :   :   :   :       Min_pixel_.9 <= 41.34003:
    :   :   :   :       :...Min_pixel_.8 <= -68880.95: asfalto (2)
```

```
:   :   :   :           Min_pixel_.8 > -68880.95:
:   :   :   :           :...Max_pixel_.2 <= 32.73749: policarbonato (4/1)
:   :   :   :               Max_pixel_.2 > 32.73749: aco_fosco (4)
:   :   :   NDVI2 > 0.078043:
:   :   :   :...Min_pixel_.3 > -123.6926:
:   :   :       :...Mean_mda <= 23.62375: argamassa_cimento_pintado_verde (2/1)
:   :   :       :   Mean_mda > 23.62375: asfalto (3)
:   :   :       Min_pixel_.3 <= -123.6926:
:   :   :       :...Min_pixel_.3 <= -939.2106: concreto (3)
:   :   :           Min_pixel_.3 > -939.2106:
:   :   :           :...NDBI <= 0.103527: concreto (3)
:   :   :               NDBI > 0.103527:
:   :   :               :...Max_pixel_.16 <= 503.2765: fibrocimento_amianto (58)
:   :   :                   Max_pixel_.16 > 503.2765: [S1]
:   :   NDVI1 > 0.15524:
:   :   :...NDSI1 > -0.393552:
:   :       :...Mean_pc8 <= -68.66246: argamassa_cimento_pintado_verde (7)
:   :       :   Mean_pc8 > -68.66246:
:   :       :   :...Mean_mda <= 26.44551: plastico (5/1)
:   :       :       Mean_mda > 26.44551: policarbonato (2)
:   :       NDSI1 <= -0.393552:
:   :       :...Min_pixel_.16 <= -388:
:   :           :...Mean_pc8 > 31.2787: fibrocimento_amianto (2)
:   :           :   Mean_pc8 <= 31.2787:
:   :           :   :...Border_ind <= 2.619048: sombra (2)
:   :           :       Border_ind > 2.619048: concreto (2/1)
:   :           Min_pixel_.16 > -388:
:   :           :...Mean_Layer.7 > 1600.553: solo_exposto (7)
:   :               Mean_Layer.7 <= 1600.553:
:   :               :...NDVI3 <= 0.352876: concreto (10/1)
:   :                   NDVI3 > 0.352876:
:   :                   :...modeMinimu.12 > 668: solo_exposto (3)
:   :                       modeMinimu.12 <= 668:
:   :                       :...modeMinimu.12 <= -411: concreto (2/1)
:   :                           modeMinimu.12 > -411:
:   :                           :...NDSI1 <= -0.499986: vegetacao_rasteira (48/3)
:   :                               NDSI1 > -0.499986: concreto (2/1)
:   modeMinimu.15 <= -6:
:   :...Mean_pc3 > 328.3356:
:       :...Min_pixel_ > -593: vegetacao_rasteira (3/1)
:       :   Min_pixel_ <= -593:
:       :   :...Min_pixel_.9 > 32.88556:
:       :       :...WVSI <= -0.221119: vegetacao_rasteira (2)
:       :       :   WVSI > -0.221119: vegetacao_arborea (4)
:       :       Min_pixel_.9 <= 32.88556:
:       :       :...Max_pixel_.1 > 1587: vegetacao_arborea (128)
:       :           Max_pixel_.1 <= 1587:
:       :           :...modeMinimu.15 <= -658: sombra (2)
:       :               modeMinimu.15 > -658: vegetacao_arborea (3)
:       Mean_pc3 <= 328.3356:
:       :...Mean_pc5 <= -106.6562:
:           :...Min_pixel_.9 <= 37.48987: ceramica (7)
:           :   Min_pixel_.9 > 37.48987: concreto (3)
:           Mean_pc5 > -106.6562:
:           :...Min_pixel_.13 > -208:
:               :...Mean_pc6 <= -20.72629:
:               :   :...WVSI <= 0.002854: concreto (4)
```

```
:               :   :   WVSI > 0.002854: fibrocimento_amianto (4/1)
:               :   Mean_pc6 > -20.72629:
:               :   :...Mean_mda > 27.79305: aco_brilhante (3/1)
:               :       Mean_mda <= 27.79305:
:               :       :...Asymmetry <= 0.487199: vegetacao_arborea (2)
:               :           Asymmetry > 0.487199: sombra (2)
:               Min_pixel_.13 <= -208:
:               :...modeMinimu.9 > -102:
:                   :...modeMinimu.14 <= -413: policarbonato (3/1)
:                   :   modeMinimu.14 > -413:
:                   :   :...modeMinimu.8 > 28.97406: vegetacao_arborea (2)
:                   :       modeMinimu.8 <= 28.97406:
:                   :       :...Mean_pc5 <= -27.10127: fibrocimento_amianto (2)
:                   :           Mean_pc5 > -27.10127: sombra (3)
:                   modeMinimu.9 <= -102:
:                   :...WVBI > 4.227654:
:                       :...BRT <= 6.084483: asfalto (2/1)
:                       :   BRT > 6.084483: sombra (2)
:                       WVBI <= 4.227654:
:                       :...WVBI <= -7.948527: vegetacao_arborea (3/1)
:                           WVBI > -7.948527:
:                           :...Min_pixel_.7 <= -1833.292: sombra (173/1)
:                               Min_pixel_.7 > -1833.292:
:                               :...Min_pixel_.9 <= 22.87647: sombra (21)
:                                   Min_pixel_.9 > 22.87647: [S2]
Mean_Layer.5 > 1164.12:
:...NDSI2 <= -0.420767:
    :...modeMinimu.1 > -279.6134:
    :   :...Density <= 0.773677: asfalto_pintado_vermelho (2)
    :   :   Density > 0.773677:
    :   :   :...NDSI1 <= -0.454906:
    :   :       :...Min_pixel_.3 <= 51.88723: solo_exposto (25/1)
    :   :       :   Min_pixel_.3 > 51.88723: ceramica (3)
    :   :       NDSI1 > -0.454906:
    :   :       :...Mean_mda <= 21.71218: bloquete (4/1)
    :   :           Mean_mda > 21.71218: concreto (9/1)
    :   modeMinimu.1 <= -279.6134:
    :   :...Min_pixel_.4 > -386.3284: argamassa_cimento_pintado_vermelho (10)
    :       Min_pixel_.4 <= -386.3284:
    :       :...WVBI <= -0.845891: ceramica (17)
    :           WVBI > -0.845891:
    :           :...WVNII <= 4.555916:
    :               :...Min_pixel_.8 <= -72575.66: ceramica (2)
    :               :   Min_pixel_.8 > -72575.66: asfalto_pintado_vermelho (38/1)
    :               WVNII > 4.555916:
    :               :...BRTpBoRE <= 0.928107: ceramica (5)
    :                   BRTpBoRE > 0.928107: concreto_pintado_vermelho (2)
    NDSI2 > -0.420767:
    :...Mean_mda > 33.20056:
        :...Min_pixel_.6 <= -1581.449:
        :   :...Max_pixel_.4 <= 3498.865:
        :   :   :...Mean_pc4 <= 54.55712: galvalume_c (3)
        :   :   :   Mean_pc4 > 54.55712: galvalume_b (36)
        :   :   Max_pixel_.4 > 3498.865:
        :   :   :...modeMinimu.8 > 50.05768: acrilico (2)
        :   :       modeMinimu.8 <= 50.05768:
        :   :       :...Mean_mda <= 36.29212: galvalume_a (2)
```

```
:   :           Mean_mda > 36.29212: galvalume_c (68)
:   Min_pixel_.6 > -1581.449:
:   :...Max_pixel_.2 <= 42.25928:
:       :...NDVI2 <= 0.129309:
:       :   :...Max_pixel_.4 <= 2903.06: aco_fosco (11)
:       :   :   Max_pixel_.4 > 2903.06: concreto_impermeabilizado (2/1)
:       :   NDVI2 > 0.129309:
:       :   :...Mean_mda <= 36.60251: pastilha_esmaltada (2)
:       :       Mean_mda > 36.60251: fibra_vidro (4)
:       Max_pixel_.2 > 42.25928:
:       :...Shape_inde > 1.541104: concreto (4/1)
:           Shape_inde <= 1.541104:
:           :...Min_pixel_.9 <= 45.72943: concreto_impermeabilizado (2/1)
:               Min_pixel_.9 > 45.72943: acrilico (10)
Mean_mda <= 33.20056:
:...Max_pixel_.2 <= 27.4527:
    :...Mean_pc3 <= -990.8645:
    :   :...Max_pixel_.11 > 6663: galvalume_a (8)
    :   :   Max_pixel_.11 <= 6663:
    :   :   :...Max_pixel_.12 > 1881:
    :   :       :...Mean_mda <= 23.6016: piscina_vinil (5)
    :   :       :   Mean_mda > 23.6016: concreto_impermeabilizado (5)
    :   :       Max_pixel_.12 <= 1881: [S3]
    :   Mean_pc3 > -990.8645:
    :   :...NDSI2 > -0.128786:
    :       :...Max_pixel_.7 <= 744.516: aco_fosco (13/1)
    :       :   Max_pixel_.7 > 744.516: piscina_azulejo (2)
    :       NDSI2 <= -0.128786:
    :       :...Mean_pc1 <= -68145.19: [S4]
    :           Mean_pc1 > -68145.19:
    :           :...Min_pixel_.6 > -339.5648: plastico (3/1)
    :               Min_pixel_.6 <= -339.5648:
    :               :...Asymmetry > 0.949928: argamassa_cimento (5/1)
    :                   Asymmetry <= 0.949928: [S5]
    Max_pixel_.2 > 27.4527:
    :...Max_pixel_.12 > 4205:
        :...Shape_inde <= 1.360147: lona (4)
        :   Shape_inde > 1.360147:
        :   :...NDBI <= 0.042605: galvalume_a (43/1)
        :       NDBI > 0.042605: aco_brilhante (3)
        Max_pixel_.12 <= 4205:
        :...WVBI > -0.251603: aco_fosco (16/1)
            WVBI <= -0.251603:
            :...Min_pixel_.1 <= -236.7:
                :...Mean_pc3 <= -763.2145: lona (2/1)
                :   Mean_pc3 > -763.2145: concreto (6/1)
                Min_pixel_.1 > -236.7:
                :...Max_pixel_.14 <= 1542:
                    :...Border_ind <= 2.115385: fibrocimento_amianto (2/1)
                    :   Border_ind > 2.115385: concreto (2)
                    Max_pixel_.14 > 1542:
                    :...Roundness <= 0.122017: [S6]
                        Roundness > 0.122017:
                        :...BRTpBoR <= 1.586538: [S7]
                            BRTpBoR > 1.586538:
                            :...BRTpBoR > 1.857542: [S8]
                                BRTpBoR <= 1.857542:
```

```
                                        :...WVSI > -0.009132: aco_fosco (2/1)
                                            WVSI <= -0.009132:
                                            :...Mean_pc8 > 20.60887: [S9]
                                                Mean_pc8 <= 20.60887: [S10]

SubTree [S1]

Mean_Layer.4 <= 503.3668: fibrocimento_amianto (11)
Mean_Layer.4 > 503.3668: concreto (4/1)

SubTree [S2]

Mean_Layer.6 <= -213.5914: vegetacao_arborea (4)
Mean_Layer.6 > -213.5914: sombra (3)

SubTree [S3]

Max_pixel_.10 <= 20.61736: concreto_pintado_amarelo (9)
Max_pixel_.10 > 20.61736:
:...Min_pixel_.4 <= -520.6105: quartzito_amarelo (14)
    Min_pixel_.4 > -520.6105: argamassa_cimento (3/1)

SubTree [S4]

Rectangula <= 0.92126: concreto_impermeabilizado (18)
Rectangula > 0.92126: fibra_vidro (4)

SubTree [S5]

Mean_pc7 <= 31.91496: asfalto_pintado_branco (33/2)
Mean_pc7 > 31.91496:
:...Min_pixel_.8 <= -69639.03: concreto_impermeabilizado (7/1)
    Min_pixel_.8 > -69639.03:
    :...Max_pixel_.10 > 2716.495: asfalto_pintado_branco (2)
        Max_pixel_.10 <= 2716.495:
        :...Max_pixel_.2 <= 23.77509: asfalto_pintado_branco (2)
            Max_pixel_.2 > 23.77509: concreto (18)

SubTree [S6]

Max_pixel_.16 <= -1118.707: lona (2)
Max_pixel_.16 > -1118.707: fibrocimento_crfs (9)

SubTree [S7]

Max_pixel_.2 <= 29.76666: aco_brilhante (4)
Max_pixel_.2 > 29.76666: aco_ferrugem (7)

SubTree [S8]

Mean_pc3 <= -331.1366: galvalume_a (4)
Mean_pc3 > -331.1366: aco_brilhante (2)

SubTree [S9]

Mean_pc4 <= -147.452: aco_brilhante (3)
Mean_pc4 > -147.452: fibrocimento_crfs (13/1)
```

```
SubTree [S10]

Border_ind <= 1.25: fibrocimento_crfs (3)
Border_ind > 1.25:
:...Mean_pc5 > -51.92517: aco_brilhante (70)
    Mean_pc5 <= -51.92517:
    :...Max_pixel_.10 <= 80.68051: aco_brilhante (5)
        Max_pixel_.10 > 80.68051: fibrocimento_crfs (4)
```

B.3 Árvore gerada por CART Nível II com MDA

```
CART Decision Tree - (Nível II com MDA)

Max_pixel_.2 < 19.368837499999998
|  Min_pixel_.2 < 2.599057
|  |  Max_pixel_.12 < 916.0
|  |  |  BRT < -27.004783500000002: sombra(6.0/0.0)
|  |  |  BRT >= -27.004783500000002: solo_exposto(3.0/5.0)
|  |  Max_pixel_.12 >= 916.0: asfalto_pintado_branco(7.0/0.0)
|  Min_pixel_.2 >= 2.599057: lago(442.0/0.0)
Max_pixel_.2 >= 19.368837499999998
|  mSWIR < -54.346544
|  |  Min_pixel_.12 < -228.0: sombra(195.0/9.0)
|  |  Min_pixel_.12 >= -228.0: asfalto(3.0/2.0)
|  mSWIR >= -54.346544
|  |  Mean_Layer.5 < -216.93235299999998
|  |  |  modeMinimu.9 < -422.5: sombra(6.0/0.0)
|  |  |  modeMinimu.9 >= -422.5: vegetacao_arborea(131.0/4.0)
|  |  Mean_Layer.5 >= -216.93235299999998
|  |  |  BRTmN2 < -147.8775245
|  |  |  |  Max_pixel_.12 < 4699.0
|  |  |  |  |  BRTmN1 < -219.821
|  |  |  |  |  |  WVBI < -0.4621745
|  |  |  |  |  |  |  Mean_Layer.4 < 480.13196600000003
|  |  |  |  |  |  |  |  modeMinimu.13 < -45.0
|  |  |  |  |  |  |  |  |  Min_pixel_.7 < -1217.434021: sombra(4.0/0.0)
|  |  |  |  |  |  |  |  |  Min_pixel_.7 >= -1217.434021: vegetacao_arborea(7.0/1.0)
|  |  |  |  |  |  |  |  modeMinimu.13 >= -45.0: vegetacao_rasteira(40.0/4.0)
|  |  |  |  |  |  |  Mean_Layer.4 >= 480.13196600000003
|  |  |  |  |  |  |  |  Mean_pc8 < 75.25930199999999
|  |  |  |  |  |  |  |  |  Mean_pc7 < -53.83829
|  |  |  |  |  |  |  |  |  |  Min_pixel_.2 < -279.619812
|  |  |  |  |  |  |  |  |  |  |  WVBI < -0.786151: ceramica(22.0/0.0)
|  |  |  |  |  |  |  |  |  |  |  WVBI >= -0.786151
|  |  |  |  |  |  |  |  |  |  |  |  Max_pixel_.6 < 230.80750999999998: argamassa_cimento_pintado_vermelho(10.0/1.0)
|  |  |  |  |  |  |  |  |  |  |  |  Max_pixel_.6 >= 230.80750999999998: argamassa_cimento_pintado_verde(3.0/5.0)
|  |  |  |  |  |  |  |  |  Min_pixel_.2 >= -279.619812: solo_exposto(26.0/13.0)
|  |  |  |  |  |  |  |  Mean_pc7 >= -53.83829
|  |  |  |  |  |  |  |  |  BRTmN1 < -744.6723884999999
|  |  |  |  |  |  |  |  |  |  NDVI1 < 0.097482: quartzito_amarelo(12.0/1.0)
|  |  |  |  |  |  |  |  |  |  NDVI1 >= 0.097482
|  |  |  |  |  |  |  |  |  |  |  WVNII < 4.4560035
|  |  |  |  |  |  |  |  |  |  |  |  NDBI < 0.088092
|  |  |  |  |  |  |  |  |  |  |  |  |  Max_pixel_.6 < 682.976532
```

```
| | | | | | | | | | | | | | | | Mean_Layer.2 < 979.3125
| | | | | | | | | | | | | | | | | modeMinimu.14 < 818.5: bloquete(4.0/6.0)
| | | | | | | | | | | | | | | | | modeMinimu.14 >= 818.5: argamassa_cimento(5.0/0.0)
| | | | | | | | | | | | | | | | Mean_Layer.2 >= 979.3125: asfalto_pintado_branco(5.0/1.0)
| | | | | | | | | | | | | | | Max_pixel_.6 >= 682.976532: fibra_vidro(4.0/0.0)
| | | | | | | | | | | | | | NDBI >= 0.088092: solo_exposto(6.0/1.0)
| | | | | | | | | | | | | WVNII >= 4.4560035: concreto_pintado_amarelo(9.0/1.0)
| | | | | | | | | | | BRTmN1 >= -744.6723884999999
| | | | | | | | | | | | Mean_pc4 < 4.8196255: concreto(36.0/9.0)
| | | | | | | | | | | | Mean_pc4 >= 4.8196255
| | | | | | | | | | | | | Mean_Layer.4 < 714.8778835: argamassa_cimento_pintado_verde(4.0/0.0)
| | | | | | | | | | | | | Mean_Layer.4 >= 714.8778835
| | | | | | | | | | | | | | Mean_pc3 < -767.9191025: concreto(4.0/3.0)
| | | | | | | | | | | | | | Mean_pc3 >= -767.9191025: asfalto_pintado_branco(16.0/5.0)
| | | | | | | | | Mean_pc8 >= 75.25930199999999
| | | | | | | | | | WVNII < 4.563407: asfalto_pintado_vermelho(35.0/4.0)
| | | | | | | | | | WVNII >= 4.563407: ceramica(4.0/5.0)
| | | | | | WVBI >= -0.4621745
| | | | | | | | Mean_pc7 < 161.086017
| | | | | | | | | Max_pixel_.2 < 27.5255145
| | | | | | | | | | Max_pixel_.9 < 162.6127705
| | | | | | | | | | | Mean_pc3 < -126.0858565
| | | | | | | | | | | | BRTpBoR < 1.635417: piscina_vinil(4.0/0.0)
| | | | | | | | | | | | BRTpBoR >= 1.635417: aco_fosco(4.0/5.0)
| | | | | | | | | | | Mean_pc3 >= -126.0858565: plastico(7.0/0.0)
| | | | | | | | | | Max_pixel_.9 >= 162.6127705: concreto_impermeabilizado(28.0/4.0)
| | | | | | | | | Max_pixel_.2 >= 27.5255145
| | | | | | | | | | Min_pixel_.9 < 29.6380025
| | | | | | | | | | | Mean_pc7 < 9.5712385
| | | | | | | | | | | | NDBI < 0.063358
| | | | | | | | | | | | | NDSI1 < -0.14664549999999998: fibrocimento_crfs(11.0/5.0)
| | | | | | | | | | | | | NDSI1 >= -0.14664549999999998
| | | | | | | | | | | | | | BRTmN1 < -552.664736: galvalume_a(7.0/0.0)
| | | | | | | | | | | | | | BRTmN1 >= -552.664736: aco_fosco(6.0/0.0)
| | | | | | | | | | | | NDBI >= 0.063358: aco_brilhante(10.0/1.0)
| | | | | | | | | | | Mean_pc7 >= 9.5712385
| | | | | | | | | | | | WVSI < -0.079256: aco_ferrugem(4.0/2.0)
| | | | | | | | | | | | WVSI >= -0.079256: aco_brilhante(61.0/1.0)
| | | | | | | | | | Min_pixel_.9 >= 29.6380025
| | | | | | | | | | | Max_pixel_.2 < 33.6353455: fibrocimento_crfs(15.0/2.0)
| | | | | | | | | | | Max_pixel_.2 >= 33.6353455
| | | | | | | | | | | | Mean_mda < 49.2778675
| | | | | | | | | | | | | BRTmN2 < -550.346663
| | | | | | | | | | | | | | NDBI < 0.0324705: galvalume_b(3.0/2.0)
| | | | | | | | | | | | | | NDBI >= 0.0324705: galvalume_c(8.0/4.0)
| | | | | | | | | | | | | BRTmN2 >= -550.346663: aco_fosco(6.0/0.0)
| | | | | | | | | | | | Mean_mda >= 49.2778675: acrilico(8.0/1.0)
| | | | | | | | Mean_pc7 >= 161.086017
| | | | | | | | | Mean_pc6 < -16.420936: galvalume_b(34.0/0.0)
| | | | | | | | | Mean_pc6 >= -16.420936: acrilico(3.0/2.0)
| | | | | | BRTmN1 >= -219.821
| | | | | | | NDBI < 0.1045355
| | | | | | | | NDBI < 0.0650725
| | | | | | | | | Max_pixel_.6 < 413.152985: asfalto_pintado_branco(6.0/2.0)
| | | | | | | | | Max_pixel_.6 >= 413.152985: asfalto(5.0/3.0)
| | | | | | | | NDBI >= 0.0650725
| | | | | | | | | Min_pixel_.13 < 289.0: concreto(4.0/6.0)
```

```
|  |  |  |  |  |  |  |  Min_pixel_.13 >= 289.0: aco_brilhante(3.0/0.0)
|  |  |  |  |  |  NDBI >= 0.1045355
|  |  |  |  |  |  |  Min_pixel_.3 < -699.5814515: concreto(8.0/0.0)
|  |  |  |  |  |  |  Min_pixel_.3 >= -699.5814515
|  |  |  |  |  |  |  |  modeMinimu.15 < -6.0: sombra(6.0/6.0)
|  |  |  |  |  |  |  |  modeMinimu.15 >= -6.0: fibrocimento_amianto(74.0/6.0)
|  |  |  |  Max_pixel_.12 >= 4699.0
|  |  |  |  |  Max_pixel_.2 < 37.901000999999994: galvalume_a(49.0/3.0)
|  |  |  |  |  Max_pixel_.2 >= 37.901000999999994: galvalume_c(63.0/0.0)
|  |  |  BRTmN2 >= -147.8775245
|  |  |  |  modeMinimu.15 < 1146.0
|  |  |  |  |  modeMinimu.16 < 83.0
|  |  |  |  |  |  Min_pixel_.9 < 21.84244: carvao_coque(9.0/2.0)
|  |  |  |  |  |  Min_pixel_.9 >= 21.84244
|  |  |  |  |  |  |  Min_pixel_.9 < 24.8409445
|  |  |  |  |  |  |  |  Min_pixel_.12 < -385.5: sombra(3.0/1.0)
|  |  |  |  |  |  |  |  Min_pixel_.12 >= -385.5: asfalto(8.0/0.0)
|  |  |  |  |  |  |  Min_pixel_.9 >= 24.8409445
|  |  |  |  |  |  |  |  WVBI < -0.49057700000000004: vidro_aramado(5.0/1.0)
|  |  |  |  |  |  |  |  WVBI >= -0.49057700000000004: aco_fosco(3.0/6.0)
|  |  |  |  |  modeMinimu.16 >= 83.0: asfalto(99.0/10.0)
|  |  |  |  modeMinimu.15 >= 1146.0
|  |  |  |  |  BRTmN1 < 443.6045625: aco_fosco(22.0/0.0)
|  |  |  |  |  BRTmN1 >= 443.6045625: piscina_azulejo(6.0/0.0)

Number of Leaf Nodes: 62

Size of the Tree: 123
```

B.4 Árvore gerada por CART Nível II sem MDA

```
CART Decision Tree - SEM MDA

Max_pixel_.4 < -1616.9703370000002
|  NDSI1 < -0.7033435: lago(440.0/0.0)
|  NDSI1 >= -0.7033435
|  |  BRT < -115.054487: sombra(13.0/0.0)
|  |  BRT >= -115.054487: piscina_azulejo(4.0/3.0)
Max_pixel_.4 >= -1616.9703370000002
|  mSWIR < -54.346544
|  |  Min_pixel_.13 < -166.5: sombra(188.0/9.0)
|  |  Min_pixel_.13 >= -166.5: asfalto(3.0/3.0)
|  mSWIR >= -54.346544
|  |  Mean_Layer.5 < -216.93235299999998
|  |  |  modeMinimu.9 < -422.5: sombra(6.0/0.0)
|  |  |  modeMinimu.9 >= -422.5: vegetacao_arborea(132.0/4.0)
|  |  Mean_Layer.5 >= -216.93235299999998
|  |  |  BRTmN2 < -166.29073549999998
|  |  |  |  Max_pixel_.12 < 4699.0
|  |  |  |  |  BRTmN1 < -219.821
|  |  |  |  |  |  NDSI1 < -0.379429
|  |  |  |  |  |  |  Mean_Layer.5 < 638.1011555
|  |  |  |  |  |  |  |  modeMinimu.13 < -45.0
|  |  |  |  |  |  |  |  |  Min_pixel_.10 < -62.5: sombra(4.0/0.0)
```

```
| | | | | | | | | Min_pixel_.10 >= -62.5: vegetacao_arborea(7.0/3.0)
| | | | | | | | modeMinimu.13 >= -45.0: vegetacao_rasteira(42.0/4.0)
| | | | | | | Mean_Layer.5 >= 638.1011555
| | | | | | | | Mean_pc8 < 75.25930199999999
| | | | | | | | | Min_pixel_.2 < -279.619812
| | | | | | | | | | Min_pixel_.4 < -358.029907
| | | | | | | | | | | Mean_pc7 < -361.1365745: asfalto_pintado_vermelho(3.0/0.0)
| | | | | | | | | | | Mean_pc7 >= -361.1365745: ceramica(23.0/1.0)
| | | | | | | | | | Min_pixel_.4 >= -358.029907: argamassa_cimento_pintado_vermelho(10.0/1.0)
| | | | | | | | | Min_pixel_.2 >= -279.619812
| | | | | | | | | | Mean_pc2 < 737.1148835: concreto(11.0/5.0)
| | | | | | | | | | Mean_pc2 >= 737.1148835
| | | | | | | | | | | Mean_pc7 < -3.3206075: solo_exposto(33.0/3.0)
| | | | | | | | | | | Mean_pc7 >= -3.3206075: concreto(3.0/3.0)
| | | | | | | | Mean_pc8 >= 75.25930199999999
| | | | | | | | | WVNII < 4.563407: asfalto_pintado_vermelho(35.0/2.0)
| | | | | | | | | WVNII >= 4.563407: ceramica(4.0/2.0)
| | | | | | NDSI1 >= -0.379429
| | | | | | | Mean_pc4 < -106.7165055
| | | | | | | | NDSI1 < -0.1553095
| | | | | | | | | Mean_Layer.2 < 827.942258: concreto(6.0/2.0)
| | | | | | | | | Mean_Layer.2 >= 827.942258
| | | | | | | | | | NDVI1 < 0.07762150000000001: aco_ferrugem(4.0/0.0)
| | | | | | | | | | NDVI1 >= 0.07762150000000001
| | | | | | | | | | | Border_ind < 1.5166665
| | | | | | | | | | | | Roundness < 0.179312
| | | | | | | | | | | | | Max_pixel_.8 < 412.7784575: fibrocimento_crfs(4.0/5.0)
| | | | | | | | | | | | | Max_pixel_.8 >= 412.7784575: acrilico(3.0/0.0)
| | | | | | | | | | | | Roundness >= 0.179312: aco_brilhante(6.0/1.0)
| | | | | | | | | | | Border_ind >= 1.5166665: concreto_impermeabilizado(7.0/2.0)
| | | | | | | | NDSI1 >= -0.1553095
| | | | | | | | | Max_pixel_.7 < 542.1188354999999: aco_brilhante(58.0/3.0)
| | | | | | | | | Max_pixel_.7 >= 542.1188354999999: galvalume_c(6.0/1.0)
| | | | | | | Mean_pc4 >= -106.7165055
| | | | | | | | Min_pixel_.6 < -1805.6899414999998
| | | | | | | | | NDSI2 < -0.1562705
| | | | | | | | | | Max_pixel_.7 < 28.300668999999996: quartzito_amarelo(14.0/0.0)
| | | | | | | | | | Max_pixel_.7 >= 28.300668999999996: aco_brilhante(4.0/4.0)
| | | | | | | | | NDSI2 >= -0.1562705: galvalume_b(36.0/1.0)
| | | | | | | | Min_pixel_.6 >= -1805.6899414999998
| | | | | | | | | WVBI < -0.476595
| | | | | | | | | | Mean_pc7 < 24.271832500000002
| | | | | | | | | | | WVSI < -0.0687625: asfalto_pintado_branco(24.0/6.0)
| | | | | | | | | | | WVSI >= -0.0687625
| | | | | | | | | | | | Asymmetry < 0.8252225: argamassa_cimento_pintado_verde(7.0/0.0)
| | | | | | | | | | | | Asymmetry >= 0.8252225: aco_brilhante(3.0/4.0)
| | | | | | | | | | Mean_pc7 >= 24.271832500000002
| | | | | | | | | | | Mean_pc2 < 409.24731099999997
| | | | | | | | | | | | Mean_pc6 < 53.815277: concreto(21.0/2.0)
| | | | | | | | | | | | Mean_pc6 >= 53.815277: asfalto_pintado_branco(3.0/0.0)
| | | | | | | | | | | Mean_pc2 >= 409.24731099999997
| | | | | | | | | | | | WVNII < 4.4560035
| | | | | | | | | | | | | Asymmetry < 0.9868159999999999: argamassa_cimento(4.0/5.0)
| | | | | | | | | | | | | Asymmetry >= 0.9868159999999999: pastilha_esmaltada(4.0/0.0)
| | | | | | | | | | | WVNII >= 4.4560035: concreto_pintado_amarelo(7.0/1.0)
| | | | | | | | | WVBI >= -0.476595
| | | | | | | | | | NDSI1 < -0.08185049999999999
```

```
| | | | | | | | | | | Min_pixel_.6 < -305.6911925
| | | | | | | | | | | | Mean_pc6 < -401.0552095: concreto_impermeabilizado(12.0/1.0)
| | | | | | | | | | | | Mean_pc6 >= -401.0552095
| | | | | | | | | | | | | Mean_pc5 < 210.747748
| | | | | | | | | | | | | | Mean_pc8 < 89.3937
| | | | | | | | | | | | | | | BRTpBoR < 1.6368185
| | | | | | | | | | | | | | | | Mean_pc7 < 35.3411275
| | | | | | | | | | | | | | | | | Compactnes < 1.1883115: piscina_vinil(4.0/0.0)
| | | | | | | | | | | | | | | | | Compactnes >= 1.1883115: aco_brilhante(4.0/6.0)
| | | | | | | | | | | | | | | | Mean_pc7 >= 35.3411275: concreto_impermeabilizado(7.0/1.0)
| | | | | | | | | | | | | | | BRTpBoR >= 1.6368185
| | | | | | | | | | | | | | | | Min_pixel_.11 < 2048.5
| | | | | | | | | | | | | | | | Mean_pc4 < 10.6190205
| | | | | | | | | | | | | | | | | Asymmetry < 0.60295: fibrocimento_crfs(3.0/0.0)
| | | | | | | | | | | | | | | | | Asymmetry >= 0.60295: aco_brilhante(3.0/0.0)
| | | | | | | | | | | | | | | | Mean_pc4 >= 10.6190205: asfalto_pintado_branco(5.0/7.0)
| | | | | | | | | | | | | | | | Min_pixel_.11 >= 2048.5: fibrocimento_crfs(20.0/9.0)
| | | | | | | | | | | | | | Mean_pc8 >= 89.3937: acrilico(6.0/1.0)
| | | | | | | | | | | | | Mean_pc5 >= 210.747748: galvalume_a(7.0/0.0)
| | | | | | | | | | | Min_pixel_.6 >= -305.6911925: plastico(8.0/0.0)
| | | | | | | | | | NDSI1 >= -0.08185049999999999: aco_fosco(12.0/2.0)
| | | | | BRTmN1 >= -219.821
| | | | | | NDBI < 0.1045355
| | | | | | | WVBI < -0.222453
| | | | | | | | NDBI < 0.0650725
| | | | | | | | | Max_pixel_.6 < 413.152985: asfalto_pintado_branco(5.0/0.0)
| | | | | | | | | Max_pixel_.6 >= 413.152985: asfalto(4.0/2.0)
| | | | | | | | NDBI >= 0.0650725
| | | | | | | | | Min_pixel_.12 < 143.5: concreto(3.0/4.0)
| | | | | | | | | Min_pixel_.12 >= 143.5: aco_brilhante(3.0/0.0)
| | | | | | | WVBI >= -0.222453: sombra(3.0/0.0)
| | | | | | NDBI >= 0.1045355
| | | | | | | Min_pixel_.3 < -699.5814515: concreto(8.0/0.0)
| | | | | | | Min_pixel_.3 >= -699.5814515
| | | | | | | | modeMinimu.15 < -38.0: sombra(6.0/5.0)
| | | | | | | | modeMinimu.15 >= -38.0: fibrocimento_amianto(72.0/6.0)
| | | | Max_pixel_.12 >= 4699.0
| | | | | Min_pixel_.5 < -483.89689599999997: galvalume_c(58.0/3.0)
| | | | | Min_pixel_.5 >= -483.89689599999997
| | | | | | Max_pixel_ < 16028.0: galvalume_a(46.0/2.0)
| | | | | | Max_pixel_ >= 16028.0: galvalume_c(5.0/0.0)
| | | BRTmN2 >= -166.29073549999998
| | | | modeMinimu.13 < 1273.5
| | | | | modeMinimu.16 < 83.0
| | | | | | Min_pixel_.3 < -87.86516950000001
| | | | | | | Max_pixel_.6 < 221.10657500000002: vidro_aramado(5.0/2.0)
| | | | | | | Max_pixel_.6 >= 221.10657500000002
| | | | | | | | Mean_pc8 < -7.3349595: sombra(3.0/1.0)
| | | | | | | | Mean_pc8 >= -7.3349595: asfalto(8.0/8.0)
| | | | | | Min_pixel_.3 >= -87.86516950000001: carvao_coque(9.0/6.0)
| | | | | modeMinimu.16 >= 83.0: asfalto(101.0/12.0)
| | | | modeMinimu.13 >= 1273.5
| | | | | BRTmN1 < 443.6045625: aco_fosco(23.0/0.0)
| | | | | BRTmN1 >= 443.6045625: piscina_azulejo(3.0/0.0)

Number of Leaf Nodes: 67
```

```
Size of the Tree: 133
```

B.5 Árvore gerada por C4.5 Nível II sem MDA

```
 J48 pruned tree - SEM MDA (90 atributos)
------------------

Max_pixel_.4 <= -1591.778076
|   NDVI1 <= 0.511419
|   |   Max_pixel_.1 <= -121: sombra (16.0)
|   |   Max_pixel_.1 > -121
|   |   |   Border_ind <= 1.823529: fibrocimento_amianto (3.0/1.0)
|   |   |   Border_ind > 1.823529: piscina_azulejo (5.0)
|   NDVI1 > 0.511419: lago (442.0)
Max_pixel_.4 > -1591.778076
|   Mean_Layer.5 <= 1164.12
|   |   modeMinimu.15 <= -6
|   |   |   Mean_pc3 <= 328.335594
|   |   |   |   Mean_pc5 <= -106.656215
|   |   |   |   |   BRTpBoR <= 0.908808: ceramica (7.0)
|   |   |   |   |   BRTpBoR > 0.908808: concreto (3.0)
|   |   |   |   Mean_pc5 > -106.656215
|   |   |   |   |   Min_pixel_.14 <= -149
|   |   |   |   |   |   modeMinimu.9 <= -60
|   |   |   |   |   |   |   Max_pixel_.3 <= -59347.671875: asfalto (2.0)
|   |   |   |   |   |   |   Max_pixel_.3 > -59347.671875
|   |   |   |   |   |   |   |   WVBI <= 4.227654
|   |   |   |   |   |   |   |   |   BRTmN1 <= 59.86284
|   |   |   |   |   |   |   |   |   |   Min_pixel_.5 <= -366.194794: sombra (22.0)
|   |   |   |   |   |   |   |   |   |   Min_pixel_.5 > -366.194794
|   |   |   |   |   |   |   |   |   |   |   Mean_pc7 <= -25.968152: vegetacao_arborea (4.0)
|   |   |   |   |   |   |   |   |   |   |   Mean_pc7 > -25.968152
|   |   |   |   |   |   |   |   |   |   |   |   BRT <= -335.122362: vegetacao_arborea (2.0)
|   |   |   |   |   |   |   |   |   |   |   |   BRT > -335.122362: sombra (8.0)
|   |   |   |   |   |   |   |   |   BRTmN1 > 59.86284
|   |   |   |   |   |   |   |   |   |   Mean_pc4 <= -70.999516
|   |   |   |   |   |   |   |   |   |   |   Mean_pc4 <= -73.396498: sombra (10.0)
|   |   |   |   |   |   |   |   |   |   |   Mean_pc4 > -73.396498: concreto (2.0)
|   |   |   |   |   |   |   |   |   |   Mean_pc4 > -70.999516: sombra (159.0)
|   |   |   |   |   |   |   |   WVBI > 4.227654
|   |   |   |   |   |   |   |   |   BRT <= 6.084483: asfalto (2.0/1.0)
|   |   |   |   |   |   |   |   |   BRT > 6.084483: sombra (2.0)
|   |   |   |   |   |   modeMinimu.9 > -60
|   |   |   |   |   |   |   Mean_Layer.2 <= -185.176471
|   |   |   |   |   |   |   |   Mean_Layer.2 <= -277.349876: sombra (2.0)
|   |   |   |   |   |   |   |   Mean_Layer.2 > -277.349876: vegetacao_arborea (3.0)
|   |   |   |   |   |   |   Mean_Layer.2 > -185.176471
|   |   |   |   |   |   |   |   modeMinimu.16 <= -191: policarbonato (2.0)
|   |   |   |   |   |   |   |   modeMinimu.16 > -191: fibrocimento_amianto (3.0)
|   |   |   |   |   Min_pixel_.14 > -149
|   |   |   |   |   |   Max_pixel_.10 <= 114.59613
|   |   |   |   |   |   |   Asymmetry <= 0.977404: aco_brilhante (3.0/1.0)
|   |   |   |   |   |   |   Asymmetry > 0.977404: fibrocimento_amianto (2.0)
|   |   |   |   |   |   Max_pixel_.10 > 114.59613
|   |   |   |   |   |   |   Compactnes <= 2.16: concreto (3.0)
```

```
|   |   |   |   |   |   |   Compactnes > 2.16: sombra (2.0/1.0)
|   |   |   Mean_pc3 > 328.335594
|   |   |   |   Min_pixel_ <= -593
|   |   |   |   |   Max_pixel_.1 <= 1587
|   |   |   |   |   |   Border_ind <= 1.84375: sombra (2.0)
|   |   |   |   |   |   Border_ind > 1.84375: vegetacao_arborea (3.0)
|   |   |   |   |   Max_pixel_.1 > 1587: vegetacao_arborea (134.0/2.0)
|   |   |   |   Min_pixel_ > -593: vegetacao_rasteira (3.0/1.0)
|   |   modeMinimu.15 > -6
|   |   |   NDVI1 <= 0.15524
|   |   |   |   NDVI2 <= 0.078043
|   |   |   |   |   WVSI <= -0.203472
|   |   |   |   |   |   Min_pixel_.8 <= -63285.695313
|   |   |   |   |   |   |   Border_ind <= 1.214286: asfalto (3.0)
|   |   |   |   |   |   |   Border_ind > 1.214286
|   |   |   |   |   |   |   |   Asymmetry <= 0.694804: argamassa_cimento_pintado_vermelho (2.0/1.0)
|   |   |   |   |   |   |   |   Asymmetry > 0.694804: vegetacao_rasteira (2.0/1.0)
|   |   |   |   |   |   Min_pixel_.8 > -63285.695313: carvao_coque (9.0)
|   |   |   |   |   WVSI > -0.203472
|   |   |   |   |   |   Mean_pc6 <= -17.036502
|   |   |   |   |   |   |   Max_pixel_.3 <= -58705.050781
|   |   |   |   |   |   |   |   Max_pixel_.3 <= -59863.351563: asfalto (3.0)
|   |   |   |   |   |   |   |   Max_pixel_.3 > -59863.351563: policarbonato (3.0)
|   |   |   |   |   |   |   Max_pixel_.3 > -58705.050781
|   |   |   |   |   |   |   |   Asymmetry <= 0.936634: fibrocimento_amianto (2.0)
|   |   |   |   |   |   |   |   Asymmetry > 0.936634: aco_fosco (3.0)
|   |   |   |   |   |   Mean_pc6 > -17.036502
|   |   |   |   |   |   |   Mean_Layer.1 <= 897.329861
|   |   |   |   |   |   |   |   modeMinimu.11 <= -70: vidro_aramado (3.0)
|   |   |   |   |   |   |   |   modeMinimu.11 > -70
|   |   |   |   |   |   |   |   |   modeMinimu.15 <= 235
|   |   |   |   |   |   |   |   |   |   Asymmetry <= 0.658237: vidro_aramado (3.0)
|   |   |   |   |   |   |   |   |   |   Asymmetry > 0.658237: asfalto (6.0)
|   |   |   |   |   |   |   |   |   modeMinimu.15 > 235: asfalto (99.0)
|   |   |   |   |   |   |   Mean_Layer.1 > 897.329861
|   |   |   |   |   |   |   |   Max_pixel_.6 <= 372.334564: asfalto_pintado_branco (6.0/1.0)
|   |   |   |   |   |   |   |   Max_pixel_.6 > 372.334564: asfalto (5.0/1.0)
|   |   |   |   NDVI2 > 0.078043
|   |   |   |   |   Min_pixel_.3 <= -123.692596
|   |   |   |   |   |   modeMinimu.2 <= -939.210632: concreto (3.0)
|   |   |   |   |   |   modeMinimu.2 > -939.210632
|   |   |   |   |   |   |   NDBI <= 0.103527: concreto (3.0)
|   |   |   |   |   |   |   NDBI > 0.103527
|   |   |   |   |   |   |   |   Max_pixel_.16 <= 503.276459: fibrocimento_amianto (58.0)
|   |   |   |   |   |   |   |   Max_pixel_.16 > 503.276459
|   |   |   |   |   |   |   |   |   Max_pixel_.16 <= 546.665344: concreto (4.0/1.0)
|   |   |   |   |   |   |   |   |   Max_pixel_.16 > 546.665344: fibrocimento_amianto (11.0)
|   |   |   |   |   Min_pixel_.3 > -123.692596
|   |   |   |   |   |   Density <= 1.299588: asfalto (3.0)
|   |   |   |   |   |   Density > 1.299588: asfalto_pintado_branco (2.0/1.0)
|   |   |   NDVI1 > 0.15524
|   |   |   |   NDSI1 <= -0.393552
|   |   |   |   |   Min_pixel_.16 <= -388
|   |   |   |   |   |   BRT <= 86.180346: sombra (3.0/1.0)
|   |   |   |   |   |   BRT > 86.180346: fibrocimento_amianto (3.0/1.0)
|   |   |   |   |   Min_pixel_.16 > -388
|   |   |   |   |   |   Mean_Layer.7 <= 1600.552632
```

```
|   |   |   |   |   |   |   NDVI3 <= 0.352876: concreto (10.0/1.0)
|   |   |   |   |   |   |   NDVI3 > 0.352876
|   |   |   |   |   |   |   |   modeMinimu.12 <= 668
|   |   |   |   |   |   |   |   |   modeMinimu.12 <= -411: vegetacao_arborea (2.0/1.0)
|   |   |   |   |   |   |   |   |   modeMinimu.12 > -411
|   |   |   |   |   |   |   |   |   |   Mean_pc1 <= -62255.825103: solo_exposto (2.0/1.0)
|   |   |   |   |   |   |   |   |   |   Mean_pc1 > -62255.825103: vegetacao_rasteira (48.0/3.0)
|   |   |   |   |   |   |   |   modeMinimu.12 > 668: solo_exposto (3.0)
|   |   |   |   |   |   Mean_Layer.7 > 1600.552632: solo_exposto (7.0)
|   |   |   |   NDSI1 > -0.393552
|   |   |   |   |   Mean_pc8 <= -68.662459: argamassa_cimento_pintado_verde (7.0)
|   |   |   |   |   Mean_pc8 > -68.662459
|   |   |   |   |   |   Compactnes <= 1.865341: policarbonato (3.0/1.0)
|   |   |   |   |   |   Compactnes > 1.865341: plastico (4.0)
|   Mean_Layer.5 > 1164.12
|   |   NDSI2 <= -0.420767
|   |   |   modeMinimu.1 <= -279.613403
|   |   |   |   Min_pixel_.4 <= -386.32843
|   |   |   |   |   WVBI <= -0.845891: ceramica (17.0)
|   |   |   |   |   WVBI > -0.845891
|   |   |   |   |   |   Mean_Layer.3 <= 1576.847826
|   |   |   |   |   |   |   Min_pixel_ <= 606: ceramica (2.0)
|   |   |   |   |   |   |   Min_pixel_ > 606: asfalto_pintado_vermelho (38.0/1.0)
|   |   |   |   |   |   Mean_Layer.3 > 1576.847826
|   |   |   |   |   |   |   Compactnes <= 1.316239: concreto_pintado_vermelho (2.0)
|   |   |   |   |   |   |   Compactnes > 1.316239: ceramica (5.0)
|   |   |   |   Min_pixel_.4 > -386.32843: argamassa_cimento_pintado_vermelho (10.0)
|   |   |   modeMinimu.1 > -279.613403
|   |   |   |   NDSI1 <= -0.454906
|   |   |   |   |   Min_pixel_.3 <= 51.887234: solo_exposto (25.0/1.0)
|   |   |   |   |   Min_pixel_.3 > 51.887234: ceramica (3.0)
|   |   |   |   NDSI1 > -0.454906
|   |   |   |   |   Max_pixel_.8 <= 643.405457
|   |   |   |   |   |   Mean_pc4 <= -62.143877: concreto (9.0/1.0)
|   |   |   |   |   |   Mean_pc4 > -62.143877: bloquete (3.0)
|   |   |   |   |   Max_pixel_.8 > 643.405457: asfalto_pintado_vermelho (3.0/1.0)
|   |   NDSI2 > -0.420767
|   |   |   Max_pixel_.5 <= 7529
|   |   |   |   Mean_pc7 <= 148.335251
|   |   |   |   |   BRTpBoR <= 1.610843
|   |   |   |   |   |   BRTpBoRE <= 1.041442
|   |   |   |   |   |   |   BRTpBoR <= 1.284638: quartzito_amarelo (13.0)
|   |   |   |   |   |   |   BRTpBoR > 1.284638: concreto_pintado_amarelo (3.0/1.0)
|   |   |   |   |   |   BRTpBoRE > 1.041442
|   |   |   |   |   |   |   Mean_pc1 <= -67218.146273
|   |   |   |   |   |   |   |   Max_pixel_.8 <= 258.958069
|   |   |   |   |   |   |   |   |   Border_ind <= 1.52381: piscina_vinil (4.0)
|   |   |   |   |   |   |   |   |   Border_ind > 1.52381: concreto_impermeabilizado (2.0)
|   |   |   |   |   |   |   |   Max_pixel_.8 > 258.958069
|   |   |   |   |   |   |   |   |   NDVI1 <= 0.076137: aco_ferrugem (7.0)
|   |   |   |   |   |   |   |   |   NDVI1 > 0.076137: aco_brilhante (2.0)
|   |   |   |   |   |   |   Mean_pc1 > -67218.146273
|   |   |   |   |   |   |   |   Mean_pc6 <= -71.905134
|   |   |   |   |   |   |   |   |   Border_ind <= 1.44186: argamassa_cimento (3.0/2.0)
|   |   |   |   |   |   |   |   |   Border_ind > 1.44186: concreto_impermeabilizado (3.0/2.0)
|   |   |   |   |   |   |   |   Mean_pc6 > -71.905134
|   |   |   |   |   |   |   |   |   Mean_pc7 <= 79.487345
```

```
|   |   |   |   |   |   |   |   |   |   Mean_pc2 <= 978.069615
|   |   |   |   |   |   |   |   |   |   |   Max_pixel_.3 <= -58740.054688
|   |   |   |   |   |   |   |   |   |   |   |   NDBI <= 0.0764
|   |   |   |   |   |   |   |   |   |   |   |   |   Max_pixel_.6 <= 439.938904: asfalto_pintado_branco (25.0/2.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   Max_pixel_.6 > 439.938904
|   |   |   |   |   |   |   |   |   |   |   |   |   |   BRTmN2 <= -803.064783: asfalto_pintado_branco (2.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   |   BRTmN2 > -803.064783: concreto (4.0)
|   |   |   |   |   |   |   |   |   |   |   |   NDBI > 0.0764: concreto (4.0)
|   |   |   |   |   |   |   |   |   |   |   Max_pixel_.3 > -58740.054688
|   |   |   |   |   |   |   |   |   |   |   |   Border_ind <= 2.913043: aco_brilhante (3.0)
|   |   |   |   |   |   |   |   |   |   |   |   Border_ind > 2.913043: asfalto_pintado_branco (2.0)
|   |   |   |   |   |   |   |   |   |   Mean_pc2 > 978.069615: argamassa_cimento (4.0)
|   |   |   |   |   |   |   |   |   Mean_pc7 > 79.487345: concreto (16.0)
|   |   |   |   |   BRTpBoR > 1.610843
|   |   |   |   |   |   BRTmN2 <= -581.533626
|   |   |   |   |   |   |   Mean_pc5 <= 181.641185
|   |   |   |   |   |   |   |   BRTpBoRE <= 1.285574
|   |   |   |   |   |   |   |   |   Density <= 1.266007
|   |   |   |   |   |   |   |   |   |   Mean_pc3 <= -363.560767: fibra_vidro (5.0)
|   |   |   |   |   |   |   |   |   |   Mean_pc3 > -363.560767: policarbonato (2.0/1.0)
|   |   |   |   |   |   |   |   |   Density > 1.266007: asfalto_pintado_branco (3.0/1.0)
|   |   |   |   |   |   |   |   BRTpBoRE > 1.285574
|   |   |   |   |   |   |   |   |   BRTmN2 <= -2557.322443
|   |   |   |   |   |   |   |   |   |   Border_ind <= 1.1: fibra_vidro (4.0/1.0)
|   |   |   |   |   |   |   |   |   |   Border_ind > 1.1: concreto_impermeabilizado (15.0)
|   |   |   |   |   |   |   |   |   BRTmN2 > -2557.322443
|   |   |   |   |   |   |   |   |   |   Mean_pc5 <= -39.743076
|   |   |   |   |   |   |   |   |   |   |   Max_pixel_.1 <= 3573: fibrocimento_amianto (2.0/1.0)
|   |   |   |   |   |   |   |   |   |   |   Max_pixel_.1 > 3573
|   |   |   |   |   |   |   |   |   |   |   |   Mean_pc8 <= -17.551482
|   |   |   |   |   |   |   |   |   |   |   |   |   Max_pixel_.9 <= 250.075623: aco_brilhante (5.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   Max_pixel_.9 > 250.075623: concreto_impermeabilizado (3.0/1.0)
|   |   |   |   |   |   |   |   |   |   |   |   Mean_pc8 > -17.551482
|   |   |   |   |   |   |   |   |   |   |   |   |   Mean_pc7 <= 34.399286: fibrocimento_crfs (15.0/1.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   Mean_pc7 > 34.399286
|   |   |   |   |   |   |   |   |   |   |   |   |   |   BRTpBoR <= 1.693648: concreto_impermeabilizado (8.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   |   BRTpBoR > 1.693648: fibrocimento_crfs (6.0/1.0)
|   |   |   |   |   |   |   |   |   Mean_pc5 > -39.743076
|   |   |   |   |   |   |   |   |   |   Rectangula <= 0.939198
|   |   |   |   |   |   |   |   |   |   |   Mean_pc8 <= 22.060468: aco_brilhante (66.0/2.0)
|   |   |   |   |   |   |   |   |   |   |   Mean_pc8 > 22.060468
|   |   |   |   |   |   |   |   |   |   |   |   BRTpBoR <= 1.731752: asfalto_pintado_branco (4.0)
|   |   |   |   |   |   |   |   |   |   |   |   BRTpBoR > 1.731752
|   |   |   |   |   |   |   |   |   |   |   |   |   Border_ind <= 1.258065: aco_brilhante (3.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   Border_ind > 1.258065
|   |   |   |   |   |   |   |   |   |   |   |   |   |   BRTmN1 <= -643.191051: fibrocimento_crfs (4.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   |   BRTmN1 > -643.191051: concreto_impermeabilizado (2.0)
|   |   |   |   |   |   |   |   |   |   Rectangula > 0.939198: fibrocimento_crfs (5.0)
|   |   |   |   |   |   |   Mean_pc5 > 181.641185
|   |   |   |   |   |   |   |   Mean_pc4 <= 19.722885
|   |   |   |   |   |   |   |   |   BRT <= 3624.372297: concreto (2.0/1.0)
|   |   |   |   |   |   |   |   |   BRT > 3624.372297: galvalume_c (4.0)
|   |   |   |   |   |   |   |   Mean_pc4 > 19.722885: galvalume_a (10.0/1.0)
|   |   |   |   |   |   BRTmN2 > -581.533626
|   |   |   |   |   |   |   Mean_pc3 <= -129.550527
|   |   |   |   |   |   |   |   NDSI2 <= -0.130317
|   |   |   |   |   |   |   |   |   Border_ind <= 1.341463: acrilico (3.0)
```

```
|   |   |   |   |   |   |   |   |   Border_ind > 1.341463
|   |   |   |   |   |   |   |   |   |   BRTmN1 <= -209.785433: asfalto_pintado_branco (3.0)
|   |   |   |   |   |   |   |   |   |   BRTmN1 > -209.785433: concreto (2.0/1.0)
|   |   |   |   |   |   |   |   NDSI2 > -0.130317
|   |   |   |   |   |   |   |   |   NDBI <= 0.056124: aco_fosco (39.0/1.0)
|   |   |   |   |   |   |   |   |   NDBI > 0.056124: aco_brilhante (3.0)
|   |   |   |   |   |   |   Mean_pc3 > -129.550527: plastico (4.0/1.0)
|   |   |   |   Mean_pc7 > 148.335251
|   |   |   |   |   NDSI2 <= -0.156084
|   |   |   |   |   |   Max_pixel_.8 <= 439.392212: concreto_pintado_amarelo (8.0/1.0)
|   |   |   |   |   |   Max_pixel_.8 > 439.392212
|   |   |   |   |   |   |   Elliptic_F <= 0.609756: pastilha_esmaltada (4.0/1.0)
|   |   |   |   |   |   |   Elliptic_F > 0.609756: acrilico (4.0/1.0)
|   |   |   |   |   NDSI2 > -0.156084
|   |   |   |   |   |   Max_pixel_.12 <= 4751: galvalume_b (35.0)
|   |   |   |   |   |   Max_pixel_.12 > 4751: piscina_azulejo (2.0)
|   |   |   Max_pixel_.5 > 7529
|   |   |   |   Min_pixel_.5 <= -851.169128: galvalume_c (52.0)
|   |   |   |   Min_pixel_.5 > -851.169128
|   |   |   |   |   Shape_inde <= 1.474024
|   |   |   |   |   |   NDVI1 <= 0.202716
|   |   |   |   |   |   |   Max_pixel_ <= 12698: lona (8.0/1.0)
|   |   |   |   |   |   |   Max_pixel_ > 12698: galvalume_c (7.0)
|   |   |   |   |   |   NDVI1 > 0.202716: acrilico (3.0)
|   |   |   |   |   Shape_inde > 1.474024
|   |   |   |   |   |   Mean_pc4 <= -137.351456
|   |   |   |   |   |   |   BRTmN1 <= -1591.796992: galvalume_c (5.0)
|   |   |   |   |   |   |   BRTmN1 > -1591.796992: aco_brilhante (6.0/1.0)
|   |   |   |   |   |   Mean_pc4 > -137.351456
|   |   |   |   |   |   |   BRTpBoR <= 1.670762: galvalume_c (2.0/1.0)
|   |   |   |   |   |   |   BRTpBoR > 1.670762: galvalume_a (48.0/1.0)

Number of Leaves  :  126

Size of the tree :  251
```

B.6 Árvore gerada por C5.0 Nível II sem MDA

```
C5.0 [Release 2.07 GPL Edition] - (Nível II sem MDA)
-------------------------------

Class specified by attribute 'outcome'

Read 1796 cases (90 attributes) from undefined.data

Decision tree:

Max_pixel_.4 <= -1591.778:
:...NDSI2 <= -0.725163: lago (442)
:   NDSI2 > -0.725163:
:   :...Max_pixel_.16 > 1115.027: piscina_azulejo (5)
:       Max_pixel_.16 <= 1115.027:
:       :...NDVI3 <= 0.221289: fibrocimento_amianto (3/1)
:           NDVI3 > 0.221289: sombra (16)
```

```
Max_pixel_.4 > -1591.778:
:...Mean_Layer.5 <= 1164.12:
    :...modeMinimu.15 <= -6:
    :   :...Mean_pc3 > 328.3356:
    :   :   :...Min_pixel_ > -593: vegetacao_rasteira (3/1)
    :   :   :   Min_pixel_ <= -593:
    :   :   :   :...Max_pixel_.1 > 1587: vegetacao_arborea (134/2)
    :   :   :       Max_pixel_.1 <= 1587:
    :   :   :       :...modeMinimu.15 <= -658: sombra (2)
    :   :   :           modeMinimu.15 > -658: vegetacao_arborea (3)
    :   :   Mean_pc3 <= 328.3356:
    :   :   :...Mean_pc5 <= -106.6562:
    :   :       :...NDBI <= 0.290015: ceramica (7)
    :   :       :   NDBI > 0.290015: concreto (3)
    :   :       Mean_pc5 > -106.6562:
    :   :       :...Min_pixel_.13 > -208:
    :   :           :...Mean_pc6 <= -20.72629:
    :   :           :   :...WVSI <= 0.002854: concreto (4)
    :   :           :   :   WVSI > 0.002854: fibrocimento_amianto (4/1)
    :   :           :   Mean_pc6 > -20.72629:
    :   :           :   :...Min_pixel_.1 <= -70.56067: aco_brilhante (3/1)
    :   :           :       Min_pixel_.1 > -70.56067:
    :   :           :       :...Asymmetry <= 0.487199: vegetacao_arborea (2)
    :   :           :           Asymmetry > 0.487199: sombra (2)
    :   :           Min_pixel_.13 <= -208:
    :   :           :...modeMinimu.9 <= -102: sombra (208/9)
    :   :               modeMinimu.9 > -102:
    :   :               :...modeMinimu.14 <= -413: policarbonato (3/1)
    :   :                   modeMinimu.14 > -413:
    :   :                   :...Mean_pc6 <= -12.41609: fibrocimento_amianto (2)
    :   :                       Mean_pc6 > -12.41609:
    :   :                       :...Max_pixel_.12 <= 143: sombra (3)
    :   :                           Max_pixel_.12 > 143: vegetacao_arborea (2)
    :   modeMinimu.15 > -6:
    :   :...NDVI1 > 0.15524:
    :       :...NDSI1 > -0.393552:
    :       :   :...Mean_pc8 <= -68.66246: argamassa_cimento_pintado_verde (7)
    :       :   :   Mean_pc8 > -68.66246:
    :       :   :   :...Compactnes <= 1.865341: policarbonato (3/1)
    :       :   :       Compactnes > 1.865341: plastico (4)
    :       :   NDSI1 <= -0.393552:
    :       :   :...Min_pixel_.16 <= -388:
    :       :       :...Mean_pc8 > 31.2787: fibrocimento_amianto (2)
    :       :       :   Mean_pc8 <= 31.2787:
    :       :       :   :...Border_ind <= 2.619048: sombra (2)
    :       :       :       Border_ind > 2.619048: concreto (2/1)
    :       :       Min_pixel_.16 > -388:
    :       :       :...Mean_Layer.7 > 1600.553: solo_exposto (7)
    :       :           Mean_Layer.7 <= 1600.553:
    :       :           :...NDVI3 <= 0.352876: concreto (10/1)
    :       :               NDVI3 > 0.352876:
    :       :               :...modeMinimu.12 > 668: solo_exposto (3)
    :       :                   modeMinimu.12 <= 668:
    :       :                   :...modeMinimu.12 <= -411: concreto (2/1)
    :       :                       modeMinimu.12 > -411:
    :       :                       :...NDSI1 <= -0.499986: vegetacao_rasteira (48/3)
    :       :                           NDSI1 > -0.499986: concreto (2/1)
```

```
:       NDVI1 <= 0.15524:
:       :...NDVI2 > 0.078043:
:           :...Min_pixel_.3 > -123.6926:
:           :   :...Density <= 1.299588: asfalto (3)
:           :   :   Density > 1.299588: argamassa_cimento_pintado_verde (2/1)
:           :   Min_pixel_.3 <= -123.6926:
:           :   :...Min_pixel_.3 <= -939.2106: concreto (3)
:           :       Min_pixel_.3 > -939.2106:
:           :       :...NDBI <= 0.103527: concreto (3)
:           :           NDBI > 0.103527:
:           :           :...Max_pixel_.16 <= 503.2765: fibrocimento_amianto (58)
:           :               Max_pixel_.16 > 503.2765: [S1]
:           NDVI2 <= 0.078043:
:           :...WVSI <= -0.203472:
:               :...Min_pixel_.12 > -39: asfalto (4/1)
:               :   Min_pixel_.12 <= -39:
:               :   :...Min_pixel_.3 <= -142.9033: concreto (2/1)
:               :       Min_pixel_.3 > -142.9033: carvao_coque (10/1)
:               WVSI > -0.203472:
:               :...Mean_pc6 <= -17.0365:
:                   :...Max_pixel_.3 <= -58705.05:
:                   :   :...Max_pixel_.3 <= -59863.35: asfalto (3)
:                   :   :   Max_pixel_.3 > -59863.35: policarbonato (3)
:                   :   Max_pixel_.3 > -58705.05:
:                   :   :...Mean_pc7 <= -10.68193: aco_fosco (3)
:                   :       Mean_pc7 > -10.68193: fibrocimento_amianto (2)
:                   Mean_pc6 > -17.0365:
:                   :...Border_ind > 4.709091: asfalto_pintado_branco (5/1)
:                       Border_ind <= 4.709091:
:                       :...modeMinimu.11 <= -70: vidro_aramado (3)
:                           modeMinimu.11 > -70:
:                           :...Mean_Layer.2 > 581.3438: aco_brilhante (2/1)
:                               Mean_Layer.2 <= 581.3438:
:                               :...Border_ind <= 3.609375: asfalto (102/1)
:                                   Border_ind > 3.609375: [S2]
Mean_Layer.5 > 1164.12:
:...NDSI2 <= -0.420767:
    :...modeMinimu.1 > -279.6134:
    :   :...Density <= 0.773677: asfalto_pintado_vermelho (2)
    :   :   Density > 0.773677:
    :   :   :...NDSI1 <= -0.454906:
    :   :       :...Min_pixel_.3 <= 51.88723: solo_exposto (25/1)
    :   :       :   Min_pixel_.3 > 51.88723: ceramica (3)
    :   :       NDSI1 > -0.454906:
    :   :       :...NDBI <= 0.065007: bloquete (4/1)
    :   :           NDBI > 0.065007: concreto (9/1)
    :   modeMinimu.1 <= -279.6134:
    :   :...Min_pixel_.4 > -386.3284: argamassa_cimento_pintado_vermelho (10)
    :       Min_pixel_.4 <= -386.3284:
    :       :...WVBI <= -0.845891: ceramica (17)
    :           WVBI > -0.845891:
    :           :...WVNII <= 4.555916:
    :               :...Min_pixel_.8 <= -72575.66: ceramica (2)
    :               :   Min_pixel_.8 > -72575.66: asfalto_pintado_vermelho (38/1)
    :               WVNII > 4.555916:
    :               :...BRTpBoRE <= 0.928107: ceramica (5)
    :                   BRTpBoRE > 0.928107: concreto_pintado_vermelho (2)
```

```
NDSI2 > -0.420767:
:...Max_pixel_.5 > 7529:
    :...Min_pixel_.5 <= -851.1691: galvalume_c (52)
    :   Min_pixel_.5 > -851.1691:
    :   :...Shape_inde <= 1.474024:
    :       :...Max_pixel_.8 > 820.5172: acrilico (3)
    :       :   Max_pixel_.8 <= 820.5172:
    :       :   :...Max_pixel_.15 <= 12328: lona (8/1)
    :       :       Max_pixel_.15 > 12328: galvalume_c (7)
    :       Shape_inde > 1.474024:
    :       :...Mean_pc4 <= -137.3515:
    :           :...Max_pixel_.4 <= 5234.897: aco_brilhante (6/1)
    :           :   Max_pixel_.4 > 5234.897: galvalume_c (5)
    :           Mean_pc4 > -137.3515:
    :           :...BRTpBoR <= 1.670762: aco_brilhante (2/1)
    :               BRTpBoR > 1.670762: galvalume_a (48/1)
    Max_pixel_.5 <= 7529:
    :...Mean_pc7 > 148.3353:
        :...Min_pixel_.8 <= -74713.39:
        :   :...NDSI1 <= -0.23413: acrilico (3)
        :   :   NDSI1 > -0.23413: galvalume_b (35)
        :   Min_pixel_.8 > -74713.39:
        :   :...Min_pixel_.11 <= 2198:
        :       :...Mean_Layer.6 <= 3134.882: pastilha_esmaltada (4/1)
        :       :   Mean_Layer.6 > 3134.882: piscina_azulejo (2)
        :       Min_pixel_.11 > 2198:
        :       :...Max_pixel_.3 <= -67743.04: piscina_vinil (2/1)
        :           Max_pixel_.3 > -67743.04: concreto_pintado_amarelo (7)
        Mean_pc7 <= 148.3353:
        :...BRTpBoR <= 1.610843:
            :...BRTpBoRE <= 1.041442:
            :   :...Max_pixel_.10 <= 3.021088: concreto_pintado_amarelo (2)
            :   :   Max_pixel_.10 > 3.021088: quartzito_amarelo (14/1)
            :   BRTpBoRE > 1.041442:
            :   :...Mean_pc1 <= -67218.15:
            :       :...NDVI1 <= 0.076137:
            :       :   :...NDBI <= 0.032102: piscina_vinil (4)
            :       :   :   NDBI > 0.032102: aco_ferrugem (7)
            :       :   NDVI1 > 0.076137: [S3]
            :       Mean_pc1 > -67218.15:
            :       :...Mean_pc6 <= -71.90514: [S4]
            :           Mean_pc6 > -71.90514:
            :           :...Min_pixel_.6 <= -1619.788: argamassa_cimento (3)
            :               Min_pixel_.6 > -1619.788:
            :               :...WVBI <= -0.633588: argamassa_cimento (3/1)
            :                   WVBI > -0.633588:
            :                   :...Mean_pc7 > 75.31554: concreto (16)
            :                       Mean_pc7 <= 75.31554:
            :                       :...Max_pixel_.3 > -58740.05: [S5]
            :                           Max_pixel_.3 <= -58740.05: [S6]
            BRTpBoR > 1.610843:
            :...BRTmN2 > -581.5336:
                :...Mean_pc3 > -129.5505: plastico (4/1)
                :   Mean_pc3 <= -129.5505:
                :   :...NDSI2 > -0.130317:
                :       :...NDBI <= 0.056124: aco_fosco (39/1)
                :       :   NDBI > 0.056124: aco_brilhante (3)
```

```
:       NDSI2 <= -0.130317:
:       :...NDBI > 0.070355: aco_brilhante (2/1)
:           NDBI <= 0.070355: [S7]
BRTmN2 <= -581.5336:
:...Mean_pc5 > 181.6412:
    :...Mean_pc7 > -12.56791: galvalume_c (5/1)
    :   Mean_pc7 <= -12.56791:
    :   :...Asymmetry <= 0.384027: acrilico (2/1)
    :       Asymmetry > 0.384027: galvalume_a (9)
    Mean_pc5 <= 181.6412:
    :...BRTpBoRE <= 1.285574:
        :...Mean_pc8 <= -37.29316: acrilico (2/1)
        :   Mean_pc8 > -37.29316: [S8]
        BRTpBoRE > 1.285574:
        :...BRTmN2 <= -2557.323: [S9]
            BRTmN2 > -2557.323:
            :...Mean_pc5 <= -39.74308:
                :...Max_pixel_.4 <= 956.4114: concreto (2/1)
                :   Max_pixel_.4 > 956.4114: [S10]
                Mean_pc5 > -39.74308: [S11]

SubTree [S1]

Mean_Layer.4 <= 503.3668: fibrocimento_amianto (11)
Mean_Layer.4 > 503.3668: concreto (4/1)

SubTree [S2]

Mean_Layer.2 <= 150.7316: vidro_aramado (3)
Mean_Layer.2 > 150.7316: asfalto (7)

SubTree [S3]

Border_ind <= 1.833333: aco_brilhante (2)
Border_ind > 1.833333: concreto_impermeabilizado (2)

SubTree [S4]

modeMinimu.16 <= 427: concreto_impermeabilizado (3/2)
modeMinimu.16 > 427: aco_brilhante (3/2)

SubTree [S5]

Border_ind <= 2.913043: aco_brilhante (3)
Border_ind > 2.913043: asfalto_pintado_branco (2)

SubTree [S6]

Max_pixel_.10 <= 97.74387: asfalto_pintado_branco (23/1)
Max_pixel_.10 > 97.74387:
:...modeMinimu.14 <= 538: asfalto_pintado_branco (3)
    modeMinimu.14 > 538: concreto (7)

SubTree [S7]

Border_ind <= 1.341463: acrilico (3)
Border_ind > 1.341463: asfalto_pintado_branco (3)
```

```
SubTree [S8]

Density <= 1.266007: fibra_vidro (5)
Density > 1.266007: asfalto_pintado_branco (3/1)

SubTree [S9]

Max_pixel_.10 <= 143.6804: fibra_vidro (3)
Max_pixel_.10 > 143.6804: concreto_impermeabilizado (16/1)

SubTree [S10]

Min_pixel_.6 <= -1827.245: aco_brilhante (3/1)
Min_pixel_.6 > -1827.245:
:...Mean_pc7 <= 34.39928:
    :...Min_pixel_.5 <= -585.3093: aco_brilhante (2)
    :   Min_pixel_.5 > -585.3093: fibrocimento_crfs (15/1)
    Mean_pc7 > 34.39928:
    :...BRTpBoR <= 1.690354: concreto_impermeabilizado (10)
        BRTpBoR > 1.690354:
        :...Mean_pc6 <= -233.2562: aco_brilhante (2/1)
            Mean_pc6 > -233.2562: fibrocimento_crfs (5)

SubTree [S11]

Max_pixel_.7 > 528.3216: concreto_impermeabilizado (2)
Max_pixel_.7 <= 528.3216:
:...Rectangula > 0.939198: fibrocimento_crfs (5)
    Rectangula <= 0.939198:
    :...Mean_pc8 <= 22.06047: aco_brilhante (66/2)
        Mean_pc8 > 22.06047:
        :...Mean_pc4 <= -110.7652: aco_brilhante (3)
            Mean_pc4 > -110.7652:
            :...BRTpBoR <= 1.747087: asfalto_pintado_branco (4)
                BRTpBoR > 1.747087: fibrocimento_crfs (4)
```

B.7 Árvore gerada por CART Nível I com MDA

```
CART Decision Tree - (Nível I com MDA)

Max_pixel_.2 < 19.368837499999998
|  Min_pixel_.2 < 2.599057
|  |  Max_pixel_.12 < 365.5
|  |  |  BRT < -27.004783500000002: sombra(6.0/0.0)
|  |  |  BRT >= -27.004783500000002: solo_exposto(3.0/2.0)
|  |  Max_pixel_.12 >= 365.5: pavimentacao_viaria(10.0/0.0)
|  Min_pixel_.2 >= 2.599057: lago(442.0/0.0)
Max_pixel_.2 >= 19.368837499999998
|  mSWIR < -54.346544
|  |  Min_pixel_.12 < -228.0: sombra(195.0/9.0)
|  |  Min_pixel_.12 >= -228.0: pavimentacao_viaria(3.0/2.0)
|  mSWIR >= -54.346544
|  |  modeMinimu.14 < 1269.5
|  |  |  NDSI1 < -0.9061170000000001
```

```
|  |  |  |  modeMinimu.10 < -447.5: sombra(7.0/3.0)
|  |  |  |  modeMinimu.10 >= -447.5
|  |  |  |  |  Min_pixel_.14 < -248.0: vegetacao_arborea(136.0/6.0)
|  |  |  |  |  Min_pixel_.14 >= -248.0: vegetacao_rasteira(4.0/0.0)
|  |  |  NDSI1 >= -0.9061170000000001
|  |  |  |  modeMinimu.8 < 25.824343
|  |  |  |  |  NDVI2 < 0.09522900000000001
|  |  |  |  |  |  Mean_Layer.2 < 26.407101500000003
|  |  |  |  |  |  |  Mean_pc3 < -87.36590050000001
|  |  |  |  |  |  |  |  Mean_mda < 23.610619: pavimentacao_nao_viaria(7.0/0.0)
|  |  |  |  |  |  |  |  Mean_mda >= 23.610619: ceramica(4.0/3.0)
|  |  |  |  |  |  |  Mean_pc3 >= -87.36590050000001: sombra(8.0/0.0)
|  |  |  |  |  |  Mean_Layer.2 >= 26.407101500000003: pavimentacao_viaria(144.0/19.0)
|  |  |  |  |  NDVI2 >= 0.09522900000000001
|  |  |  |  |  |  NDVI1 < 0.4090905
|  |  |  |  |  |  |  BRTpBoRE < 0.835503
|  |  |  |  |  |  |  |  modeMinimu.1 < -279.619812
|  |  |  |  |  |  |  |  |  Mean_pc7 < -329.409865: pavimentacao_viaria(6.0/1.0)
|  |  |  |  |  |  |  |  |  Mean_pc7 >= -329.409865: ceramica(12.0/1.0)
|  |  |  |  |  |  |  |  modeMinimu.1 >= -279.619812
|  |  |  |  |  |  |  |  |  Mean_Layer < 1636.8597045000001
|  |  |  |  |  |  |  |  |  |  WVSI < -0.303513
|  |  |  |  |  |  |  |  |  |  |  BRTmN2 < -776.0641865: materiais_mistos(4.0/0.0)
|  |  |  |  |  |  |  |  |  |  |  BRTmN2 >= -776.0641865: pavimentacao_viaria(3.0/3.0)
|  |  |  |  |  |  |  |  |  |  WVSI >= -0.303513: vegetacao_rasteira(5.0/1.0)
|  |  |  |  |  |  |  |  |  Mean_Layer >= 1636.8597045000001: solo_exposto(23.0/3.0)
|  |  |  |  |  |  |  BRTpBoRE >= 0.835503
|  |  |  |  |  |  |  |  modeMinimu.5 < -212.7274625
|  |  |  |  |  |  |  |  |  Max_pixel_.8 < 788.7038574999999
|  |  |  |  |  |  |  |  |  |  Max_pixel_.16 < -995.8473509999999: pavimentacao_nao_viaria(7.0/0.0)
|  |  |  |  |  |  |  |  |  |  Max_pixel_.16 >= -995.8473509999999
|  |  |  |  |  |  |  |  |  |  |  Max_pixel_.2 < 23.45627
|  |  |  |  |  |  |  |  |  |  |  |  Min_pixel_.6 < -858.5167845: pavimentacao_viaria(10.0/2.0)
|  |  |  |  |  |  |  |  |  |  |  |  Min_pixel_.6 >= -858.5167845
|  |  |  |  |  |  |  |  |  |  |  |  |  Mean_pc2 < 712.790157: materiais_mistos(5.0/2.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  Mean_pc2 >= 712.790157: solo_exposto(3.0/0.0)
|  |  |  |  |  |  |  |  |  |  |  Max_pixel_.2 >= 23.45627
|  |  |  |  |  |  |  |  |  |  |  |  BRT < 466.6711275: coberturas_diversas(4.0/0.0)
|  |  |  |  |  |  |  |  |  |  |  |  BRT >= 466.6711275: materiais_mistos(59.0/11.0)
|  |  |  |  |  |  |  |  |  Max_pixel_.8 >= 788.7038574999999
|  |  |  |  |  |  |  |  |  |  modeMinimu.10 < 3263.5: pavimentacao_viaria(14.0/1.0)
|  |  |  |  |  |  |  |  |  |  modeMinimu.10 >= 3263.5: ceramica(3.0/3.0)
|  |  |  |  |  |  |  |  modeMinimu.5 >= -212.7274625: coberturas_diversas(14.0/4.0)
|  |  |  |  |  |  NDVI1 >= 0.4090905: vegetacao_rasteira(36.0/5.0)
|  |  |  |  modeMinimu.8 >= 25.824343
|  |  |  |  |  Mean_pc2 < -256.8398555
|  |  |  |  |  |  Mean_pc3 < -338.6283365: coberturas_metalicas(7.0/1.0)
|  |  |  |  |  |  Mean_pc3 >= -338.6283365
|  |  |  |  |  |  |  Min_pixel_.3 < -699.5814515: materiais_mistos(5.0/0.0)
|  |  |  |  |  |  |  Min_pixel_.3 >= -699.5814515
|  |  |  |  |  |  |  |  BRTmN2 < 150.93595349999998: coberturas_diversas(83.0/7.0)
|  |  |  |  |  |  |  |  BRTmN2 >= 150.93595349999998: coberturas_metalicas(4.0/2.0)
|  |  |  |  |  Mean_pc2 >= -256.8398555
|  |  |  |  |  |  WVSI < -0.17975750000000001
|  |  |  |  |  |  |  Max_pixel_.2 < 30.565309499999998: ceramica(9.0/0.0)
|  |  |  |  |  |  |  Max_pixel_.2 >= 30.565309499999998: solo_exposto(5.0/3.0)
|  |  |  |  |  |  WVSI >= -0.17975750000000001
```

```
|  |  |  |  |  |  |  Mean_pc4 < -24.033853999999998
|  |  |  |  |  |  |  |  Min_pixel_.5 < -443.820511: materiais_mistos(7.0/5.0)
|  |  |  |  |  |  |  |  Min_pixel_.5 >= -443.820511: coberturas_metalicas(11.0/2.0)
|  |  |  |  |  |  |  Mean_pc4 >= -24.033853999999998: coberturas_diversas(14.0/5.0)
|  |  modeMinimu.14 >= 1269.5
|  |  |  Max_pixel_.2 < 26.199709
|  |  |  |  NDSI2 < -0.177992
|  |  |  |  |  NDSI1 < -0.23997049999999998
|  |  |  |  |  |  BRT < 5089.2547555
|  |  |  |  |  |  |  BRT < 3733.517137: pavimentacao_viaria(3.0/0.0)
|  |  |  |  |  |  |  BRT >= 3733.517137: pavimentacao_nao_viaria(3.0/0.0)
|  |  |  |  |  |  BRT >= 5089.2547555: coberturas_diversas(4.0/0.0)
|  |  |  |  |  NDSI1 >= -0.23997049999999998: materiais_mistos(28.0/1.0)
|  |  |  |  NDSI2 >= -0.177992
|  |  |  |  |  Max_pixel_.8 < 77.624401: piscina(9.0/1.0)
|  |  |  |  |  Max_pixel_.8 >= 77.624401: coberturas_metalicas(10.0/1.0)
|  |  |  Max_pixel_.2 >= 26.199709
|  |  |  |  Compactnes < 1.2964105
|  |  |  |  |  modeMinimu.16 < 4404.0: coberturas_diversas(23.0/4.0)
|  |  |  |  |  modeMinimu.16 >= 4404.0: coberturas_metalicas(7.0/0.0)
|  |  |  |  Compactnes >= 1.2964105: coberturas_metalicas(256.0/18.0)

Number of Leaf Nodes: 47

Size of the Tree: 93
```

B.8 Árvore gerada por CART Nível I sem MDA

```
CART Decision Tree - (Nível I sem MDA)

Max_pixel_.4 < -1616.9703370000002
|  NDSI1 < -0.7033435: lago(440.0/0.0)
|  NDSI1 >= -0.7033435
|  |  BRT < -115.054487: sombra(13.0/0.0)
|  |  BRT >= -115.054487: piscina(4.0/3.0)
Max_pixel_.4 >= -1616.9703370000002
|  modeMinimu.14 < 1269.5
|  |  mSWIR < -54.346544
|  |  |  Min_pixel_.13 < -166.5: sombra(188.0/9.0)
|  |  |  Min_pixel_.13 >= -166.5: pavimentacao_viaria(3.0/3.0)
|  |  mSWIR >= -54.346544
|  |  |  Mean_Layer.5 < -216.93235299999998
|  |  |  |  modeMinimu.9 < -422.5: sombra(6.0/0.0)
|  |  |  |  modeMinimu.9 >= -422.5: vegetacao_arborea(132.0/4.0)
|  |  |  Mean_Layer.5 >= -216.93235299999998
|  |  |  |  NDVI2 < 0.09471450000000001
|  |  |  |  |  modeMinimu.16 < 45.5
|  |  |  |  |  |  NDSI2 < -0.45760500000000004: ceramica(8.0/2.0)
|  |  |  |  |  |  NDSI2 >= -0.45760500000000004
|  |  |  |  |  |  |  WVSI < -0.1937235
|  |  |  |  |  |  |  |  Max_pixel_.9 < 75.227745: pavimentacao_nao_viaria(8.0/0.0)
|  |  |  |  |  |  |  |  Max_pixel_.9 >= 75.227745: materiais_mistos(3.0/2.0)
|  |  |  |  |  |  |  WVSI >= -0.1937235
|  |  |  |  |  |  |  |  modeMinimu.15 < 267.0
|  |  |  |  |  |  |  |  |  Max_pixel_.10 < 100.076168: sombra(3.0/2.0)
```

```
|  |  |  |  |  |  |  |  |  |  Max_pixel_.10 >= 100.076168: coberturas_diversas(7.0/4.0)
|  |  |  |  |  |  |  |  |  modeMinimu.15 >= 267.0: pavimentacao_viaria(6.0/2.0)
|  |  |  |  |  modeMinimu.16 >= 45.5
|  |  |  |  |  |  Mean_pc6 < -9.5304105
|  |  |  |  |  |  |  Min_pixel_.6 < -363.594696
|  |  |  |  |  |  |  |  WVBI < -0.366224: pavimentacao_viaria(6.0/1.0)
|  |  |  |  |  |  |  |  WVBI >= -0.366224: coberturas_metalicas(7.0/2.0)
|  |  |  |  |  |  |  Min_pixel_.6 >= -363.594696: coberturas_diversas(5.0/0.0)
|  |  |  |  |  |  Mean_pc6 >= -9.5304105
|  |  |  |  |  |  |  Mean_Layer.3 < 1568.5348835
|  |  |  |  |  |  |  |  modeMinimu.12 < 212.5
|  |  |  |  |  |  |  |  |  BRTpBoR < 1.332319: coberturas_diversas(4.0/0.0)
|  |  |  |  |  |  |  |  |  BRTpBoR >= 1.332319: pavimentacao_viaria(3.0/3.0)
|  |  |  |  |  |  |  |  modeMinimu.12 >= 212.5
|  |  |  |  |  |  |  |  |  Density < 1.7295345: pavimentacao_viaria(124.0/3.0)
|  |  |  |  |  |  |  |  |  Density >= 1.7295345
|  |  |  |  |  |  |  |  |  |  BRT < 757.1012754999999: pavimentacao_viaria(8.0/0.0)
|  |  |  |  |  |  |  |  |  |  BRT >= 757.1012754999999: materiais_mistos(6.0/0.0)
|  |  |  |  |  |  |  Mean_Layer.3 >= 1568.5348835: materiais_mistos(3.0/1.0)
|  |  |  |  NDVI2 >= 0.09471450000000001
|  |  |  |  |  BRTpBoRE < 1.3490795000000002
|  |  |  |  |  |  NDVI1 < 0.40652699999999997
|  |  |  |  |  |  |  WVBI < -0.8272195
|  |  |  |  |  |  |  |  Mean_pc7 < -172.252794: ceramica(14.0/0.0)
|  |  |  |  |  |  |  |  Mean_pc7 >= -172.252794
|  |  |  |  |  |  |  |  |  Mean_pc2 < 741.781919
|  |  |  |  |  |  |  |  |  |  BRTmN2 < -752.4249265000001: materiais_mistos(11.0/4.0)
|  |  |  |  |  |  |  |  |  |  BRTmN2 >= -752.4249265000001: ceramica(3.0/10.0)
|  |  |  |  |  |  |  |  |  Mean_pc2 >= 741.781919: solo_exposto(32.0/3.0)
|  |  |  |  |  |  |  WVBI >= -0.8272195
|  |  |  |  |  |  |  |  WVSI < -0.2995425
|  |  |  |  |  |  |  |  |  Max_pixel_.1 < 6211.5: pavimentacao_viaria(19.0/4.0)
|  |  |  |  |  |  |  |  |  Max_pixel_.1 >= 6211.5: ceramica(4.0/1.0)
|  |  |  |  |  |  |  |  WVSI >= -0.2995425
|  |  |  |  |  |  |  |  |  Max_pixel_.16 < -1031.595886: pavimentacao_nao_viaria(7.0/1.0)
|  |  |  |  |  |  |  |  |  Max_pixel_.16 >= -1031.595886
|  |  |  |  |  |  |  |  |  |  Max_pixel_.7 < 138.639046
|  |  |  |  |  |  |  |  |  |  |  Mean_pc4 < 135.26042: materiais_mistos(47.0/16.0)
|  |  |  |  |  |  |  |  |  |  |  Mean_pc4 >= 135.26042: pavimentacao_viaria(3.0/0.0)
|  |  |  |  |  |  |  |  |  |  Max_pixel_.7 >= 138.639046
|  |  |  |  |  |  |  |  |  |  |  NDVI2 < 0.167613
|  |  |  |  |  |  |  |  |  |  |  |  Max_pixel_.9 < 222.129425: pavimentacao_viaria(9.0/4.0)
|  |  |  |  |  |  |  |  |  |  |  |  Max_pixel_.9 >= 222.129425: materiais_mistos(8.0/1.0)
|  |  |  |  |  |  |  |  |  |  |  NDVI2 >= 0.167613
|  |  |  |  |  |  |  |  |  |  |  |  modeMinimu.11 < 1461.5: coberturas_diversas(18.0/5.0)
|  |  |  |  |  |  |  |  |  |  |  |  modeMinimu.11 >= 1461.5: materiais_mistos(5.0/6.0)
|  |  |  |  |  |  NDVI1 >= 0.40652699999999997
|  |  |  |  |  |  |  modeMinimu.15 < -4.5
|  |  |  |  |  |  |  |  Min_pixel_.16 < -92.5: sombra(5.0/4.0)
|  |  |  |  |  |  |  |  Min_pixel_.16 >= -92.5: vegetacao_arborea(7.0/0.0)
|  |  |  |  |  |  |  modeMinimu.15 >= -4.5
|  |  |  |  |  |  |  |  Max_pixel_.1 < 3632.5: vegetacao_rasteira(40.0/3.0)
|  |  |  |  |  |  |  |  Max_pixel_.1 >= 3632.5: solo_exposto(3.0/1.0)
|  |  |  |  |  BRTpBoRE >= 1.3490795000000002
|  |  |  |  |  |  Min_pixel_.3 < -699.5814515: materiais_mistos(5.0/1.0)
|  |  |  |  |  |  Min_pixel_.3 >= -699.5814515
|  |  |  |  |  |  |  Mean_pc3 < -340.329833
```

```
|  |  |  |  |  |  |  |  BRT < 2669.4978815: coberturas_metalicas(6.0/0.0)
|  |  |  |  |  |  |  |  BRT >= 2669.4978815: coberturas_diversas(3.0/1.0)
|  |  |  |  |  |  |  Mean_pc3 >= -340.329833
|  |  |  |  |  |  |  |  modeMinimu.12 < 42.5: sombra(3.0/0.0)
|  |  |  |  |  |  |  |  modeMinimu.12 >= 42.5: coberturas_diversas(79.0/6.0)
|  modeMinimu.14 >= 1269.5
|  |  Max_pixel_.7 < 178.843597
|  |  |  Max_pixel_.3 < -71698.28125: coberturas_metalicas(19.0/3.0)
|  |  |  Max_pixel_.3 >= -71698.28125
|  |  |  |  Border_ind < 1.2650375
|  |  |  |  |  Max_pixel_.16 < -728.0300904999999
|  |  |  |  |  |  BRT < 4933.5516824999995: piscina(5.0/3.0)
|  |  |  |  |  |  BRT >= 4933.5516824999995: coberturas_diversas(3.0/0.0)
|  |  |  |  |  Max_pixel_.16 >= -728.0300904999999: coberturas_diversas(19.0/0.0)
|  |  |  |  Border_ind >= 1.2650375
|  |  |  |  |  Mean_pc4 < -132.0655355: coberturas_metalicas(6.0/0.0)
|  |  |  |  |  Mean_pc4 >= -132.0655355
|  |  |  |  |  |  Mean_pc7 < 15.712261: coberturas_diversas(3.0/1.0)
|  |  |  |  |  |  Mean_pc7 >= 15.712261: materiais_mistos(13.0/4.0)
|  |  Max_pixel_.7 >= 178.843597
|  |  |  Mean_pc6 < -448.1860585
|  |  |  |  BRTmN1 < -85.8671875: materiais_mistos(9.0/2.0)
|  |  |  |  BRTmN1 >= -85.8671875: piscina(3.0/0.0)
|  |  |  Mean_pc6 >= -448.1860585
|  |  |  |  Border_ind < 1.0662395
|  |  |  |  |  Max_pixel_.4 < 7495.0058595: coberturas_diversas(9.0/3.0)
|  |  |  |  |  Max_pixel_.4 >= 7495.0058595: coberturas_metalicas(6.0/0.0)
|  |  |  |  Border_ind >= 1.0662395: coberturas_metalicas(240.0/15.0)

Number of Leaf Nodes: 56

Size of the Tree: 111
```

B.9 Árvore gerada por C4.5 Nível I com MDA

```
J48 pruned tree - (Nível I com MDA)
------------------

Max_pixel_.4 <= -1591.778076
|   NDVI1 <= 0.511419
|   |   Max_pixel_.1 <= -121: sombra (16.0)
|   |   Max_pixel_.1 > -121
|   |   |   Border_ind <= 1.823529: coberturas_diversas (3.0/1.0)
|   |   |   Border_ind > 1.823529: piscina (5.0)
|   NDVI1 > 0.511419: lago (442.0)
Max_pixel_.4 > -1591.778076
|   Mean_Layer.4 <= 124.285714
|   |   BRTmN1 <= -325.236842
|   |   |   Min_pixel_.14 <= -259
|   |   |   |   WVSI <= -0.107006
|   |   |   |   |   Mean_Layer.2 <= -332.681818: vegetacao_arborea (120.0)
|   |   |   |   |   Mean_Layer.2 > -332.681818
|   |   |   |   |   |   Asymmetry <= 0.785621: vegetacao_arborea (15.0/1.0)
|   |   |   |   |   |   Asymmetry > 0.785621: vegetacao_rasteira (2.0)
|   |   |   |   WVSI > -0.107006
```

```
|   |   |   |   |   Min_pixel_ <= -1009: sombra (4.0)
|   |   |   |   |   Min_pixel_ > -1009
|   |   |   |   |   |   BRTpBoR <= 5.336353: vegetacao_arborea (6.0)
|   |   |   |   |   |   BRTpBoR > 5.336353: vegetacao_rasteira (2.0)
|   |   |   Min_pixel_.14 > -259: vegetacao_rasteira (10.0)
|   |   BRTmN1 > -325.236842
|   |   |   modeMinimu.9 <= -60
|   |   |   |   Max_pixel_.7 <= 620.852356
|   |   |   |   |   BRTmN1 <= 59.86284
|   |   |   |   |   |   Min_pixel_.8 <= -60460.632813
|   |   |   |   |   |   |   Mean_Layer.6 <= -50.871429
|   |   |   |   |   |   |   |   Max_pixel_.4 <= 413.819702
|   |   |   |   |   |   |   |   |   Max_pixel_.2 <= 26.899782: sombra (3.0)
|   |   |   |   |   |   |   |   |   Max_pixel_.2 > 26.899782: vegetacao_arborea (2.0)
|   |   |   |   |   |   |   |   Max_pixel_.4 > 413.819702: sombra (24.0)
|   |   |   |   |   |   |   Mean_Layer.6 > -50.871429: pavimentacao_viaria (2.0/1.0)
|   |   |   |   |   |   Min_pixel_.8 > -60460.632813: vegetacao_arborea (3.0)
|   |   |   |   |   BRTmN1 > 59.86284
|   |   |   |   |   |   Mean_pc4 <= -70.999516
|   |   |   |   |   |   |   Mean_pc4 <= -81.130268: sombra (9.0)
|   |   |   |   |   |   |   Mean_pc4 > -81.130268: materiais_mistos (2.0)
|   |   |   |   |   |   Mean_pc4 > -70.999516: sombra (157.0)
|   |   |   |   Max_pixel_.7 > 620.852356: coberturas_diversas (3.0/1.0)
|   |   |   modeMinimu.9 > -60
|   |   |   |   modeMinimu.16 <= -279: sombra (2.0/1.0)
|   |   |   |   modeMinimu.16 > -279: coberturas_diversas (4.0)
|   Mean_Layer.4 > 124.285714
|   |   BRTpBoRE <= 1.044126
|   |   |   NDVI3 <= 0.352876
|   |   |   |   NDSI2 <= -0.423025
|   |   |   |   |   Mean_pc8 <= 72.822729
|   |   |   |   |   |   WVSI <= -0.397283
|   |   |   |   |   |   |   Asymmetry <= 0.839338
|   |   |   |   |   |   |   |   Max_pixel_.2 <= 21.466433: sombra (2.0/1.0)
|   |   |   |   |   |   |   |   Max_pixel_.2 > 21.466433: ceramica (19.0)
|   |   |   |   |   |   |   Asymmetry > 0.839338: pavimentacao_viaria (3.0)
|   |   |   |   |   |   WVSI > -0.397283
|   |   |   |   |   |   |   BRTmN1 <= -679.096591
|   |   |   |   |   |   |   |   Max_pixel_.5 <= 2780
|   |   |   |   |   |   |   |   |   Mean_Layer.2 <= 196.5: solo_exposto (21.0)
|   |   |   |   |   |   |   |   |   Mean_Layer.2 > 196.5: pavimentacao_viaria (4.0/1.0)
|   |   |   |   |   |   |   |   Max_pixel_.5 > 2780
|   |   |   |   |   |   |   |   |   BRTpBoRE <= 0.961902: ceramica (14.0/2.0)
|   |   |   |   |   |   |   |   |   BRTpBoRE > 0.961902: materiais_mistos (2.0)
|   |   |   |   |   |   |   BRTmN1 > -679.096591
|   |   |   |   |   |   |   |   NDBI <= 0.047362: ceramica (2.0/1.0)
|   |   |   |   |   |   |   |   NDBI > 0.047362: materiais_mistos (25.0/1.0)
|   |   |   |   |   Mean_pc8 > 72.822729
|   |   |   |   |   |   WVNII <= 4.555916: pavimentacao_viaria (36.0/1.0)
|   |   |   |   |   |   WVNII > 4.555916: ceramica (5.0/1.0)
|   |   |   |   NDSI2 > -0.423025
|   |   |   |   |   NDVI1 <= 0.102211
|   |   |   |   |   |   Max_pixel_.7 <= 127.786697
|   |   |   |   |   |   |   Density <= 0.810602: vegetacao_rasteira (2.0/1.0)
|   |   |   |   |   |   |   Density > 0.810602
|   |   |   |   |   |   |   |   Compactnes <= 1.391695: materiais_mistos (2.0)
|   |   |   |   |   |   |   |   Compactnes > 1.391695: pavimentacao_nao_viaria (22.0/2.0)
```

```
|   |   |   |   |   |   Max_pixel_.7 > 127.786697: pavimentacao_viaria (3.0/1.0)
|   |   |   |   |   NDVI1 > 0.102211: materiais_mistos (9.0)
|   |   |   NDVI3 > 0.352876
|   |   |   |   mSWIR <= 1882.697874
|   |   |   |   |   modeMinimu.16 <= -270
|   |   |   |   |   |   Asymmetry <= 0.561596: sombra (2.0)
|   |   |   |   |   |   Asymmetry > 0.561596: materiais_mistos (2.0)
|   |   |   |   |   modeMinimu.16 > -270
|   |   |   |   |   |   Min_pixel_.11 <= -457: coberturas_diversas (2.0)
|   |   |   |   |   |   Min_pixel_.11 > -457
|   |   |   |   |   |   |   NDVI1 <= 0.38647
|   |   |   |   |   |   |   |   Min_pixel_ <= -8: solo_exposto (3.0/1.0)
|   |   |   |   |   |   |   |   Min_pixel_ > -8: vegetacao_rasteira (3.0)
|   |   |   |   |   |   |   NDVI1 > 0.38647
|   |   |   |   |   |   |   |   Max_pixel_.9 <= 39.003502: vegetacao_arborea (3.0/1.0)
|   |   |   |   |   |   |   |   Max_pixel_.9 > 39.003502: vegetacao_rasteira (30.0)
|   |   |   |   mSWIR > 1882.697874: solo_exposto (12.0)
|   |   BRTpBoRE > 1.044126
|   |   |   Mean_mda <= 25.566669
|   |   |   |   modeMinimu.16 <= 1198
|   |   |   |   |   NDVI2 <= 0.078766
|   |   |   |   |   |   Mean_Layer.5 <= 1705.22291
|   |   |   |   |   |   |   WVBI <= -0.711179
|   |   |   |   |   |   |   |   BRTmN1 <= 40.566542: pavimentacao_viaria (2.0)
|   |   |   |   |   |   |   |   BRTmN1 > 40.566542: pavimentacao_nao_viaria (3.0/1.0)
|   |   |   |   |   |   |   WVBI > -0.711179
|   |   |   |   |   |   |   |   Mean_pc4 <= -54.405567
|   |   |   |   |   |   |   |   |   Border_ind <= 1.653846: pavimentacao_viaria (8.0)
|   |   |   |   |   |   |   |   |   Border_ind > 1.653846: coberturas_diversas (2.0)
|   |   |   |   |   |   |   |   Mean_pc4 > -54.405567: pavimentacao_viaria (111.0)
|   |   |   |   |   |   Mean_Layer.5 > 1705.22291: coberturas_metalicas (2.0/1.0)
|   |   |   |   |   NDVI2 > 0.078766
|   |   |   |   |   |   Mean_pc3 <= -162.943146
|   |   |   |   |   |   |   Mean_pc7 <= 30.195955
|   |   |   |   |   |   |   |   Max_pixel_.8 <= 134.329102: materiais_mistos (3.0)
|   |   |   |   |   |   |   |   Max_pixel_.8 > 134.329102
|   |   |   |   |   |   |   |   |   Density <= 0.944094: coberturas_metalicas (3.0/1.0)
|   |   |   |   |   |   |   |   |   Density > 0.944094: pavimentacao_viaria (26.0/2.0)
|   |   |   |   |   |   |   Mean_pc7 > 30.195955
|   |   |   |   |   |   |   |   Max_pixel_.2 <= 23.721865: pavimentacao_viaria (4.0/1.0)
|   |   |   |   |   |   |   |   Max_pixel_.2 > 23.721865: materiais_mistos (32.0)
|   |   |   |   |   |   Mean_pc3 > -162.943146
|   |   |   |   |   |   |   Min_pixel_.4 <= -51.941666
|   |   |   |   |   |   |   |   Mean_pc8 <= -54.655744: materiais_mistos (7.0)
|   |   |   |   |   |   |   |   Mean_pc8 > -54.655744
|   |   |   |   |   |   |   |   |   Border_ind <= 3.066667: coberturas_diversas (17.0)
|   |   |   |   |   |   |   |   |   Border_ind > 3.066667: materiais_mistos (2.0)
|   |   |   |   |   |   |   Min_pixel_.4 > -51.941666
|   |   |   |   |   |   |   |   Asymmetry <= 0.63352: pavimentacao_viaria (2.0)
|   |   |   |   |   |   |   |   Asymmetry > 0.63352: sombra (2.0)
|   |   |   |   modeMinimu.16 > 1198
|   |   |   |   |   NDSI1 <= -0.165599
|   |   |   |   |   |   Border_ind <= 1.1
|   |   |   |   |   |   |   Compactnes <= 1.02439: coberturas_diversas (4.0)
|   |   |   |   |   |   |   Compactnes > 1.02439: pavimentacao_viaria (2.0/1.0)
|   |   |   |   |   |   Border_ind > 1.1
|   |   |   |   |   |   |   Mean_pc6 <= 58.148927: materiais_mistos (23.0)
```

```
|   |   |   |   |   |   |   Mean_pc6 > 58.148927: pavimentacao_nao_viaria (2.0/1.0)
|   |   |   |   |   NDSI1 > -0.165599
|   |   |   |   |   |   Max_pixel_.8 <= 389.33075
|   |   |   |   |   |   |   Roundness <= 0.044241: piscina (4.0)
|   |   |   |   |   |   |   Roundness > 0.044241
|   |   |   |   |   |   |   |   BRTmN1 <= 345.590164: coberturas_metalicas (20.0)
|   |   |   |   |   |   |   |   BRTmN1 > 345.590164: piscina (2.0)
|   |   |   |   |   |   Max_pixel_.8 > 389.33075: coberturas_diversas (2.0/1.0)
|   |   |   Mean_mda > 25.566669
|   |   |   |   modeMinimu.15 <= 985
|   |   |   |   |   Max_pixel_.8 <= 625.277893
|   |   |   |   |   |   BRTmN2 <= 150.913424
|   |   |   |   |   |   |   Mean_mda <= 30.234095
|   |   |   |   |   |   |   |   Mean_pc3 <= -109.651305
|   |   |   |   |   |   |   |   |   Min_pixel_.9 <= 25.229982
|   |   |   |   |   |   |   |   |   |   Asymmetry <= 0.662958: materiais_mistos (2.0)
|   |   |   |   |   |   |   |   |   |   Asymmetry > 0.662958: pavimentacao_viaria (2.0)
|   |   |   |   |   |   |   |   |   Min_pixel_.9 > 25.229982: coberturas_metalicas (9.0)
|   |   |   |   |   |   |   |   Mean_pc3 > -109.651305
|   |   |   |   |   |   |   |   |   Mean_mda <= 29.631647
|   |   |   |   |   |   |   |   |   |   Min_pixel_.3 <= -180.824081: coberturas_diversas (21.0)
|   |   |   |   |   |   |   |   |   |   Min_pixel_.3 > -180.824081: coberturas_metalicas (2.0/1.0)
|   |   |   |   |   |   |   |   |   Mean_mda > 29.631647: coberturas_metalicas (2.0)
|   |   |   |   |   |   |   Mean_mda > 30.234095
|   |   |   |   |   |   |   |   modeMinimu.11 <= -560: materiais_mistos (3.0)
|   |   |   |   |   |   |   |   modeMinimu.11 > -560
|   |   |   |   |   |   |   |   |   modeMinimu.11 <= 1051: coberturas_diversas (64.0/1.0)
|   |   |   |   |   |   |   |   |   modeMinimu.11 > 1051
|   |   |   |   |   |   |   |   |   |   Max_pixel_.2 <= 45.279846: coberturas_diversas (4.0)
|   |   |   |   |   |   |   |   |   |   Max_pixel_.2 > 45.279846: materiais_mistos (3.0)
|   |   |   |   |   |   BRTmN2 > 150.913424
|   |   |   |   |   |   |   Asymmetry <= 0.956603: pavimentacao_viaria (3.0)
|   |   |   |   |   |   |   Asymmetry > 0.956603: coberturas_metalicas (4.0/1.0)
|   |   |   |   |   Max_pixel_.8 > 625.277893
|   |   |   |   |   |   Min_pixel_.3 <= -574.601868: materiais_mistos (8.0)
|   |   |   |   |   |   Min_pixel_.3 > -574.601868
|   |   |   |   |   |   |   Border_ind <= 1.821429: sombra (2.0)
|   |   |   |   |   |   |   Border_ind > 1.821429: coberturas_diversas (3.0/1.0)
|   |   |   |   modeMinimu.15 > 985
|   |   |   |   |   NDVI1 <= 0.197732
|   |   |   |   |   |   Compactnes <= 1.285714
|   |   |   |   |   |   |   Max_pixel_.2 <= 33.16687
|   |   |   |   |   |   |   |   Max_pixel_.10 <= 132.281052: coberturas_diversas (11.0)
|   |   |   |   |   |   |   |   Max_pixel_.10 > 132.281052: coberturas_metalicas (4.0/1.0)
|   |   |   |   |   |   |   Max_pixel_.2 > 33.16687: coberturas_metalicas (7.0)
|   |   |   |   |   |   Compactnes > 1.285714
|   |   |   |   |   |   |   Mean_pc3 <= -1271.186967: coberturas_metalicas (164.0)
|   |   |   |   |   |   |   Mean_pc3 > -1271.186967
|   |   |   |   |   |   |   |   NDVI1 <= 0.094796
|   |   |   |   |   |   |   |   |   Shape_inde <= 1.347151: coberturas_diversas (3.0/1.0)
|   |   |   |   |   |   |   |   |   Shape_inde > 1.347151: coberturas_metalicas (87.0/2.0)
|   |   |   |   |   |   |   |   NDVI1 > 0.094796
|   |   |   |   |   |   |   |   |   Max_pixel_.16 <= 171.374435
|   |   |   |   |   |   |   |   |   |   NDSI1 <= -0.166757: coberturas_diversas (17.0)
|   |   |   |   |   |   |   |   |   |   NDSI1 > -0.166757
|   |   |   |   |   |   |   |   |   |   |   Asymmetry <= 0.467505: coberturas_diversas (2.0)
|   |   |   |   |   |   |   |   |   |   |   Asymmetry > 0.467505: coberturas_metalicas (5.0)
```

```
|   |   |   |   |   |   |   |   Max_pixel_.16 > 171.374435: coberturas_metalicas (3.0)
|   |   |   |   |   NDVI1 > 0.197732: coberturas_diversas (12.0)

Number of Leaves  :  96

Size of the tree :  191
```

B.10 Árvore gerada por C4.5 Nível I sem MDA

```
J48 pruned tree - (Nível I sem MDA)
------------------

Max_pixel_.4 <= -1591.778076
|   NDVI1 <= 0.511419
|   |   Max_pixel_.1 <= -121: sombra (16.0)
|   |   Max_pixel_.1 > -121
|   |   |   Border_ind <= 1.823529: coberturas_diversas (3.0/1.0)
|   |   |   Border_ind > 1.823529: piscina (5.0)
|   NDVI1 > 0.511419: lago (442.0)
Max_pixel_.4 > -1591.778076
|   Mean_Layer.4 <= 124.285714
|   |   BRTmN1 <= -325.236842
|   |   |   Min_pixel_.14 <= -259
|   |   |   |   WVSI <= -0.107006
|   |   |   |   |   Mean_Layer.2 <= -332.681818: vegetacao_arborea (120.0)
|   |   |   |   |   Mean_Layer.2 > -332.681818
|   |   |   |   |   |   Asymmetry <= 0.785621: vegetacao_arborea (15.0/1.0)
|   |   |   |   |   |   Asymmetry > 0.785621: vegetacao_rasteira (2.0)
|   |   |   |   WVSI > -0.107006
|   |   |   |   |   Min_pixel_ <= -1009: sombra (4.0)
|   |   |   |   |   Min_pixel_ > -1009
|   |   |   |   |   |   BRTpBoR <= 5.336353: vegetacao_arborea (6.0)
|   |   |   |   |   |   BRTpBoR > 5.336353: vegetacao_rasteira (2.0)
|   |   |   Min_pixel_.14 > -259: vegetacao_rasteira (10.0)
|   |   BRTmN1 > -325.236842
|   |   |   modeMinimu.9 <= -60
|   |   |   |   Max_pixel_.7 <= 620.852356
|   |   |   |   |   BRTmN1 <= 59.86284
|   |   |   |   |   |   Min_pixel_.8 <= -60460.632813
|   |   |   |   |   |   |   Mean_Layer.6 <= -50.871429
|   |   |   |   |   |   |   |   Max_pixel_.4 <= 413.819702
|   |   |   |   |   |   |   |   |   Max_pixel_.8 <= 112.06443: vegetacao_arborea (2.0)
|   |   |   |   |   |   |   |   |   Max_pixel_.8 > 112.06443: sombra (3.0)
|   |   |   |   |   |   |   |   Max_pixel_.4 > 413.819702: sombra (24.0)
|   |   |   |   |   |   |   Mean_Layer.6 > -50.871429: pavimentacao_viaria (2.0/1.0)
|   |   |   |   |   |   Min_pixel_.8 > -60460.632813: vegetacao_arborea (3.0)
|   |   |   |   |   BRTmN1 > 59.86284
|   |   |   |   |   |   Mean_pc4 <= -70.999516
|   |   |   |   |   |   |   Mean_pc4 <= -81.130268: sombra (9.0)
|   |   |   |   |   |   |   Mean_pc4 > -81.130268: materiais_mistos (2.0)
|   |   |   |   |   |   Mean_pc4 > -70.999516: sombra (157.0)
|   |   |   |   Max_pixel_.7 > 620.852356: coberturas_diversas (3.0/1.0)
|   |   |   modeMinimu.9 > -60
|   |   |   |   modeMinimu.16 <= -279: sombra (2.0/1.0)
|   |   |   |   modeMinimu.16 > -279: coberturas_diversas (4.0)
```

```
|   Mean_Layer.4 > 124.285714
|   |   BRTpBoRE <= 1.044126
|   |   |   NDVI3 <= 0.352876
|   |   |   |   NDSI2 <= -0.423025
|   |   |   |   |   Mean_pc8 <= 72.822729
|   |   |   |   |   |   WVSI <= -0.397283
|   |   |   |   |   |   |   Asymmetry <= 0.839338: ceramica (21.0/2.0)
|   |   |   |   |   |   |   Asymmetry > 0.839338: pavimentacao_viaria (3.0)
|   |   |   |   |   |   WVSI > -0.397283
|   |   |   |   |   |   |   BRTmN1 <= -679.096591
|   |   |   |   |   |   |   |   Max_pixel_.5 <= 2780
|   |   |   |   |   |   |   |   |   Mean_Layer.2 <= 196.5: solo_exposto (21.0)
|   |   |   |   |   |   |   |   |   Mean_Layer.2 > 196.5: pavimentacao_viaria (4.0/1.0)
|   |   |   |   |   |   |   |   Max_pixel_.5 > 2780
|   |   |   |   |   |   |   |   |   BRTpBoRE <= 0.961902: ceramica (14.0/2.0)
|   |   |   |   |   |   |   |   |   BRTpBoRE > 0.961902: materiais_mistos (2.0)
|   |   |   |   |   |   |   BRTmN1 > -679.096591
|   |   |   |   |   |   |   |   NDBI <= 0.047362: ceramica (2.0/1.0)
|   |   |   |   |   |   |   |   NDBI > 0.047362: materiais_mistos (25.0/1.0)
|   |   |   |   |   Mean_pc8 > 72.822729
|   |   |   |   |   |   WVNII <= 4.555916: pavimentacao_viaria (36.0/1.0)
|   |   |   |   |   |   WVNII > 4.555916: ceramica (5.0/1.0)
|   |   |   |   NDSI2 > -0.423025
|   |   |   |   |   NDVI1 <= 0.102211
|   |   |   |   |   |   Max_pixel_.7 <= 127.786697
|   |   |   |   |   |   |   Density <= 0.810602: vegetacao_rasteira (2.0/1.0)
|   |   |   |   |   |   |   Density > 0.810602
|   |   |   |   |   |   |   |   Compactnes <= 1.391695: materiais_mistos (2.0)
|   |   |   |   |   |   |   |   Compactnes > 1.391695: pavimentacao_nao_viaria (22.0/2.0)
|   |   |   |   |   |   Max_pixel_.7 > 127.786697: pavimentacao_viaria (3.0/1.0)
|   |   |   |   |   NDVI1 > 0.102211: materiais_mistos (9.0)
|   |   |   NDVI3 > 0.352876
|   |   |   |   mSWIR <= 1882.697874
|   |   |   |   |   modeMinimu.16 <= -270
|   |   |   |   |   |   Asymmetry <= 0.561596: sombra (2.0)
|   |   |   |   |   |   Asymmetry > 0.561596: materiais_mistos (2.0)
|   |   |   |   |   modeMinimu.16 > -270
|   |   |   |   |   |   Min_pixel_.11 <= -457: coberturas_diversas (2.0)
|   |   |   |   |   |   Min_pixel_.11 > -457
|   |   |   |   |   |   |   NDVI1 <= 0.38647
|   |   |   |   |   |   |   |   Min_pixel_ <= -8: solo_exposto (3.0/1.0)
|   |   |   |   |   |   |   |   Min_pixel_ > -8: vegetacao_rasteira (3.0)
|   |   |   |   |   |   |   NDVI1 > 0.38647
|   |   |   |   |   |   |   |   Max_pixel_.9 <= 39.003502: vegetacao_arborea (3.0/1.0)
|   |   |   |   |   |   |   |   Max_pixel_.9 > 39.003502: vegetacao_rasteira (30.0)
|   |   |   |   mSWIR > 1882.697874: solo_exposto (12.0)
|   |   BRTpBoRE > 1.044126
|   |   |   modeMinimu.14 <= 1259
|   |   |   |   NDVI2 <= 0.078766
|   |   |   |   |   Mean_Layer.5 <= 1705.22291
|   |   |   |   |   |   modeMinimu.11 <= -49
|   |   |   |   |   |   |   Mean_pc8 <= -7.4478
|   |   |   |   |   |   |   |   BRT <= 701.269129: sombra (2.0)
|   |   |   |   |   |   |   |   BRT > 701.269129: pavimentacao_viaria (2.0)
|   |   |   |   |   |   |   Mean_pc8 > -7.4478
|   |   |   |   |   |   |   |   modeMinimu.12 <= -75: coberturas_metalicas (4.0/1.0)
|   |   |   |   |   |   |   |   modeMinimu.12 > -75: coberturas_diversas (8.0/1.0)
```

```
|   |   |   |   |   |   modeMinimu.11 > -49
|   |   |   |   |   |   |   Mean_Layer.2 <= 489.926471
|   |   |   |   |   |   |   |   Mean_pc6 <= -28.193601: coberturas_diversas (3.0)
|   |   |   |   |   |   |   |   Mean_pc6 > -28.193601
|   |   |   |   |   |   |   |   |   modeMinimu.15 <= 233
|   |   |   |   |   |   |   |   |   |   Max_pixel_.10 <= 103.753143: pavimentacao_viaria (7.0)
|   |   |   |   |   |   |   |   |   |   Max_pixel_.10 > 103.753143: coberturas_diversas (4.0)
|   |   |   |   |   |   |   |   |   modeMinimu.15 > 233
|   |   |   |   |   |   |   |   |   |   BRTpBoR <= 1.080609
|   |   |   |   |   |   |   |   |   |   |   Border_ind <= 1.555556: pavimentacao_viaria (2.0)
|   |   |   |   |   |   |   |   |   |   |   Border_ind > 1.555556: pavimentacao_nao_viaria (2.0)
|   |   |   |   |   |   |   |   |   |   BRTpBoR > 1.080609: pavimentacao_viaria (113.0)
|   |   |   |   |   |   |   Mean_Layer.2 > 489.926471
|   |   |   |   |   |   |   |   Max_pixel_.7 <= 264.990662: pavimentacao_viaria (3.0/1.0)
|   |   |   |   |   |   |   |   Max_pixel_.7 > 264.990662: coberturas_metalicas (4.0/1.0)
|   |   |   |   |   Mean_Layer.5 > 1705.22291
|   |   |   |   |   |   Asymmetry <= 0.540947: materiais_mistos (2.0)
|   |   |   |   |   |   Asymmetry > 0.540947: coberturas_metalicas (4.0)
|   |   |   |   NDVI2 > 0.078766
|   |   |   |   |   Mean_pc3 <= -166.275678
|   |   |   |   |   |   WVSI <= -0.071614
|   |   |   |   |   |   |   Mean_pc7 <= 30.195955
|   |   |   |   |   |   |   |   modeMinimu.10 <= 2814
|   |   |   |   |   |   |   |   |   Max_pixel_.14 <= 1214
|   |   |   |   |   |   |   |   |   |   Border_ind <= 1.95: coberturas_diversas (2.0)
|   |   |   |   |   |   |   |   |   |   Border_ind > 1.95: materiais_mistos (2.0)
|   |   |   |   |   |   |   |   |   Max_pixel_.14 > 1214
|   |   |   |   |   |   |   |   |   |   Max_pixel_.6 <= 653.210876
|   |   |   |   |   |   |   |   |   |   |   WVBI <= -0.454478: pavimentacao_viaria (20.0/1.0)
|   |   |   |   |   |   |   |   |   |   |   WVBI > -0.454478: coberturas_metalicas (3.0/1.0)
|   |   |   |   |   |   |   |   |   |   Max_pixel_.6 > 653.210876: coberturas_diversas (2.0)
|   |   |   |   |   |   |   |   modeMinimu.10 > 2814: materiais_mistos (3.0)
|   |   |   |   |   |   |   Mean_pc7 > 30.195955
|   |   |   |   |   |   |   |   Mean_pc8 <= 16.31713
|   |   |   |   |   |   |   |   |   Mean_pc2 <= -83.634262
|   |   |   |   |   |   |   |   |   |   BRT <= 1229.61: materiais_mistos (3.0)
|   |   |   |   |   |   |   |   |   |   BRT > 1229.61: coberturas_metalicas (3.0/1.0)
|   |   |   |   |   |   |   |   |   Mean_pc2 > -83.634262: materiais_mistos (29.0/1.0)
|   |   |   |   |   |   |   |   Mean_pc8 > 16.31713
|   |   |   |   |   |   |   |   |   BRT <= 1626.324519: materiais_mistos (3.0)
|   |   |   |   |   |   |   |   |   BRT > 1626.324519: coberturas_diversas (4.0/1.0)
|   |   |   |   |   |   WVSI > -0.071614
|   |   |   |   |   |   |   BRTpBoR <= 1.574117
|   |   |   |   |   |   |   |   Asymmetry <= 0.983793: materiais_mistos (3.0)
|   |   |   |   |   |   |   |   Asymmetry > 0.983793: coberturas_diversas (2.0)
|   |   |   |   |   |   |   BRTpBoR > 1.574117
|   |   |   |   |   |   |   |   Compactnes <= 1.420455: pavimentacao_viaria (2.0)
|   |   |   |   |   |   |   |   Compactnes > 1.420455
|   |   |   |   |   |   |   |   |   Max_pixel_.6 <= 155.030411: coberturas_metalicas (6.0)
|   |   |   |   |   |   |   |   |   Max_pixel_.6 > 155.030411
|   |   |   |   |   |   |   |   |   |   Max_pixel_.10 <= 57.128319: coberturas_metalicas (3.0)
|   |   |   |   |   |   |   |   |   |   Max_pixel_.10 > 57.128319
|   |   |   |   |   |   |   |   |   |   |   Elliptic_F <= 0.133333: coberturas_metalicas (2.0/1.0)
|   |   |   |   |   |   |   |   |   |   |   Elliptic_F > 0.133333
|   |   |   |   |   |   |   |   |   |   |   |   Max_pixel_.9 <= 257.521179: coberturas_diversas (12.0)
|   |   |   |   |   |   |   |   |   |   |   |   Max_pixel_.9 > 257.521179: coberturas_metalicas (3.0/1.0)
|   |   |   |   |   Mean_pc3 > -166.275678
```

```
|   |   |   |   |   |   BRTpBoR <= 2.470544
|   |   |   |   |   |   |   WVSI <= 0.086165
|   |   |   |   |   |   |   |   Min_pixel_.7 <= -361.988647
|   |   |   |   |   |   |   |   |   modeMinimu.2 <= -960.490356: materiais_mistos (3.0)
|   |   |   |   |   |   |   |   |   modeMinimu.2 > -960.490356
|   |   |   |   |   |   |   |   |   |   BRTpBoRE <= 1.237534
|   |   |   |   |   |   |   |   |   |   |   BRTmN1 <= -507.648601: coberturas_diversas (3.0)
|   |   |   |   |   |   |   |   |   |   |   BRTmN1 > -507.648601: materiais_mistos (3.0/1.0)
|   |   |   |   |   |   |   |   |   |   BRTpBoRE > 1.237534
|   |   |   |   |   |   |   |   |   |   |   Min_pixel_.16 <= -550
|   |   |   |   |   |   |   |   |   |   |   |   Mean_pc7 <= 32.337679: coberturas_diversas (6.0)
|   |   |   |   |   |   |   |   |   |   |   |   Mean_pc7 > 32.337679: materiais_mistos (2.0)
|   |   |   |   |   |   |   |   |   |   |   Min_pixel_.16 > -550: coberturas_diversas (73.0)
|   |   |   |   |   |   |   |   Min_pixel_.7 > -361.988647: coberturas_metalicas (2.0/1.0)
|   |   |   |   |   |   |   WVSI > 0.086165: materiais_mistos (6.0)
|   |   |   |   |   |   BRTpBoR > 2.470544: sombra (4.0/1.0)
|   |   |   modeMinimu.14 > 1259
|   |   |   |   Border_ind <= 1.0625
|   |   |   |   |   Mean_Layer.7 <= 10729.3125
|   |   |   |   |   |   BRTpBoR <= 1.607727
|   |   |   |   |   |   |   Compactnes <= 1.171631: piscina (4.0)
|   |   |   |   |   |   |   Compactnes > 1.171631
|   |   |   |   |   |   |   |   Asymmetry <= 0.580538: pavimentacao_viaria (2.0)
|   |   |   |   |   |   |   |   Asymmetry > 0.580538: coberturas_diversas (2.0)
|   |   |   |   |   |   BRTpBoR > 1.607727
|   |   |   |   |   |   |   BRTmN1 <= -355.203261: coberturas_diversas (29.0/2.0)
|   |   |   |   |   |   |   BRTmN1 > -355.203261: coberturas_metalicas (3.0)
|   |   |   |   |   Mean_Layer.7 > 10729.3125: coberturas_metalicas (14.0)
|   |   |   |   Border_ind > 1.0625
|   |   |   |   |   Mean_pc6 <= -456.371406
|   |   |   |   |   |   NDVI1 <= 0.161182
|   |   |   |   |   |   |   Border_ind <= 1.884615: coberturas_metalicas (3.0)
|   |   |   |   |   |   |   Border_ind > 1.884615: piscina (2.0)
|   |   |   |   |   |   NDVI1 > 0.161182: materiais_mistos (14.0)
|   |   |   |   |   Mean_pc6 > -456.371406
|   |   |   |   |   |   BRTpBoR <= 1.644237
|   |   |   |   |   |   |   BRTpBoRE <= 1.209719: pavimentacao_nao_viaria (2.0/1.0)
|   |   |   |   |   |   |   BRTpBoRE > 1.209719
|   |   |   |   |   |   |   |   Mean_pc3 <= -1041.600834
|   |   |   |   |   |   |   |   |   NDVI1 <= 0.088413: coberturas_metalicas (13.0)
|   |   |   |   |   |   |   |   |   NDVI1 > 0.088413: materiais_mistos (5.0/1.0)
|   |   |   |   |   |   |   |   Mean_pc3 > -1041.600834: materiais_mistos (10.0/1.0)
|   |   |   |   |   |   BRTpBoR > 1.644237
|   |   |   |   |   |   |   Min_pixel_.6 <= -1585.233887: coberturas_metalicas (166.0)
|   |   |   |   |   |   |   Min_pixel_.6 > -1585.233887
|   |   |   |   |   |   |   |   Mean_pc6 <= -286.161124: materiais_mistos (3.0)
|   |   |   |   |   |   |   |   Mean_pc6 > -286.161124
|   |   |   |   |   |   |   |   |   Mean_pc5 <= -37.892119: coberturas_diversas (10.0/2.0)
|   |   |   |   |   |   |   |   |   Mean_pc5 > -37.892119
|   |   |   |   |   |   |   |   |   |   WVBI <= -0.412102
|   |   |   |   |   |   |   |   |   |   |   Border_ind <= 1.794118: pavimentacao_viaria (2.0)
|   |   |   |   |   |   |   |   |   |   |   Border_ind > 1.794118: coberturas_metalicas (2.0)
|   |   |   |   |   |   |   |   |   |   WVBI > -0.412102
|   |   |   |   |   |   |   |   |   |   |   Compactnes <= 1.636364
|   |   |   |   |   |   |   |   |   |   |   |   NDSI1 <= -0.083354: coberturas_diversas (4.0/1.0)
|   |   |   |   |   |   |   |   |   |   |   |   NDSI1 > -0.083354: coberturas_metalicas (7.0)
|   |   |   |   |   |   |   |   |   |   |   Compactnes > 1.636364: coberturas_metalicas (66.0/1.0)
```

```
Number of Leaves  :  108

Size of the tree :  215
```

B.11 Árvore gerada por C5.0 Nível I com MDA

```
C5.0 [Release 2.07 GPL Edition] - (Nível I com MDA)
-------------------------------

Class specified by attribute 'outcome'

Read 1796 cases (94 attributes) from undefined.data

Decision tree:

Max_pixel_.4 <= -1591.778:
:...NDSI2 <= -0.725163: lago (442)
:   NDSI2 > -0.725163:
:   :...Max_pixel_.16 > 1115.027: piscina (5)
:       Max_pixel_.16 <= 1115.027:
:       :...NDVI3 <= 0.221289: coberturas_diversas (3/1)
:           NDVI3 > 0.221289: sombra (16)
Max_pixel_.4 > -1591.778:
:...Mean_Layer.4 <= 124.2857:
    :...Mean_pc3 <= 328.3356:
    :   :...Mean_mda > 54.73729: materiais_mistos (2)
    :   :   Mean_mda <= 54.73729:
    :   :   :...Min_pixel_.11 <= -773: sombra (194/6)
    :   :       Min_pixel_.11 > -773:
    :   :       :...Asymmetry > 0.84915: coberturas_diversas (4)
    :   :           Asymmetry <= 0.84915:
    :   :           :...Max_pixel_.2 > 26.80567: vegetacao_arborea (5)
    :   :               Max_pixel_.2 <= 26.80567:
    :   :               :...Mean_pc7 <= -50.62038: vegetacao_arborea (2)
    :   :                   Mean_pc7 > -50.62038: sombra (9)
    :   Mean_pc3 > 328.3356:
    :   :...Min_pixel_.14 > -259: vegetacao_rasteira (10)
    :       Min_pixel_.14 <= -259:
    :       :...Asymmetry > 0.904824: vegetacao_rasteira (2)
    :           Asymmetry <= 0.904824:
    :           :...Min_pixel_ > -702:
    :               :...Max_pixel_.8 <= 178.4789: vegetacao_arborea (4)
    :               :   Max_pixel_.8 > 178.4789: vegetacao_rasteira (3)
    :               Min_pixel_ <= -702:
    :               :...Max_pixel_.1 > 1587: vegetacao_arborea (130)
    :                   Max_pixel_.1 <= 1587:
    :                   :...modeMinimu.15 <= -658: sombra (2)
    :                       modeMinimu.15 > -658: vegetacao_arborea (3)
    Mean_Layer.4 > 124.2857:
    :...BRTpBoRE <= 1.044126:
        :...NDVI3 > 0.352876:
        :   :...mSWIR > 1882.698: solo_exposto (12)
        :   :   mSWIR <= 1882.698:
        :   :   :...NDBI > 0.209133: coberturas_diversas (3/1)
```

```
:   :       NDBI <= 0.209133:
:   :       :...modeMinimu.15 <= -22: sombra (3/1)
:   :           modeMinimu.15 > -22:
:   :           :...BRTpBoRE <= 0.849527: vegetacao_rasteira (37/3)
:   :               BRTpBoRE > 0.849527: materiais_mistos (2/1)
:   NDVI3 <= 0.352876:
:   :...NDSI2 > -0.423025:
:       :...NDVI1 > 0.102211: materiais_mistos (9)
:       :   NDVI1 <= 0.102211:
:       :   :...Min_pixel_.2 <= -305.3598: materiais_mistos (2)
:       :       Min_pixel_.2 > -305.3598:
:       :       :...Min_pixel_.7 <= -2084.617: materiais_mistos (2/1)
:       :           Min_pixel_.7 > -2084.617:
:       :           :...Rectangula <= 0.916928: pavimentacao_nao_viaria (23/3)
:       :               Rectangula > 0.916928: pavimentacao_viaria (2)
:       NDSI2 <= -0.423025:
:       :...Mean_pc8 > 72.82273:
:           :...BRTmN2 <= -2200.609: ceramica (4)
:           :   BRTmN2 > -2200.609:
:           :   :...Asymmetry <= 0.430605: materiais_mistos (2/1)
:           :       Asymmetry > 0.430605: pavimentacao_viaria (35)
:           Mean_pc8 <= 72.82273:
:           :...WVSI <= -0.397283:
:               :...Asymmetry > 0.839338: pavimentacao_viaria (3)
:               :   Asymmetry <= 0.839338:
:               :   :...Max_pixel_.2 <= 21.46643: pavimentacao_viaria (2/1)
:               :       Max_pixel_.2 > 21.46643: ceramica (19)
:               WVSI > -0.397283:
:               :...BRTmN1 > -679.0966:
:                   :...Mean_pc5 <= -78.82233: materiais_mistos (25/1)
:                   :   Mean_pc5 > -78.82233: pavimentacao_viaria (2/1)
:                   BRTmN1 <= -679.0966:
:                   :...Max_pixel_.12 <= 380: solo_exposto (19)
:                       Max_pixel_.12 > 380:
:                       :...Mean_pc7 <= -55.17622:
:                           :...Max_pixel_.4 <= 2509.087: solo_exposto (3)
:                           :   Max_pixel_.4 > 2509.087: ceramica (13/1)
:                           Mean_pc7 > -55.17622: [S1]
BRTpBoRE > 1.044126:
:...Mean_mda <= 25.56667:
    :...modeMinimu.16 > 1198:
    :   :...NDSI1 <= -0.165599:
    :   :   :...NDSI1 <= -0.240224:
    :   :   :   :...Compactnes <= 1.16129: coberturas_diversas (4)
    :   :   :   :   Compactnes > 1.16129: pavimentacao_nao_viaria (2/1)
    :   :   :   NDSI1 > -0.240224:
    :   :   :   :...Mean_pc7 <= 131.8085: materiais_mistos (23)
    :   :   :       Mean_pc7 > 131.8085: coberturas_metalicas (2/1)
    :   :   NDSI1 > -0.165599:
    :   :   :...Max_pixel_.8 > 389.3307: coberturas_diversas (2/1)
    :   :       Max_pixel_.8 <= 389.3307:
    :   :       :...Mean_pc5 <= -67.4882: piscina (4)
    :   :           Mean_pc5 > -67.4882:
    :   :           :...WVSI <= 0.060472: coberturas_metalicas (20)
    :   :               WVSI > 0.060472: piscina (2)
    :   modeMinimu.16 <= 1198:
    :   :...NDVI2 <= 0.078766:
```

```
:       :...Mean_Layer.5 > 1705.223: coberturas_metalicas (2/1)
:       :   Mean_Layer.5 <= 1705.223:
:       :   :...WVBI <= -0.711179:
:       :       :...BRTmN1 <= 40.56654: pavimentacao_viaria (2)
:       :       :   BRTmN1 > 40.56654: pavimentacao_nao_viaria (3/1)
:       :       WVBI > -0.711179:
:       :       :...Elliptic_F > 0.077551: pavimentacao_viaria (106)
:       :           Elliptic_F <= 0.077551:
:       :           :...Mean_pc4 <= -23.44537: coberturas_diversas (2)
:       :               Mean_pc4 > -23.44537: pavimentacao_viaria (13)
:       NDVI2 > 0.078766:
:       :...Mean_pc3 <= -162.9431:
:           :...WVBI > -0.384092: coberturas_metalicas (2)
:           :   WVBI <= -0.384092:
:           :   :...Mean_pc7 > 30.19596: [S2]
:           :       Mean_pc7 <= 30.19596: [S3]
:           Mean_pc3 > -162.9431:
:           :...modeMinimu.9 <= -375: sombra (2)
:               modeMinimu.9 > -375:
:               :...Min_pixel_.9 > 23.62909: pavimentacao_viaria (2)
:                   Min_pixel_.9 <= 23.62909:
:                   :...Mean_pc8 <= -54.65574: materiais_mistos (7)
:                       Mean_pc8 > -54.65574: [S4]
Mean_mda > 25.56667:
:...modeMinimu.15 > 985:
    :...Min_pixel_.9 > 46.86615: coberturas_diversas (10)
    :   Min_pixel_.9 <= 46.86615:
    :   :...Rectangula > 0.945603:
    :       :...Min_pixel_.4 <= -102.7313: coberturas_diversas (11)
    :       :   Min_pixel_.4 > -102.7313: coberturas_metalicas (6/1)
    :       Rectangula <= 0.945603:
    :       :...NDSI1 <= -0.162248:
    :           :...Min_pixel_.8 <= -84592.38: coberturas_metalicas (53)
    :           :   Min_pixel_.8 > -84592.38:
    :           :   :...Mean_pc8 <= -17.06399: coberturas_metalicas (8/1)
    :           :       Mean_pc8 > -17.06399: [S5]
    :           NDSI1 > -0.162248:
    :           :...Asymmetry > 0.2518: coberturas_metalicas (190/1)
    :               Asymmetry <= 0.2518:
    :               :...Mean_pc8 <= 20.60887: coberturas_metalicas (10)
    :                   Mean_pc8 > 20.60887:
    :                   :...BRT <= 3891.362: materiais_mistos (2)
    :                       BRT > 3891.362: coberturas_diversas (2)
    modeMinimu.15 <= 985:
    :...Max_pixel_.8 > 625.2779:
        :...Min_pixel_.3 <= -574.6019: materiais_mistos (8)
        :   Min_pixel_.3 > -574.6019:
        :   :...Rectangula <= 0.703834: coberturas_diversas (3/1)
        :       Rectangula > 0.703834: sombra (2)
        Max_pixel_.8 <= 625.2779:
        :...BRTmN2 > 150.9134:
            :...Density <= 1.049681: coberturas_metalicas (4/1)
            :   Density > 1.049681: pavimentacao_viaria (3)
            BRTmN2 <= 150.9134:
            :...modeMinimu.10 <= -669: materiais_mistos (5/1)
                modeMinimu.10 > -669:
                :...Mean_mda > 30.23409: [S6]
```

```
                                Mean_mda <= 30.23409:
                                :...Mean_pc3 <= -109.6513: [S7]
                                    Mean_pc3 > -109.6513: [S8]

SubTree [S1]

modeMinimu.9 <= 2070: pavimentacao_viaria (3)
modeMinimu.9 > 2070: materiais_mistos (3)

SubTree [S2]

Max_pixel_.2 <= 23.72186: pavimentacao_viaria (4/1)
Max_pixel_.2 > 23.72186: materiais_mistos (32)

SubTree [S3]

Max_pixel_.2 > 26.50727: coberturas_metalicas (2/1)
Max_pixel_.2 <= 26.50727:
:...modeMinimu.10 <= 2814: pavimentacao_viaria (25/1)
    modeMinimu.10 > 2814: materiais_mistos (3)

SubTree [S4]

Border_ind <= 3.066667: coberturas_diversas (17)
Border_ind > 3.066667: materiais_mistos (2)

SubTree [S5]

Max_pixel_.6 <= 105.0572: coberturas_metalicas (2)
Max_pixel_.6 > 105.0572: coberturas_diversas (21/1)

SubTree [S6]

modeMinimu.12 <= 1105: coberturas_diversas (66/1)
modeMinimu.12 > 1105: materiais_mistos (3/1)

SubTree [S7]

modeMinimu.8 > 27.09198: coberturas_metalicas (9)
modeMinimu.8 <= 27.09198:
:...Asymmetry <= 0.662958: materiais_mistos (2)
    Asymmetry > 0.662958: pavimentacao_viaria (2)

SubTree [S8]

Mean_mda > 29.63165: coberturas_metalicas (2)
Mean_mda <= 29.63165:
:...Min_pixel_.3 <= -180.8241: coberturas_diversas (21)
    Min_pixel_.3 > -180.8241: coberturas_metalicas (2/1)
```

B.12 Árvore gerada por C5.0 Nível I sem MDA

```
C5.0 [Release 2.07 GPL Edition] - (Nível I sem MDA)
-------------------------------
```

```
Class specified by attribute ‘outcome’

Read 1796 cases (90 attributes) from undefined.data

Decision tree:

Max_pixel_.4 <= -1591.778:
:...NDSI2 <= -0.725163: lago (442)
:   NDSI2 > -0.725163:
:   :...Max_pixel_.16 > 1115.027: piscina (5)
:       Max_pixel_.16 <= 1115.027:
:       :...NDVI3 <= 0.221289: coberturas_diversas (3/1)
:           NDVI3 > 0.221289: sombra (16)
Max_pixel_.4 > -1591.778:
:...Mean_Layer.4 <= 124.2857:
    :...Mean_pc3 > 328.3356:
    :   :...Min_pixel_.14 > -259: vegetacao_rasteira (10)
    :   :   Min_pixel_.14 <= -259:
    :   :   :...Asymmetry > 0.904824: vegetacao_rasteira (2)
    :   :       Asymmetry <= 0.904824:
    :   :       :...Min_pixel_ > -702:
    :   :           :...Max_pixel_.8 <= 178.4789: vegetacao_arborea (4)
    :   :           :   Max_pixel_.8 > 178.4789: vegetacao_rasteira (3)
    :   :           Min_pixel_ <= -702:
    :   :           :...Max_pixel_.1 > 1587: vegetacao_arborea (130)
    :   :               Max_pixel_.1 <= 1587:
    :   :               :...modeMinimu.15 <= -658: sombra (2)
    :   :                   modeMinimu.15 > -658: vegetacao_arborea (3)
    :   Mean_pc3 <= 328.3356:
    :   :...Min_pixel_.2 <= -500.2198: materiais_mistos (2)
    :       Min_pixel_.2 > -500.2198:
    :       :...modeMinimu.9 > -60:
    :           :...modeMinimu.16 > -244: coberturas_diversas (4)
    :           :   modeMinimu.16 <= -244:
    :           :   :...Mean_Layer.2 <= -277.3499: sombra (2)
    :           :       Mean_Layer.2 > -277.3499: vegetacao_arborea (3)
    :           modeMinimu.9 <= -60:
    :           :...Max_pixel_.7 > 620.8524: coberturas_diversas (3/1)
    :               Max_pixel_.7 <= 620.8524:
    :               :...WVBI <= -2.632097:
    :                   :...Asymmetry <= 0.546282: vegetacao_arborea (2)
    :                   :   Asymmetry > 0.546282: sombra (3/1)
    :                   WVBI > -2.632097:
    :                   :...BRTmN1 > 77.2381: sombra (165)
    :                       BRTmN1 <= 77.2381:
    :                       :...BRTmN2 <= -29.34511: sombra (27/1)
    :                           BRTmN2 > -29.34511: [S1]
    Mean_Layer.4 > 124.2857:
    :...BRTpBoRE <= 1.044126:
        :...NDVI3 > 0.352876:
        :   :...mSWIR > 1882.698: solo_exposto (12)
        :   :   mSWIR <= 1882.698:
        :   :   :...NDBI > 0.209133: coberturas_diversas (3/1)
        :   :       NDBI <= 0.209133:
        :   :       :...modeMinimu.15 <= -22: sombra (3/1)
        :   :           modeMinimu.15 > -22:
        :   :           :...BRTpBoRE <= 0.849527: vegetacao_rasteira (37/3)
```

```
:   :               BRTpBoRE > 0.849527: materiais_mistos (2/1)
:   NDVI3 <= 0.352876:
:   :...NDSI2 > -0.423025:
:       :...NDVI1 > 0.102211: materiais_mistos (9)
:       :   NDVI1 <= 0.102211:
:       :   :...Min_pixel_.2 <= -305.3598: materiais_mistos (2)
:       :       Min_pixel_.2 > -305.3598:
:       :       :...Min_pixel_.7 <= -2084.617: materiais_mistos (2/1)
:       :           Min_pixel_.7 > -2084.617:
:       :           :...Rectangula <= 0.916928: pavimentacao_nao_viaria (23/3)
:       :               Rectangula > 0.916928: pavimentacao_viaria (2)
:       NDSI2 <= -0.423025:
:       :...Mean_pc8 > 72.82273:
:           :...BRTmN2 <= -2200.609: ceramica (4)
:           :   BRTmN2 > -2200.609:
:           :   :...Asymmetry <= 0.430605: materiais_mistos (2/1)
:           :       Asymmetry > 0.430605: pavimentacao_viaria (35)
:           Mean_pc8 <= 72.82273:
:           :...WVSI <= -0.397283:
:               :...Asymmetry <= 0.839338: ceramica (21/2)
:               :   Asymmetry > 0.839338: pavimentacao_viaria (3)
:               WVSI > -0.397283:
:               :...BRTmN1 > -679.0966:
:                   :...Mean_pc5 <= -78.82233: materiais_mistos (25/1)
:                   :   Mean_pc5 > -78.82233: pavimentacao_viaria (2/1)
:                   BRTmN1 <= -679.0966:
:                   :...Max_pixel_.12 <= 380: solo_exposto (19)
:                       Max_pixel_.12 > 380:
:                       :...Mean_pc7 <= -55.17622:
:                           :...Max_pixel_.4 <= 2509.087: solo_exposto (3)
:                           :   Max_pixel_.4 > 2509.087: ceramica (13/1)
:                           Mean_pc7 > -55.17622: [S2]
BRTpBoRE > 1.044126:
:...modeMinimu.14 <= 1259:
    :...NDVI2 <= 0.078766:
    :   :...Mean_Layer.5 > 1705.223:
    :   :   :...Min_pixel_.7 <= -2119.419: coberturas_metalicas (4)
    :   :   :   Min_pixel_.7 > -2119.419: materiais_mistos (2)
    :   :   Mean_Layer.5 <= 1705.223:
    :   :   :...modeMinimu.11 <= -49:
    :   :       :...Mean_Layer.5 > 1141: pavimentacao_viaria (2)
    :   :       :   Mean_Layer.5 <= 1141:
    :   :       :   :...modeMinimu.15 <= -226: sombra (3/1)
    :   :       :       modeMinimu.15 > -226: [S3]
    :   :       modeMinimu.11 > -49:
    :   :       :...Mean_Layer.2 > 489.9265:
    :   :           :...Max_pixel_.16 > 279.4071: materiais_mistos (2/1)
    :   :           :   Max_pixel_.16 <= 279.4071: [S4]
    :   :           Mean_Layer.2 <= 489.9265:
    :   :           :...Mean_pc6 <= -28.1936: coberturas_diversas (3)
    :   :               Mean_pc6 > -28.1936:
    :   :               :...modeMinimu.15 <= 233: [S5]
    :   :                   modeMinimu.15 > 233:
    :   :                   :...BRTpBoR > 1.080609: pavimentacao_viaria (113)
    :   :                       BRTpBoR <= 1.080609: [S6]
    :   NDVI2 > 0.078766:
    :   :...Mean_pc3 > -166.2757:
```

```
:       :...BRTpBoR > 2.470544: sombra (4/1)
:       :   BRTpBoR <= 2.470544:
:       :   :...Mean_pc8 <= -87.0129: materiais_mistos (6)
:       :       Mean_pc8 > -87.0129:
:       :       :...Min_pixel_.7 > -361.9886: coberturas_metalicas (2/1)
:       :           Min_pixel_.7 <= -361.9886: [S7]
:       Mean_pc3 <= -166.2757:
:       :...WVSI <= -0.071614:
:           :...Mean_pc7 > 30.19596:
:           :   :...Mean_pc8 > 16.31713:
:           :   :   :...BRT <= 1626.324: materiais_mistos (3)
:           :   :   :   BRT > 1626.324: coberturas_diversas (4/1)
:           :   :   Mean_pc8 <= 16.31713:
:           :   :   :...Mean_pc2 > -83.63426: materiais_mistos (29/1)
:           :   :       Mean_pc2 <= -83.63426:
:           :   :       :...BRT <= 1229.61: materiais_mistos (3)
:           :   :           BRT > 1229.61: coberturas_metalicas (3/1)
:           :   Mean_pc7 <= 30.19596:
:           :   :...modeMinimu.10 > 2814: materiais_mistos (3)
:           :       modeMinimu.10 <= 2814:
:           :       :...BRTmN1 > -236.8096: [S8]
:           :           BRTmN1 <= -236.8096: [S9]
:           WVSI > -0.071614:
:           :...Max_pixel_.4 > 4210.683: pavimentacao_viaria (2)
:               Max_pixel_.4 <= 4210.683:
:               :...Elliptic_F > 0.783582: materiais_mistos (2/1)
:                   Elliptic_F <= 0.783582: [S10]
modeMinimu.14 > 1259:
:...Border_ind <= 1.0625:
    :...Max_pixel_.15 > 13167: coberturas_metalicas (13)
    :   Max_pixel_.15 <= 13167:
    :   :...BRTpBoR > 1.607727:
    :       :...NDSI2 <= -0.083358: coberturas_diversas (30/3)
    :       :   NDSI2 > -0.083358: coberturas_metalicas (3)
    :       BRTpBoR <= 1.607727:
    :       :...Max_pixel_.1 > 8881: coberturas_diversas (2)
    :           Max_pixel_.1 <= 8881:
    :           :...Min_pixel_.6 <= -1287.797: piscina (4)
    :               Min_pixel_.6 > -1287.797: pavimentacao_viaria (2)
    Border_ind > 1.0625:
    :...Mean_pc6 <= -456.3714:
        :...NDSI1 > 0.015258: piscina (2)
        :   NDSI1 <= 0.015258:
        :   :...Max_pixel_.6 <= 64.30662: coberturas_metalicas (3)
        :       Max_pixel_.6 > 64.30662: materiais_mistos (14)
        Mean_pc6 > -456.3714:
        :...BRTpBoR <= 1.644237:
            :...NDSI1 <= -0.236498: pavimentacao_nao_viaria (2/1)
            :   NDSI1 > -0.236498:
            :   :...Mean_pc3 > -1041.601: materiais_mistos (10/1)
            :       Mean_pc3 <= -1041.601:
            :       :...NDVI3 <= 0.103925: coberturas_metalicas (13)
            :           NDVI3 > 0.103925: materiais_mistos (5/1)
            BRTpBoR > 1.644237:
            :...Min_pixel_.6 <= -1585.234: coberturas_metalicas (166)
                Min_pixel_.6 > -1585.234:
                :...Mean_pc6 <= -287.1251: materiais_mistos (3)
```

```
                                Mean_pc6 > -287.1251:
                                :...Mean_pc5 <= -27.10127: [S11]
                                    Mean_pc5 > -27.10127:
                                    :...WVBI <= -0.412102: [S12]
                                        WVBI > -0.412102: [S13]

SubTree [S1]

Max_pixel_.13 <= 178: vegetacao_arborea (3)
Max_pixel_.13 > 178: pavimentacao_viaria (2/1)

SubTree [S2]

modeMinimu.9 <= 2070: pavimentacao_viaria (3)
modeMinimu.9 > 2070: materiais_mistos (3)

SubTree [S3]

Max_pixel_.10 <= 85.14075: coberturas_metalicas (3)
Max_pixel_.10 > 85.14075: coberturas_diversas (8/1)

SubTree [S4]

Max_pixel_.7 <= 264.9907: pavimentacao_viaria (2)
Max_pixel_.7 > 264.9907: coberturas_metalicas (3)

SubTree [S5]

Max_pixel_.10 <= 103.7531: pavimentacao_viaria (7)
Max_pixel_.10 > 103.7531: coberturas_diversas (4)

SubTree [S6]

Border_ind <= 1.555556: pavimentacao_viaria (2)
Border_ind > 1.555556: pavimentacao_nao_viaria (2)

SubTree [S7]

Min_pixel_.3 <= -960.4904: materiais_mistos (3)
Min_pixel_.3 > -960.4904: coberturas_diversas (87/5)

SubTree [S8]

Border_ind <= 1.95: coberturas_diversas (2)
Border_ind > 1.95: materiais_mistos (2)

SubTree [S9]

Max_pixel_.6 > 653.2109: coberturas_diversas (2)
Max_pixel_.6 <= 653.2109:
:...WVBI <= -0.454478: pavimentacao_viaria (20/1)
    WVBI > -0.454478: coberturas_metalicas (3/1)

SubTree [S10]

Min_pixel_.7 > -1676.546: coberturas_metalicas (9/1)
Min_pixel_.7 <= -1676.546:
```

```
:...NDBI <= 0.014932: coberturas_metalicas (3/1)
    NDBI > 0.014932:
    :...Mean_pc7 <= 9.056426: coberturas_diversas (11)
        Mean_pc7 > 9.056426:
        :...NDSI2 <= -0.235711: coberturas_diversas (3)
            NDSI2 > -0.235711: coberturas_metalicas (3/1)

SubTree [S11]

Min_pixel_.5 <= -501.5183: coberturas_metalicas (2/1)
Min_pixel_.5 > -501.5183: coberturas_diversas (9)

SubTree [S12]

Border_ind <= 1.794118: pavimentacao_viaria (2)
Border_ind > 1.794118: coberturas_metalicas (2)

SubTree [S13]

Compactnes > 1.636364: coberturas_metalicas (65)
Compactnes <= 1.636364:
:...NDVI1 <= 0.072262: coberturas_metalicas (8)
    NDVI1 > 0.072262: coberturas_diversas (3)
```

APÊNDICE C - MATRIZES DE CONFUSÃO C4.5, C5.0 E CART PARA NÍVEL I E NÍVEL II

Tabela C.1 - Matriz de confusão para classificação por CART Nível I com MDA.

Tabela C.2 - Matriz de confusão para classificação por C4.5 Nível I sem MDA.

Tabela C.3 - Matriz de confusão para classificação por C5.0 Nível I com MDA.

Tabela C.4 - Matriz de confusão para classificação por C5.0 Nível I sem MDA.

Tabela C.5 - Matriz de confusão para classificação por CART Nível II sem MDA.

Tabela C.6 - Matriz de confusão para classificação por C4.5 Nível II com MDA.

Tabela C.7 - Matriz de confusão para classificação por C4.5 Nível II sem MDA.

Tabela C.8 - Matriz de confusão para classificação por C5.0 Nível II sem MDA.

APÊNDICE D - IMAGENS TEMÁTICAS CLASSIFICADAS

Figura D.1 - Imagens temáticas classificadas para as legendas do Nível I, que foram omitidas por apresentarem índices Kappa inferiores aos dados justificados no capítulo de discussão dos resultados.

Figura D.2 - Imagens temáticas classificadas para o Nível de Legenda I que foram omitidas da discussão de resultado por apresentarem alta similaridade estatística no teste Z.

APÊNDICE E - EXPERIMENTOS REALIZADOS NO LARAD PARA O ESTUDO DO COMPORTAMENTO ESPECTRAL DE ALVOS URBANOS

Figura E.1 - Galvalume Ondulado.

Figura E.2 - Galvalume Trapezoidal.

Figura E.3 - Galvalume Trapezoidal Tamanho 40 com Pintura Eletrostática Branca.

Figura E.4 - Aço Galvanizado Trapezoidal Tamanho 25.

Figura E.5 - Galvalume tipo sanduíche com pintura eletrostática branca.

Figura E.6 - Telha de fibrocimento CRFS.

Figura E.7 - Telha cerâmica resinada tipo portuguesa.

Figura E.8 - Quartzito amarelo (Pedra mineira).

Figura E.9 - Telha de concreto.

APÊNDICE F - SCRIPT [R] PARA PROCESSAMENTO, CLASSIFICAÇÃO E VALIDAÇÃO DOS SEGMENTOS DA IMAGEM WV-2

```
#################################################################################################
#                                                                                               #
# Disclaimer: This is a code stub with solely educational purposes. Every external package and library          #
# used in here belongs to its respective owners. The present document is a very simple script intended to       #
# help somehow in the optimization of the automatic classification process of remote sensing images,            #
# through means of machine learning fed with CSV data. I would like to apologize beforehand for the             #
# unoptimized "dirty-coding" and the almost-non-existant commenting in the sections to come. Don't mind         #
# the directory folders path, just replace it with yours if you please. I would also like to thank for the many #
# hands and brains borrowed from my fellows at INPE in the development of this patchwork. Special thanks to     #
# Sacha, Marujo and Pompeu for their contributions in the validation process. Feedback, comments                #
# and tomatoes can be thrown away at dvdgmf@gmail.com                                                           #
#                                                                                               #
#################################################################################################

require(sp)
require(raster)
library(rgdal)
library(maptools)
#install.packages("raster")
#library(RWeka)
library(foreign)
library(caret)

#PC ENVIRONMENT HISTORY
load("C:/Users/David Monteiro/Google Drive/MASTER/.RData")

# PC
setwd("C:/Users/David Monteiro/Google Drive/MASTER/")
# NTBK
setwd("C:/Users/David/Google Drive/MASTER/")

oldmapa = mapa
mapa = readOGR("C:/Users/David Monteiro/Google Drive/MASTER/ecognition/SHPs/pcmda23.shp",layer = "pcmda23")

colnames(mapa@data)
#oldt = tabela
tabela = mapa@data

amostras = tabela
amostras = amostras[!(amostras$Class_name=="unclassified"),]
write.csv(amostras, file="C:/Users/David Monteiro/Google Drive/MASTER/CSV/pcmda23_amostras.csv", quote = F, row.names = F)
```

```
noclass = tabela
noclass$Class_name = "unclassified"
write.csv(noclass, file="C:/Users/David Monteiro/Google Drive/MASTER/CSV/pcmda23_noclass.csv", quote = F, row.names = F)

#write.arff(noclass, file="C:/Users/David Monteiro/Google Drive/MASTER/CSV/pcmda16_noclass.arff")
# CLASSIFICATION

#pcmda7lv3c45 = read.arff("C:/Users/David Monteiro/Google Drive/MASTER/CSV/pcmda7_lv3_c45_bag_boost.arff")
#pcmda8c45bab = read.arff("C:/Users/David Monteiro/Google Drive/MASTER/CSV/pcmda8_c45_bag_boost.arff")
pcmda = read.arff("C:/Users/David Monteiro/Google Drive/MASTER/CSV/pcmda23_cart_lv3_nomda_prd.arff")

colnames(pcmda)[9]
colnames(pcmda)[9] = "prdclass"
tabela$crt3nmd = pcmda$prdclass
colnames(tabela)
mapa@data = tabela
writeOGR(mapa,dsn="C:/Users/David Monteiro/Google Drive/MASTER/ecognition/SHPs/pcmda23_lv36_prd.shp",
         layer = "pcmda23_lv36_prd",
         driver = "ESRI Shapefile")

# REMAPEAR CLASSES DO NÍVEL II PARA O NÍVEL I (lv6 -> lv3) ----------------------------------------

# exporta comandos do console para janela externa
library(rite)
sinkstart(echo=FALSE)

library("foreign") # permite manipulação de arquivos .arff do WEKA
amostralv3 = read.arff("C:/Users/David Monteiro/Google Drive/MASTER/CSV/pcmda23_amostras.csv.arff")
colnames(amostralv3)
unique(amostralv3$Class_name)

#testar classe
amostralv3$Class_name[which(amostralv3$Class_name=="concreto_pintado_vermelho")]

# classes que não precisam mudar de nome do lv6 -> lv3

#levels(amostralv3$Class_name)[levels(amostralv3$Class_name)=="sombra"]                          = "sombra"
#levels(amostralv3$Class_name)[levels(amostralv3$Class_name)=="solo_exposto"]                     = "solo_exposto"
#levels(amostralv3$Class_name)[levels(amostralv3$Class_name)=="lago"]                             = "lago"
#levels(amostralv3$Class_name)[levels(amostralv3$Class_name)=="vegetacao_rasteira"]               = "vegetacao_rasteira"
#levels(amostralv3$Class_name)[levels(amostralv3$Class_name)=="vegetacao_arborea"]                = "vegetacao_arborea"
#levels(amostralv3$Class_name)[levels(amostralv3$Class_name)=="ceramica"]                         = "ceramica"

levels(amostralv3$Class_name)[levels(amostralv3$Class_name)=="piscina_azulejo"]                   = "piscina"
levels(amostralv3$Class_name)[levels(amostralv3$Class_name)=="piscina_vinil"]                     = "piscina"
```

```
levels(amostralv3$Class_name)[levels(amostralv3$Class_name)=="quartzito_amarelo"]                    = "pavimentacao_nao_viaria"
levels(amostralv3$Class_name)[levels(amostralv3$Class_name)=="carvao_coque"]                         = "pavimentacao_nao_viaria"
levels(amostralv3$Class_name)[levels(amostralv3$Class_name)=="asfalto"]                              = "pavimentacao_viaria"
levels(amostralv3$Class_name)[levels(amostralv3$Class_name)=="asfalto_pintado_vermelho"]             = "pavimentacao_viaria"
levels(amostralv3$Class_name)[levels(amostralv3$Class_name)=="asfalto_pintado_branco"]               = "pavimentacao_viaria"
levels(amostralv3$Class_name)[levels(amostralv3$Class_name)=="bloquete"]                             = "pavimentacao_viaria"
levels(amostralv3$Class_name)[levels(amostralv3$Class_name)=="argamassa_cimento"]                    = "materiais_mistos"
levels(amostralv3$Class_name)[levels(amostralv3$Class_name)=="argamassa_cimento_pintado_vermelho"]   = "materiais_mistos"
levels(amostralv3$Class_name)[levels(amostralv3$Class_name)=="argamassa_cimento_pintado_verde"]      = "materiais_mistos"
levels(amostralv3$Class_name)[levels(amostralv3$Class_name)=="concreto_pintado_amarelo"]             = "materiais_mistos"
levels(amostralv3$Class_name)[levels(amostralv3$Class_name)=="concreto_pintado_vermelho"]            = "materiais_mistos"
levels(amostralv3$Class_name)[levels(amostralv3$Class_name)=="concreto"]                             = "materiais_mistos"
levels(amostralv3$Class_name)[levels(amostralv3$Class_name)=="concreto_impermeabilizado"]            = "materiais_mistos"
levels(amostralv3$Class_name)[levels(amostralv3$Class_name)=="vidro_aramado"]                        = "coberturas_diversas"
levels(amostralv3$Class_name)[levels(amostralv3$Class_name)=="fibra_vidro"]                          = "coberturas_diversas"
levels(amostralv3$Class_name)[levels(amostralv3$Class_name)=="pastilha_esmaltada"]                   = "coberturas_diversas"
levels(amostralv3$Class_name)[levels(amostralv3$Class_name)=="fibrocimento_crfs"]                    = "coberturas_diversas"
levels(amostralv3$Class_name)[levels(amostralv3$Class_name)=="fibrocimento_amianto"]                 = "coberturas_diversas"
levels(amostralv3$Class_name)[levels(amostralv3$Class_name)=="policarbonato"]                        = "coberturas_diversas"
levels(amostralv3$Class_name)[levels(amostralv3$Class_name)=="acrilico"]                             = "coberturas_diversas"
levels(amostralv3$Class_name)[levels(amostralv3$Class_name)=="plastico"]                             = "coberturas_diversas"
levels(amostralv3$Class_name)[levels(amostralv3$Class_name)=="lona"]                                 = "coberturas_diversas"
levels(amostralv3$Class_name)[levels(amostralv3$Class_name)=="aco_brilhante"]                        = "coberturas_metalicas"
levels(amostralv3$Class_name)[levels(amostralv3$Class_name)=="aco_ferrugem"]                         = "coberturas_metalicas"
levels(amostralv3$Class_name)[levels(amostralv3$Class_name)=="aco_fosco"]                            = "coberturas_metalicas"
levels(amostralv3$Class_name)[levels(amostralv3$Class_name)=="galvalume_a"]                          = "coberturas_metalicas"
levels(amostralv3$Class_name)[levels(amostralv3$Class_name)=="galvalume_b"]                          = "coberturas_metalicas"
levels(amostralv3$Class_name)[levels(amostralv3$Class_name)=="galvalume_c"]                          = "coberturas_metalicas"

# excluída
# levels(amostralv3$Class_name)[levels(amostralv3$Class_name)=="manta_asfaltica_aluminizada"]       = "coberturas_diversas"

# SALVAR NOVO CSV COM AS CLASSES REMAPEADAS
write.csv(amostralv3, file="C:/Users/David Monteiro/Google Drive/MASTER/CSV/pcmda23_amostras_lv3.csv", quote = F, row.names = F)

# GERAR PONTOS ALEATORIOS ---------------------------------------------------

# NTBK
validamapa = readOGR("C:/Users/David/Google Drive/MASTER/ecognition/SHPs/pcmda23_lv36_prd.shp",
                     layer = "pcmda23_lv36_prd")
# PC
validamapa = readOGR("C:/Users/David Monteiro/Google Drive/MASTER/ecognition/SHPs/pcmda23_lv36_prd.shp",
                     layer = "pcmda23_lv36_prd")

# Contador de amostras por classe ------------------------------------------
```

```
contamostra = table(unlist(validamapa@data$Clss_nm))
as.data.frame(contamostra)

# Gera coluna com valores de área por classe
validamapa@data$area = sapply(validamapa@polygons, FUN=function(x) {slot(x, 'area')})

# Gera SHP de pontos
mapapontos = s.samples(numPts = 1080,
                       thematicMap = validamapa,
                       classAttr = validamapa@data$c456cmd,
                       areaAttr = validamapa@data$area,
                       #minPtsPerClass = 30,
                       #maxPtsPerClass = 60
                       )

#biblioteca da função count()
library(plyr)

count(pts$c456cmd)
count(validamapa@data$c456cmd)
count(oldvalidapontos@data$c456cmd)

#gera coluna com nome das classes cujos pontos sobrepuserem o segmento
mapapontos = SpatialPointsDataFrame(data = data.frame(over(mapapontos, validamapa)$c456cmd), coords = mapapontos@coords)

#data.frame(levels(validamapa@data$c456cmd))
plot(validamapa)
plot(mapapontos,col="red", add=T)

writeOGR(mapapontos,
         dsn ="C:/Users/David Monteiro/Google Drive/MASTER/ecognition/SHPs/pontos.shp",
         layer = "pontos",
         driver = "ESRI Shapefile",
         overwrite_layer = T)

#writeSpatialShape(mapapontos,"C:/Users/David/Google Drive/MASTER/ecognition/SHPs/pcmda23_pontos.shp")

################################################################################
# FUNCAO s.samples() cedida por Sacha: GERA PONTOS ALEATORIOS PARA UM MAPA VETORIAL #
################################################################################

# Random sample points based on classes proportion (s.samples) ----------------

# DOCUMENTATION
# requires the package{sp}
```

```
# Arguments
#   numPts
#       number of points you would like to sample. Maybe the function results in a
#       bigger number of points, because the number of points per class is rounded
#       for the nearest larger integer;
#   thematicMap
#       must be a SpatialPolygonsDataFrame{sp};
#   classAttr
#       must be something like this: thematicMap@data$class;
#   areaAttr
#       must be something like this: thematicMap@data$area;
#   minPtsPerClass
#       minimum number of points per class;
#   maxPtsPerClass
#       maximum number of points per class;
# Return
#   A SpatialPoints object with the random sample points based on classes proportion.

s.samples <- function(numPts = 100, thematicMap, classAttr, areaAttr, minPtsPerClass = 1, maxPtsPerClass = 10000){
  require(raster)

  ptsPerClass <- data.frame(levels(classAttr))
  names(ptsPerClass) <- c("class")

  for (class in 1:length(levels(classAttr))) {

    ptsPerClass$numPts[class] <- ceiling(numPts * sum(areaAttr[classAttr == levels(classAttr)[class]]) / sum(areaAttr))

    if(ptsPerClass$numPts[class] < minPtsPerClass){

      ptsPerClass$numPts[class] = minPtsPerClass
    }
    if(ptsPerClass$numPts[class] > maxPtsPerClass){

      ptsPerClass$numPts[class] = maxPtsPerClass
    }
    if (class == 1){

      pts <- spsample(x = thematicMap[classAttr == levels(classAttr)[class],],
                      n = ptsPerClass$numPts[ptsPerClass$class == levels(classAttr)[class]],
                      type = "random")

    } else {pts <- bind(pts, spsample(x = thematicMap[classAttr == levels(classAttr)[class],],
                                      n = ptsPerClass$numPts[ptsPerClass$class == levels(classAttr)[class]],
                                      type = "random"))
```

```
    }

  }
  print(ptsPerClass)
  #pts = as(pts,"SpatialPolygonsDataFrame")
  #pts = as(pts,"SpatialPointsDataFrame")
  return(pts)
}

##################################################
# GERA ARVORE DO MODELO DE CLASSIFICACAO POR C5.0 #
##################################################

# PC
setwd("C:/Users/David Monteiro/Google Drive/MASTER/")
# NTBK
setwd("C:/Users/David/Google Drive/MASTER/")

#-----------------------------------------------------------------
# http://connor-johnson.com/2014/08/30/modeling-in-r-with-the-caret-package/
install.packages("caret")
library( caret )
library( randomForest )
data( iris )
set.seed(318)

m = caret::train( Species ~ ., data=iris, method="rf" )
m

#-----------------------------------------------------------------
install.packages("C50")

library(C50)
library("foreign")

#amostras
#dados = read.csv(paste(getwd(),"/CSV/modelo_v8_lv3.csv",sep=""), header=T, quote="")
dados = read.arff("C:/Users/David Monteiro/Google Drive/MASTER/CSV/pcmda23_amostras.csv.arff")
colnames(dados)[8]
ggplot(dados$Class_name)

#noclass
#brutos = read.csv(paste(getwd(),"/CSV/amostras_noclass_v5.csv",sep=""), header=T, quote="")
brutos = read.arff("C:/Users/David Monteiro/Google Drive/MASTER/CSV/pcmda23_noclass.arff")
```

```
#reamostragem aleatoria
#dados = dados[ sample( nrow( dados ) ), ]

xb = brutos[,c(1:94)]
X = dados[,c(1:7,9:94)]
y = dados[,8]

#trainX = X[1:1000,]
#trainy = y[1:1000]
#testX = X[1001:1528,]
#testy = y[1001:1528]

modelo = C50::C5.0( X, y, trials=10)
#modelo = C50::C5.0( X, y)
summary( modelo )
previsao4 = predict(modelo,xb,type = "class")


#-------------------------------------------------------------
library(rgdal)
library(maptools)

mapa = readOGR("C:/Users/David Monteiro/Google Drive/MASTER/ecognition/SHPs/pcmda23.shp",
                layer = "pcmda21c_prd")

tabela = mapa@data

colnames(tabela)

tabela$c506cmd = previsao
tabela$c506nmd = previsao2
tabela$c503cmd = previsao3
tabela$c503nmd = previsao4

mapa@data = tabela

writeOGR(mapa,dsn=paste(getwd(),"/CSV/pcmda21d_prd.shp",sep=""),
                        layer = "pcmda21d_prd",
                        driver = "ESRI Shapefile")

##############################################################
# RENOMEIA CLASSES PARA CONFECCAO DAS MATRIZES DE CONFUSAO #
##############################################################

library(plyr)
```

```
pts$crt6nmd = revalue(pts$crt6nmd, c("aco_brilhante"="C1",
             "aco_ferrugem"="C2",
             "aco_fosco"="C3",
             "acrilico"="C4",
             "argamassa_cimento"="C5",
             "argamassa_cimento_pintado_verde"="C6",
             "argamassa_cimento_pintado_vermelho"="C7",
             "asfalto"="C8",
             "asfalto_pintado_branco"="C9",
             "asfalto_pintado_vermelho"="C10",
             "bloquete"="C11",
             "carvao_coque"="C12",
             "ceramica"="C13",
             "concreto"="C14",
             "concreto_impermeabilizado"="C15",
             "concreto_pintado_amarelo"="C16",
             "concreto_pintado_vermelho"="C17",
             "fibra_vidro"="C18",
             "fibrocimento_amianto"="C19",
             "fibrocimento_crfs"="C20",
             "galvalume_a"="C21",
             "galvalume_b"="C22",
             "galvalume_c"="C23",
             "lago"="C24",
             "lona"="C25",
             "pastilha_esmaltada"="C26",
             "piscina_azulejo"="C27",
             "piscina_vinil"="C28",
             "plastico"="C29",
             "policarbonato"="C30",
             "quartzito_amarelo"="C31",
             "solo_exposto"="C32",
             "sombra"="C33",
             "vegetacao_arborea"="C34",
             "vegetacao_rasteira"="C35",
             "vidro_aramado"="C36"
             ))

temp$Reference = revalue(temp$Reference, c(
#temp$Prediction = revalue(temp$Prediction, c(
  "C1" = "pavimentacao_viaria",
  "C2" = "coberturas_metalicas",
  "C3" = "materiais_mistos",
  "C4" = "coberturas_diversas",
  "C5" = "sombra",
```

```
  "C6" = "pavimentacao_nao_viaria",
  "C7" = "vegetacao_arborea",
  "C8" = "solo_exposto",
  "C9" = "vegetacao_rasteira",
  "C10"= "ceramica",
  "C11"= "lago",
  "C12"= "piscina"
))
#> [1] alpha two   three alpha two
#> Levels: alpha two three

mapvalues(x, from = c("beta", "gamma"), to = c("two", "three"))
#> [1] alpha two   three alpha two
#> Levels: alpha two three

levels(legendas)[1] ="Aço Galvanizado\nBrilhante"
levels(legendas)[2] ="Aço Galvanizado\ncom Ferrugem"
levels(legendas)[3] ="Aço Galvanizado\nFosco"
levels(legendas)[4] ="Acrilico"
levels(legendas)[5] ="Argamassa de Cimento"
levels(legendas)[6] ="Argamassa de Cimento\nPintado de Verde"
levels(legendas)[7] ="Argamassa de Cimento\nPintado de Vermelho"
levels(legendas)[8] ="Asfalto"
levels(legendas)[9] ="Asfalto Pintado\nde Branco"
levels(legendas)[10]="Asfalto Pintado\nde Vermelho"
levels(legendas)[11]="Bloquete"
levels(legendas)[12]="Carvão Coque"
levels(legendas)[13]="Cerâmica"
levels(legendas)[14]="Concreto"
levels(legendas)[15]="Concreto\nImpermeabilizado"
levels(legendas)[16]="Concreto Pintado\nde Amarelo"
levels(legendas)[17]="Concreto Pintado\nde vermelho"
levels(legendas)[18]="Fibra de Vidro"
levels(legendas)[19]="Fibrocimento\ncom Amianto"
levels(legendas)[20]="Fibrocimento\nCRFS"
levels(legendas)[21]="Galvalume A"
levels(legendas)[22]="Galvalume B"
levels(legendas)[23]="Galvalume C"
levels(legendas)[24]="Lago"
levels(legendas)[25]="Lona"
levels(legendas)[26]="Pastilha\nEsmaltada"
levels(legendas)[27]="Piscina de\nAzulejo"
levels(legendas)[28]="Piscina de\nVinil"
levels(legendas)[29]="Plástico"
levels(legendas)[30]="Policarbonato"
```

```
levels(legendas)[31]="Quartzito\nAmarelo"
levels(legendas)[32]="Solo Exposto"
levels(legendas)[33]="Sombra"
levels(legendas)[34]="Não Classificado"
levels(legendas)[35]="Vegetação\nÁrborea"
levels(legendas)[36]="Vegetação\nHerbácea"
levels(legendas)[37]="Vidro Aramado"

legenda = as.factor(c(
  "Aço Galvanizado Brilhante - C1",
  "Aço Galvanizado com Ferrugem - C2",
  "Aço Galvanizado Fosco - C3",
  "Acrilico - C4",
  "Argamassa de Cimento - C5",
  "Argamassa de Cimento Pintado de Verde - C6",
  "Argamassa de Cimento Pintado de Vermelho - C7",
  "Asfalto - C8",
  "Asfalto Pintado de Branco - C9",
  "Asfalto Pintado de Vermelho - C10",
  "Bloquete - C11",
  "Carvão Coque - C12",
  "Cerâmica - C13",
  "Concreto - C14",
  "Concreto Impermeabilizado - C15",
  "Concreto Pintado de Amarelo - C16",
  "Concreto Pintado de vermelho - C17",
  "Fibra de Vidro - C18",
  "Fibrocimento com Amianto - C19",
  "Fibrocimento CRFS - C20",
  "Galvalume A - C21",
  "Galvalume B - C22",
  "Galvalume C - C23",
  "Lago - C24",
  "Lona - C25",
  "Pastilha Esmaltada - C26",
  "Piscina de Azulejo - C27",
  "Piscina de Vinil - C28",
  "Plástico - C29",
  "Policarbonato - C30",
  "Quartzito Amarelo - C31",
  "Solo Exposto - C32",
  "Sombra - C33",
  #"Não Classificado - C34",
  "Vegetação Árborea - C34",
  "Vegetação Herbácea - C35",
```

```
  "Vidro Aramado - C36"
))

legenda2 = as.factor(c(
  "Aço Galvanizado Brilhante",
  "Aço Galvanizado com Ferrugem",
  "Aço Galvanizado Fosco",
  "Acrilico",
  "Argamassa de Cimento",
  "Argamassa de Cimento Pintado de Verde",
  "Argamassa de Cimento Pintado de Vermelho",
  "Asfalto",
  "Asfalto Pintado de Branco",
  "Asfalto Pintado de Vermelho",
  "Bloquete",
  "Carvão Coque",
  "Cerâmica",
  "Concreto",
  "Concreto Impermeabilizado",
  "Concreto Pintado de Amarelo",
  "Concreto Pintado de vermelho",
  "Fibra de Vidro",
  "Fibrocimento com Amianto",
  "Fibrocimento CRFS",
  "Galvalume A",
  "Galvalume B",
  "Galvalume C",
  "Lago",
  "Lona",
  "Pastilha Esmaltada",
  "Piscina de Azulejo",
  "Piscina de Vinil",
  "Plástico",
  "Policarbonato",
  "Quartzito Amarelo",
  "Solo Exposto",
  "Sombra",
  "Vegetação Árborea",
  "Vegetação Herbácea",
  "Vidro Aramado"
))

legenda_12 = as.factor(c(
  "Pavimentação Viária - C1",
  "Cerâmica - C2",
```

```
  "Lago - C3",
  "Piscina - C4",
  "Coberturas Metálicas - C5",
  "Materiais Mistos - C6",
  "Coberturas Diversas - C7",
  "Sombra - C8",
  "Pavimentação não-Viária - C9",
  "Vegetação Árborea - C10",
  "Solo Exposto - C11",
  "Vegetação Herbácea - C12"
))

pts12 = as.factor(c(
  "C1","C2","C3","C4","C5","C6","C7","C8","C9","C10","C11","C12"
))

##########################################################
# LIBS PARA GERACAO DE GRAFICOS DE COORDENADAS PARALELAS: #
##########################################################

test <- read.csv("C:/Users/David/Google Drive/MASTER/ecognition/CSV/amostras_qgis/amostras_lv6_49att_noclass.csv", sep=";")

### DELETE COLUMN
test$Max..diff. <- NULL
test$Length.Thickness <- NULL
test$Thickness <- NULL
test$Brightness <- NULL

write.table(test,
file="C:/Users/David/Google Drive/MASTER/ecognition/CSV/amostras_qgis/r_lv6_49att_noclass.csv",
sep=",",
fileEncoding='UTF-8')

test <- read.csv("C:/Users/David/Google Drive/MASTER/ecognition/CSV/amostras_qgis/r_lv6_49att_noclass.csv", sep=",")

library(MASS)
#parcoord(carvao_coque,col=rainbow(length(test[,1])),var.label=TRUE)
parcoord(carvao_coque,var.label=TRUE)
carvao_coque = amostra_classes[(amostra_classes$Class=="carvao_coque"),]
carvao_coque$Class = NULL

validamapa = readOGR("C:/Users/David Monteiro/Google Drive/MASTER/SHP_FILES/join_pcmda23.shp",layer = "join_pcmda23")

dados = validamapa@data
```

```
carvao_coque = dados[(dados$Clss_nm=="carvao_coque"),]

carvao_coque = carvao_coque[c(1,108:115)]
carvao_coque[2:9] = carvao_coque[2:9]/100
colnames(carvao_coque) = c("FID","B1","B2","B3","B4","B5","B6","B7","B8")
m.carvao = melt(carvao_coque,id="FID")

#colnames(amostra_classes)[1] = "Class"
#colnames(amostra_classes)[2] = "B1"
#colnames(amostra_classes)[3] = "B2"
#colnames(amostra_classes)[4] = "B3"
#colnames(amostra_classes)[5] = "B4"
#colnames(amostra_classes)[6] = "B5"
#colnames(amostra_classes)[7] = "B6"
#colnames(amostra_classes)[8] = "B7"
#colnames(amostra_classes)[9] = "B8"

#windows(family='serif')

png("C:/Users/David Monteiro/Desktop/testeSacha.png", width = 2.8, height = 2, units = 'in', res = 1024)

ggplot(data = m.carvao, aes(x=variable, y= value, group = FID)) +
  theme_bw() +
  geom_line(alpha=0.8) +
  #geom_point() +
  #scale_y_continuous(breaks=c(0,5,10,15,20,25,30,35,40,45,50,55,60,65,70,75,80,85,90,95,100))
  #scale_y_continuous(breaks=seq(0,100,5))
  ylim(0,1) +
  #ylab(expression(~rho)) +
  xlab("") +
  ylab("") +
  ggtitle("carvao_coque") +
  annotate('text', 8, 0.90, label="rho", parse=TRUE, hjust=1, size=6,angle = 354) +
  theme(legend.position="none",
        panel.grid = element_blank())

dev.off()

write.csv(carvao_coque, file="C:/Users/David Monteiro/Google Drive/MASTER/CSV/plot_amostras/carvao_coque.csv",
                        quote = F, row.names = F)

install.packages("cowplot")
require(gridExtra)

require(cowplot)
```

```
grid.arrange(c1, c2, ncol=2)
plot_grid(plot1, plot2, align='h', labels=c('A', 'B'))

##############################################################
# GERA GRAFICOS DE REFLECTANCIA PARA TIPOLOGIAS DAS CLASSES #
# COM BASE NA CURVA ESPECTRAL DAS AMOSTRAS                  #
##############################################################

library(ggplot2)
library(reshape2)

nomecolunas = c("FID","B1","B2","B3","B4","B5","B6","B7","B8")

#loop de 1:36 para criar as variaveis do ggplot
for(i in seq_along(classes)){
  #criar variavel e atribuir dados
  assign(paste(toString(classes[i])),dados[(dados$Clss_nm == toString(classes[i])),])
  #remover demais colunas
  eq = paste(classes[i]," = ",classes[i],"[c(1,108:115)]",sep = "")
  eval(parse(text=eq))
  #mudar escala de 100 para 1
  eq = paste(classes[i],"[2:9]"," = ",classes[i],"[2:9]/100",sep = "")
  eval(parse(text=eq))
  #renomear colunas
  eq = paste("colnames(",classes[i],") = nomecolunas",sep = "")
  eval(parse(text=eq))
  #remover NA
  eq = paste(classes[i]," = ",classes[i],"[complete.cases(",classes[i],"),]",sep = "")
  eval(parse(text=eq))
  #preparar plotagem
  eq = paste("m.",classes[i]," = melt(",classes[i],",id = colnames(",classes[i],"[1]))",sep = "")
  eval(parse(text=eq))
}

#loop de 1:36 para armazenar os plots em arquivos de imagem
for(i in seq_along(classes)){

  caminho = paste("C:/Users/David Monteiro/Google Drive/MASTER/DISSERTACAO/v1_word/teste_p/",classes[i],".png",sep = "")
  #abre dispositivo de plot
  png(caminho, width = 2.8, height = 2, units = 'in', res = 900)

  eq = paste("m.",classes[i],sep = "")

  print(
  ggplot(data = eval(parse(text=eq)), aes(x=variable, y=value, group = FID)) +
```

```
    theme_bw() +
    geom_line(alpha=0.2) +
    ylim(0,1) +
    xlab("") +
    ylab("") +
    ggtitle(legendas[i]) +
    annotate('text', 8, 0.90, label="rho", parse=TRUE, hjust=1, size=6,angle = 354) +
    theme(legend.position="none",panel.grid = element_blank())
  )
  #fecha e salva dispositivo de plot
  dev.off()
}

##############################
# GERA MATRIZES DE CONFUSAO #
##############################

oldvalidapontos = readOGR("C:/Users/David Monteiro/Google Drive/MASTER/SHP_FILES/validapontos_12.shp",layer = "validapontos_12")

# REMOVE COLUNAS COM VEICULOS E PONTOS NAO OBSERVADOS EM CAMPO
pts = validapontos@data
badpts = pts[(pts$obs=="veiculo" | pts$obs=="naobs"),]
pts = pts[!(pts$obs=="veiculo" | pts$obs=="naobs"),]

validapontos@data = pts

# SALVA SHAPE DE PONTOS DESCARTADOS
validapontos@data = badpts
writeOGR(validapontos,
         dsn ="C:/Users/David Monteiro/Google Drive/MASTER/SHP_FILES/lv12_revpts.shp",
         layer = "lv12_revpts",
         driver = "ESRI Shapefile",
         overwrite_layer = T)

mapa = validamapa
#NOTEBOOK
#validamapa = readOGR("C:/Users/David/Google Drive/MASTER/SHP_FILES/lv12_revpts.shp",layer = "lv12_revpts")
#DESKTOP
validamapa = readOGR("C:/Users/David Monteiro/Google Drive/MASTER/SHP_FILES/lv12_revpts.shp",layer = "lv12_revpts")
pts = validamapa@data

# GERA MATRIZ DE CONFUSAO
# LV6 COM MDA
l2_c50cmd.cmatrix = confusionMatrix(pts$c506cmd, pts$obs_lv6)
l2_c45cmd.cmatrix = confusionMatrix(pts$c456cmd, pts$obs_lv6)
```

```
l2_crtcmd.cmatrix = confusionMatrix(pts$crt6cmd, pts$obs_lv6)
# LV6 SEM MDA
l2_c50nmd.cmatrix = confusionMatrix(pts$c506nmd, pts$obs_lv6)
l2_c45nmd.cmatrix = confusionMatrix(pts$c456nmd, pts$obs_lv6)
l2_crtnmd.cmatrix = confusionMatrix(pts$crt6nmd, pts$obs_lv6)
# LV3 COM MDA
l1_c50cmd.cmatrix = confusionMatrix(pts$c503cmd, pts$obs_lv3)
l1_c45cmd.cmatrix = confusionMatrix(pts$c453cmd, pts$obs_lv3)
l1_crtcmd.cmatrix = confusionMatrix(pts$crt3cmd, pts$obs_lv3)
# LV3 SEM MDA
l1_c50nmd.cmatrix = confusionMatrix(pts$c503nmd, pts$obs_lv3)
l1_c45nmd.cmatrix = confusionMatrix(pts$c453nmd, pts$obs_lv3)
l1_crtnmd.cmatrix = confusionMatrix(pts$crt3nmd, pts$obs_lv3)
#confpts = confusionMatrix(pts$pred, pts$obs)

#VAR KAPPA
cm = l2_c50cmd.cmatrix$table # 0.6858477 | 0.0001856738
cm = l2_c45cmd.cmatrix$table # 0.7092944 | 0.0001766550
cm = l2_crtcmd.cmatrix$table # 0.7247778 | 0.0001720917
cm = l2_c50nmd.cmatrix$table # 0.6965478 | 0.0001827731
cm = l2_c45nmd.cmatrix$table # 0.7168415 | 0.0001755872
cm = l2_crtnmd.cmatrix$table # 0.7145529 | 0.0001779491

cm = l1_c50cmd.cmatrix$table # 0.7338436 | 0.0001934546
cm = l1_c45cmd.cmatrix$table # 0.7418370 | 0.0001899773
cm = l1_crtcmd.cmatrix$table # 0.7298324 | 0.0001951512
cm = l1_c50nmd.cmatrix$table # 0.7336602 | 0.0001930517
cm = l1_c45nmd.cmatrix$table # 0.7351210 | 0.0001937642
cm = l1_crtnmd.cmatrix$table # 0.7450042 | 0.0001883820

#Experimento de plotagem

library(ggplot2)
library(reshape2)

melted_l2_c50cmd = melt(l2_c50cmd.cmatrix$table)
melted_l2_c45cmd = melt(l2_c45cmd.cmatrix$table)
melted_l2_crtcmd = melt(l2_crtcmd.cmatrix$table)

melted_l2_c50nmd = melt(l2_c50nmd.cmatrix$table)
melted_l2_c45nmd = melt(l2_c45nmd.cmatrix$table)
melted_l2_crtnmd = melt(l2_crtnmd.cmatrix$table)

melted_l1_c50cmd = melt(l1_c50cmd.cmatrix$table)
melted_l1_c45cmd = melt(l1_c45cmd.cmatrix$table)
```

```
melted_l1_crtcmd = melt(l1_crtcmd.cmatrix$table)

melted_l1_c50nmd = melt(l1_c50nmd.cmatrix$table)
melted_l1_c45nmd = melt(l1_c45nmd.cmatrix$table)
melted_l1_crtnmd = melt(l1_crtnmd.cmatrix$table)

l2_c50cmd.cmatrix$overall
l2_c45cmd.cmatrix$overall
l2_crtcmd.cmatrix$overall

l2_c50nmd.cmatrix$overall
l2_c45nmd.cmatrix$overall
l2_crtnmd.cmatrix$overall

l1_c50cmd.cmatrix$overall
l1_c45cmd.cmatrix$overall
l1_crtcmd.cmatrix$overall

l1_c50nmd.cmatrix$overall
l1_c45nmd.cmatrix$overall
l1_crtnmd.cmatrix$overall

#head(melted_c506cmd)

temp = melted_l2_c45nmd
#temp = melted_l2_c45cmd
#temp = melted_l2_crtcmd
#temp = melted_l2_c50nmd
#temp = melted_l2_c45nmd
#temp = melted_l2_crtnmd

#temp = melted_l1_c50cmd
#temp = melted_l1_c45cmd
#temp = melted_l1_crtcmd
#temp = melted_l1_c50nmd
#temp = melted_l1_c45nmd
#temp = melted_l1_crtnmd

View(melted_l1_c50cmd)
View(temp)

#36 CLASSES
temp$Prediction = factor(temp$Prediction,levels(temp$Prediction)[unique(pts$c456cmd)])
temp$Reference = factor(temp$Reference,levels(temp$Reference)[unique(pts$c456cmd)])
```

```
#12 CLASSES
#ts_pts = pts[with(pts, order(obs_lv3)), ]

#temp$Prediction = factor(temp$Prediction,levels(temp$Prediction)[unique(pts12)])
#temp$Reference = factor(temp$Reference,levels(temp$Reference)[unique(pts12)])

#REMOVE ZEROS DA MATRIZ DE CONFUSAO
temp$value[temp$value == 0] <- NA

for(class in levels(temp$Prediction)){
  temp$tot[temp$Prediction == class] = sum(temp$value[temp$Prediction == class], na.rm = T)
}

rotulolegenda = expression("italic(Kappa)\nclasse")
temp$perc = temp$value/temp$tot



  ggplot(temp, aes(x = Prediction, y = Reference)) +
  geom_tile(aes(fill = perc*1)) +
  scale_fill_continuous(name=expression(italic("Kappa\nclasse")),low="white", high="#ff3333",
                        guide="colorbar", na.value=alpha("white",0.1)) +
  scale_y_discrete(labels = legenda) +
  #scale_x_discrete(labels = pts12) +
  theme_bw() +
  theme(axis.text.x = element_text(angle = 90, vjust = .5, hjust = 1)#,legend.position = "none"
        )+
  geom_text(aes(x = Prediction, y = Reference, label = value), size=3) +
  #scale_y_reverse()+
  ggtitle("C4.5 Nível II sem MDA")+
  xlab("")+
  ylab("")

#####################################################
# Adaptado de Pompeu e Marujo:                      #
# GERA VARICANCIA DO KAPPA DAS IMAGENS CLASSIFICADAS #
#####################################################

  sum_diag=0

tabela=matrix(0,ncol=ncol(cm)+1, nrow=nrow(cm)+1)

for (i in 1:nrow(cm)){
  for (j in 1:ncol(cm)) {
    tabela[i,j]=cm[i,j]
```

```
    if (i==j){
      sum_diag=sum_diag+cm[i,j]
    }
  }
}
for (i in 1:nrow(tabela)){
  tabela[i,ncol(tabela)] = sum(tabela[i,])
  tabela[nrow(tabela),i]=sum(tabela[,i])
}
tabela[nrow(tabela),ncol(tabela)]=tabela[nrow(tabela),ncol(tabela)]/2
# EXATIDÃO GLOBAL
OA = sum_diag/sum(cm)
# PRODUCER'S ACCURACY
PA = array(0,nrow(cm))
# USER'S ACCURACY
UA = array(0,nrow(cm))

for (i in 1:nrow(cm)){
  PA[i]=tabela[i,i]/tabela[nrow(tabela),i]
  UA[i]=tabela[i,i]/tabela[i,ncol(tabela)]
}

CE = 1-UA #Comission Error
OE = 1-PA #Omission Error

sum_diag_rand=0

rand_matrix=matrix(0,ncol=ncol(cm)+1, nrow=nrow(cm)+1)

for (i in 1:nrow(cm)){
  for (j in 1:ncol(cm)) { #com dois for ele vai percorrer a tabela toda
    rand_matrix[i,j]=tabela[i,ncol(tabela)]*tabela[nrow(tabela),j]/tabela[nrow(tabela),ncol(tabela)]
  }
  sum_diag_rand=sum_diag_rand+rand_matrix[i,i]
}
tot_rand=sum_diag_rand/sum(cm)

# KAPPA
k = (OA - tot_rand)/(1-tot_rand)
k # valor Kappa
# Termos necessários para cálculo da variância do Kappa
theta3<-0
for (i in 1:nrow(cm)){
  theta3 <- theta3 + tabela[i,i]*(tabela[i,ncol(tabela)]+
                                  tabela[nrow(tabela),i])/(sum(cm)^2)
```

```
}
theta4<-0
for (i in 1:nrow(cm)){
  for (j in 1:ncol(cm)) {
    theta4=theta4 + tabela[i,j]*((tabela[i,ncol(tabela)]+
                                  tabela[nrow(tabela),j])^2)/(sum(cm)^3)
  }
}
theta1=OA
theta2=tot_rand
a=((theta1)*(1-theta1)/(1-theta2)^2)
b=(2*(1-theta1)*((2*theta1*theta2-theta3)/(1-theta2)^3))
c=((1-theta1)^2)*(theta4-4*(theta2^2))/(1-theta2)^4
# VARIÂNCIA DO KAPPA
varK=(a+b+c)/sum(cm)
varK # valor da variância do kappa

teste=list("Kappa"=k,"Variancia"=varK)
read.csv(teste)

##############################################
# Adaptado de Pompeu e Marujo: GERA TESTE-Z #
##############################################

#LV 2
mat = matrix(c(0.6858477,
               0.7092944,
               0.7247778,
               0.6965478,
               0.7168415,
               0.7145529,

               # a partir daqui são os valores de variância do kappa em ordem
               0.0001856738,
               0.000176655,
               0.0001720917,
               0.0001827731,
               0.0001755872,
               0.0001779491),
             nrow=6, # QUANTIDADE DE CLASSIFICAÇÕES
             ncol=2)

#LV 1
mat = matrix(c(0.7338436,
               0.7418370,
```

```
                0.7298324,
                0.7336602,
                0.7351210,
                0.7450042,

                # a partir daqui são os valores de variância do kappa em ordem
                0.0001934546,
                0.0001899773,
                0.0001951512,
                0.0001930517,
                0.0001937642,
                0.0001883820),
              nrow=6, # QUANTIDADE DE CLASSIFICAÇÕES
              ncol=2)

colnames(mat)=c("Kappa", "Var_k")
rownames(mat)=c("1","2","3","4","5","6")
mat

# Faz uma matriz com os números dos modelos nas linhas e colunas

tabela=matrix(0,ncol=nrow(mat), nrow=nrow(mat))
colnames(tabela)=c("1","2","3","4","5","6")
rownames(tabela)=c("1","2","3","4","5","6")

# cálculo dos valores Z

for (i in 1:nrow(mat)){
  for(j in 1:nrow(mat)){
    z=((mat[i,1])-(mat[j,1]))/(sqrt((mat[i,2])+(mat[j,2])))
    tabela[j,i]=z
  }
}
View(tabela) # Resultado

#######################################################
# PLOTAGEM DOS DIAGRAMAS DE ESPALHAMENTO DAS AMOSTRAS #
# PARA ATRIBUTOS DE INTERESSE                         #
#######################################################

mapa = readOGR("C:/Users/Cliente/Downloads/DISSERTACAO_GUARA/SHP/join_pcmda23.shp",layer = "join_pcmda23")
colnames(mapa@data)
#oldt = tabela
tabela = mapa@data
plot(tabela$BRT,tabela$NDBI)
```

```
library(foreign)
#amostras = read.arff("C:/Users/Cliente/Downloads/DISSERTACAO_GUARA/pcmda23_amostras.csv.arff")
amostras = read.arff("C:/Users/David Monteiro/Google Drive/MASTER/CSV/pcmda23_amostras.csv.arff")

plot(amostras$BRT,amostras$NDBI)

#HEXA DAS CORES DE CLASSES DA LEGENDA
myColors = c(
  "#FFF18F",
  "#FFBC8F",
  "#AED8F7",
  "#AF87FF",
  "#BA9B9B",
  "#9DAE31",
  "#D16B6F",
  "#4E4E4E",
  "#DCDCDC",
  "#FF0000",
  "#CE9DBA",
  "#87645B",
  "#FF9C4B",
  "#A5A5D0",
  "#7D889B",
  "#FFC639",
  "#D15E51",
  "#FFD1FF",
  "#AA475D",
  "#AA7581",
  "#FFFFFF",
  "#FFCCCC",
  "#B9CCFF",
  "#5788B5",
  "#FF68A8",
  "#00FF00",
  "#017CFF",
  "#A2D2FF",
  "#00FFC4",
  "#AEFFED",
  "#CAB06E",
  "#977050",
  "#000000",
  "#677F4D",
  "#A5C187",
  "#639796"
```

```
)

#Create a custom color scale
classes = levels(legenda2)
names(myColors) = levels(classes)
colScale = scale_colour_manual(name = "Classes",values = myColors)

library(ggplot2)

qplot(amostras$BRT,amostras$NDBI, colour = amostras$Class_name,ylim=c(0,0.5))
qplot(amostras$Mean_mda,amostras$NDBI, colour = colScale)

p = ggplot(dat,aes(amostras.NDVI1,amostras.NDVI2,colour = amostras.Class_name)) +
  geom_point()+
  #LEGENDA DENTRO DO PLOT
  #theme(legend.justification=c(1,1), legend.position=c(1,1)) + guides(col = guide_legend(nrow = 9)) +
  xlab("NDVI 1") +
  ylab("NDVI 2") +
  #annotate("rect", xmin=-2, xmax=2, ymin=-2, ymax=2, linetype=1,color="red",fill=NA) +
  #ggtitle("Scatterplot de Distribuição das Amostras") +
  xlim(-1,1) +
  ylim(-1,1)
  #theme(legend.position="none")


p1 = p + colScale
plot(p1)

#dat <- data.frame(amostras$BRT,amostras$NDBI,amostras$Class_name)
dat = NULL
dat = data.frame(amostras$Class_name,
                 amostras$Mean_mda,
                 amostras$Border_ind,
                 amostras$BRT,
                 amostras$Asymmetry,
                 amostras$NDBI,
                 amostras$Mean_pc7,
                 amostras$BRTpBoR,
                 amostras$NDSI1,
                 amostras$Min_pixel_.3,
                 amostras$Mean_pc8,
                 amostras$Mean_pc4,
                 amostras$WVBI,
                 amostras$Mean_pc3,
                 amostras$NDSI2,
```

```
                    amostras$NDVI1,
                    amostras$NDVI2,
                    amostras$NDVI3,
                    amostras$Mean_pc6,
                    amostras$WVSI,
                    amostras$Max_pixel_.2,
                    amostras$BRTpBoRE
                    )

levels(dat$amostras.Class_name) = legenda2

# MATRIZ DE CORRELAÇÃO
intall.packages("performanceAnalytics")
library(PerformanceAnalytics)
colnames(dat)
dat$amostras.Class_name = NULL
cor(dat)
chart.Correlation(dat,histogram=F,lowess=F)

# RENOMEIA OS ROTULOS DA DIAGONAL
colnames(dat)[1] = "Mean_mda"
colnames(dat)[2] = "Border_ind"
colnames(dat)[3] = "BRT"
colnames(dat)[4] = "Asymmetry"
colnames(dat)[5] = "NDBI"
colnames(dat)[6] = "Mean_PC7"
colnames(dat)[7] = "BRT+B/R"
colnames(dat)[8] = "NDSI1"
colnames(dat)[9] = "Min_px_PC6"
colnames(dat)[10] = "Mean_PC8"
colnames(dat)[11] = "Mean_PC4"
colnames(dat)[12] = "WVBI"
colnames(dat)[13] = "Mean_PC3"
colnames(dat)[14] = "NDSI2"
colnames(dat)[15] = "NDVI1"
colnames(dat)[16] = "NDVI2"
colnames(dat)[17] = "NDVI3"
colnames(dat)[18] = "Mean_PC6"
colnames(dat)[19] = "WVSI"
colnames(dat)[20] = "Max_px_MDA"
colnames(dat)[21] = "BRT+B/RE"
```